From the Rational to the Moral: A Utilitarian Journey

Abstract. Assume that what we call Rational Utilitarianism (RU) is true: that (very roughly) we have most reason to do what produces most well-being. Does this imply that a moral counterpart theory, what we call Moral Utilitarianism (MU), according to which (again, very roughly) what is morally right to do is what produces most well-being, is also true? It is this Question (Q) that occupies us in this paper – a question that we intend to answer in the affirmative. We start by explaining why we take RU to be a promising theory of practical reason. Then we turn to Q, dismissing arguments that attempt to derive moral from rational utilitarianism based on the idea that morality has rational authority. After this we present our argument for an affirmative answer to Q and discuss two objections. Next we elaborate on what we call the utilitarian landscape: the overall picture we get when considering the implications of RU for all normative areas. Finally, we conclude by reflecting on our findings.

1. Introduction

In his defence of act utilitarianism against rule utilitarianism (which he calls 'restricted utilitarianism'), J.J.C. Smart (1956: 353) writes:

> The restricted utilitarian might say that he is talking only of *morality* […] I would reply that as a philosopher I conceive of ethics as the study of how it would be *most rational* to act. If my opponent wishes to restrict the word 'morality' to a narrower use he can have the word. The fundamental question is the question of rationality *in general*.[1]

Similarly, Roger Crisp (2006: 1-2) summarizes the first chapter of his *Reasons and the Good* in this way:

> The book opens with the suggestion that a – perhaps *the* – fundamental question in philosophical normative ethics concerns what each of us has reason to do. That leads immediately into the issue of whether we have *moral* reasons to act, in the sense of ultimate or non-derivative reasons the correct description of which makes ineliminable use of the moral concepts – right, wrong, good, bad, virtuous, kind, cruel, and so on. Drawing an analogy between morality and the law as social phenomena, I claim that there is a strong case for thinking that morality in itself provides no such reasons, though there may well be derivative reasons for doing what some actual morality or other prescribes. If I am right, then the correct theory of reasons for action should be stated without using the moral concepts.

---

[1] Likewise, Mackie (1977: 106) prominently distinguished between morality in the broad sense, understood as 'a general all-inclusive theory of conduct', and morality in the narrow sense as a system of constraints 'whose central task is to protect the interests of persons other than the agent and which present themselves to an agent as checks on his natural inclinations or spontaneous tendencies to act'.

Let us assume that, as Smart and (with some qualifications) Crisp suggest, a version of act-utilitarianism is the best theory of practical reason, i.e., the best theory about ultimate (i.e., non-derivative) reasons for action, and that this version of utilitarianism can be stated without using moral concepts. We are mainly interested in the following

> ***Question (Q)****:* If a version of act-utilitarianism – call it *rational utilitarianism* – is the best theory of practical reason (in the sense just specified), would this imply the claim that a moral counterpart theory – call it *moral utilitarianism* – is the best moral theory (the best theory about moral reasons and moral obligations)? Put simply, does rational utilitarianism imply moral utilitarianism?

Let us state the theories under consideration more fully.

> ***Rational Utilitarianism (RU)****:* For any alternative acts *a* and *b,* there is more reason to perform *a* than *b* – and, accordingly, less reason to perform *b* than *a* – to the degree that, and because, the outcome of *a* contains more well-being than the outcome of *b*.

> ***Moral Utilitarianism (MU)****:* An act is morally right if and only if (iff), and because, none of its alternatives leads to an outcome with more well-being. An act is morally wrong iff it is not morally right. An act is morally obligatory iff none of its alternatives is morally right.

We argue for an affirmative answer to Q: RU implies MU. The implication under consideration here is *broad* implication, meaning that MU deductively follows from RU together with other plausible propositions.

Q is an interesting question for several reasons. First of all, assume that RU can be shown to be the best theory of practical reason. It would then be natural to ask what the ramifications of this would be for morality. Perhaps, rather than accepting MU, it would be justified not to accept any moral theory, as the quotes by Smart and Crisp might be taken to suggest. Or perhaps, even if RU is true, the best moral theory might be different from MU.

Secondly, given that MU is at odds with widespread moral intuitions, an affirmative answer to Q would pave the way for a new kind of argument for MU, an argument that would be immune to objections based on moral intuitions. Together with an argument for RU an affirmative answer to Q would yield what is sometimes called a *foundationalist* argument for MU.[2]

Thirdly, a convincing answer to Q promises to improve our understanding of the relation between morality and practical reason. Fourthly, a convincing answer would reveal considerations that could be used to answer similar questions regarding other theories, such as: does rational egoism lead to ethical egoism?

There is a final reason why Q is an interesting question, a reason that, albeit important, can easily be overlooked and requires a bit of explanation. We take the main approach in moral or normative epistemology to be the method of wide reflective equilibrium, and we shall assume this approach in this paper. Moreover, we make standard assumptions about criteria of adequacy for moral and other normative theories, such as that they should be clear and (other things being equal) simple, have explanatory power and scope, and be internally coherent. Despite its problems (which we will address in our paper), MU scores points with regard to these criteria of adequacy and explains (or justifies) many of our considered moral judgements (cf. Bykvist 2010: 22-5). And the same can be said about RU as a theory of practical reason. Now, importantly, whatever can be said in support of RU, if RU implies MU, then this may

---

[2] For more on this, see Jamieson (1993).

lend *additional* credence to RU. For RU will then indirectly, via moral utilitarianism, yield explanations for many considered *moral* judgements and thus score points in terms of explanatory scope. This would be an advantage of RU compared to other theories of practical reason insofar as these other theories do not imply moral counterpart theories and thus have less explanatory scope.

If RU has the appealing feature of implying MU and thus yielding explanations of morality, it is natural to ask if RU also implies counterpart theories in other normative areas, thus scoring even more points in terms of explanatory scope. We shall argue that this is the case. Most notably, RU has the resources to explain prudence and plausibly also politics, and other fields.

While Q is a conditional question, it is fair to ask what reasons there are for thinking that RU is the best theory of practical reason. Also, one might worry that RU is less plausible as a theory of practical reason than MU is as a moral theory. For these reasons, our investigation will begin by outlining an argument in support of RU. Obviously, since our focus in this paper is on Q, the argument for RU will be neither conclusive nor, in its main parts, original. Our aim is to show that RU is sufficiently promising as a theory of practical reason to warrant the investigation of Q.

Our journey, then, begins with an argument for RU, then takes us from RU to MU and ends in other normative areas. It is, perhaps surprisingly, also in a sense a reverse journey back in time, namely from Henry Sidgwick to John Stuart Mill. For even though we are not concerned with exegetical questions, the central parts of the argument for RU stem from Derek Parfit's work on Sidgwick's analysis of 'the profoundest problem in ethics'. Connecting RU

with other normative areas, on the other hand, arguably began life as a focus of philosophical study with Mill's 'art of life'.[3]

The plan of the paper is as follows. We start by explaining why we take RU to be a promising theory of practical reason (section 2). Then we turn to Q, dismissing arguments that attempt to derive moral from rational utilitarianism based on the idea that morality has rational authority (section 3). Then we present our argument for an affirmative answer to Q (section 4) and discuss two objections (sections 5 to 7). After that, we elaborate on what we call the utilitarian landscape: the overall picture we get when considering the implications of RU for all normative areas (section 8). Finally, we conclude by reflecting on our findings (section 9).

2. Rational Utilitarianism

The introduction presented several reasons for thinking that Q is an interesting question. Still, one might worry that RU is – at least prima facie – even more implausible as a theory of practical reason than MU is as a moral theory. The worry could possibly be substantiated by arguing that the most pressing interpretation of the so-called over-demandingness objection against MU focuses on practical reason, claiming that MU fails because we lack sufficient reason to do what is morally required according to MU.[4] Moreover, while it seems hard to deny that morality is, at least to some extent, characterized by impartiality, which supports MU, practical reason arguably seems to be in some sense agent-relative, focusing on what is good for the agent or what the agent cares about.

The worry threatens to undermine the relevance of answering Q. If arguing for MU is more promising than arguing for RU, then it does not seem to matter much whether MU can

---

[3] Two caveats. First, Mill seems to have understood the most fundamental principle underlying all of normativity, the principle of utility, in purely axiological terms, whereas RU is a deontic principle. Second, Bentham of course also famously applied utilitarianism to several normative domains, though he did not spend as much time elaborating on this idea as did Mill. REF

[4] REF (I suppose Portmore 2011 and Dorsey 2016)

be based on RU. In this section, therefore, we shall outline why one might think that a case for RU is more promising than a case for MU.[5] Needless to say, it would go beyond the scope of this paper to provide a conclusive argument for RU. Our goal is rather to motivate RU enough to make our investigation seem sensible. Therefore, what we will do is to provide a sketch of a possible argument for RU comprising of three main steps: the first step rejects Rational Egoism, the second step rejects Instrumentalism, the third step rejects views that attribute normative significance to the self. The first two steps are not original, building mainly on the work of Parfit. However, we need them to provide context and introduce material we use later. The third step is an original contribution. We emerge with the claim that only considerations of welfare ground reasons and they do so in a person-neutral, impartial way. Hence RU.

Our argument's first step starts with the claim that practical reason seems to be in some sense agent-relative, focusing on what is good for the agent or what the agent cares about. The most straightforward theories based on that claim seem to be Rational Egoism and Instrumentalism:

Rational Egoism: For any alternative acts *a* and *b,* the agent has more reason to perform *a* than *b* – and, accordingly, less reason to perform *b* than *a* – to the degree that, and only because, the outcome of *a* is overall better for the agent than the outcome of *b*.

Instrumentalism: For any alternative acts *a* and *b,* the agent has more reason to perform *a* than *b* – and, accordingly, less reason to perform *b* than *a* – to the degree that, and only because, the agent's (rational) desires would overall be better satisfied in the outcome of *a* than in the outcome of *b*.

---

[5] Of course, we do not mean here to include the indirect case for MU (the attempt to derive MU from RU) that we defend in the next sections. Also, strictly speaking, to substantiate the claim, one would also have to consider the difficulties arguments for MU are confronted with. These difficulties will be presented in section 5. Still, our aim here can only be very modest as we cannot discuss elaborate attempts to argue for MU.

As noted in the introduction, the discussion of RU versus Rational Egoism is a major topic of Sidgwick, which has been crucially advanced by Derek Parfit in more recent times.[6] Rational Egoism corresponds to what Parfit calls 'the Self-interest Theory', which gives each person 'one supremely rational ultimate aim: that his life go, for him, as well as possible'. Parfit compares the Self-Interest Theory with 'Neutralism', which corresponds to RU, and also with 'the Present-aim Theory', which corresponds to Instrumentalism and claims that every agent ought to do what will best achieve his or her present aims. Parfit argues:

> In rejecting Neutralism, a Self-interest Theorist must claim that a reason may have force only for the agent. But the grounds for this claim support a further claim. If a reason can have force only for the agent, it can have force for the agent only at the time of acting. The Self-interest Theorist must reject this claim. He must attack the notion of a time-relative reason. But arguments to show that reasons must be temporally neutral, thus refuting the Present-aim Theory, may also show that reasons must be neutral between different people, thus refuting the Self-interest Theory.[7]

In combining temporal neutrality with personal relativity, Rational Egoism seems unstable. RU and Instrumentalism each seem more coherent, RU because it is both temporally and personally neutral and Instrumentalism because it is both temporally and personally relative. While Parfit's conclusion – that we should reject the Self-Interest Theory, and hence Rational Egoism – can be contested, Parfit's argument is strong enough for our purposes. It sufficiently motivates the rejection of Rational Egoism, a main competitor of RU. This is our first step in arguing for RU.

---

[6] Our case for RU begins similarly to de Lazari-Radek and Singer (2014: ch. 6), who provide a more extensive discussion but at some point turn their focus on evolutionary debunking arguments. For a strong criticism of such arguments in support of RU, which de Lazari-Radek and Singer discuss, see Kahane (2011).
[7] Parfit (1984): 144.

As for the rejection of Instrumentalism, the second step of our argument for RU, we can again turn to Parfit. First, take a version of Instrumentalism that does not add any rational constraints on what counts as the agent's relevant desires, implying that whatever the agent currently happens to desire gives the agent reasons for action. Obviously, agents can desire things that they should not desire (and fail to desire things that they should desire):

> Suppose, for example, that we must choose which of two possible ordeals we shall later undergo. If one of these ordeals would be much more painful, this fact gives us a strong reason to prefer the other. If we have no other relevant reason, it would be contrary to reason, and in this way irrational, knowingly to prefer the more painful ordeal.[8]

This gives us strong reason to reject a simple version of Instrumentalism. Let us now add rational constraints. We can distinguish between substantive and procedural constraints.[9] Substantive constraints make the rationality of an agent (not) desiring an object dependent on whether the object is worth desiring. Instrumentalism is not compatible with substantive constraints: On the assumption of substantive constraints, our reasons eventually do not depend on what we desire but on what we ought or have reason to desire. By contrast, deliberative constraints make the rationality of an agent (not) desiring an object dependent on whether the agent would desire the object after fully informed and procedurally rational deliberation. Now, even if Instrumentalism can convincingly be combined with deliberative constraints, there appear to be strong reasons for rejecting this kind of Instrumentalism – *viz.*, Parfit's Agony Argument, which is based on the following case:

---

[8] Parfit (2011a): 56.
[9] Cf. Parfit (2011a): 62–3.

I know that some future event would cause me to have some period of agony. Even after ideal deliberation, I have no desire to avoid this agony. Nor do I have any other desire or aim whose fulfilment would be prevented either by this agony, or by my having no desire to avoid this agony.[10]

As Parfit points out: 'We all have a reason to want to avoid, and try to avoid, all future agony'.[11] But Instrumentalism, even when combined with procedural rational constraints ('ideal deliberation'), implies that the agent featuring in the case does not have such reason. Hence, Instrumentalism is false.

Again, not everyone is convinced that Parfit's criticism of Instrumentalism is conclusive.[12] For our purposes, however, we have heard enough to motivate the rejection of Instrumentalism. With Rational Egoism and Instrumentalism both being off the table for reasons that pose no problem for RU, we can presently turn to other theories of practical reason. Here, of course, we part company with Parfit, who eventually affirms that we have '*personal* and *partial* reasons to care about the well-being of ourselves and those to whom we have close ties' as well as '*impartial* reasons to care about everyone's well-being'.[13] But we think we have good grounds to do so: for Parfit's claim falls short of a *theory* of practical reason. Parfit and his followers seem to hold that we have reached normative bedrock with Parfit's claim: that, apart from investigations into the relative strengths of personal, partial, and impartial reasons compared to each other, no further theory is to be had. But that would be an unsatisfying result. For we would not only want to know *how* strong personal, partial, and impartial reasons are relative to each other, but also *why*. To answer the why-question, we need a theory. Therefore, Parfit's claim cannot yet be normative bedrock.

---

[10] Parfit (2011a): 73–4.
[11] Parfit (2011a): 76.
[12] Lin etc. (Sobel), Schroeder
[13] Parfit (2011a): 136.

With Rational Egoism and Instrumentalism out of the way, we can turn to the third step of our argument for RU: the rejection of remaining theories of practical reason. Of course, we cannot address all proposals but need to focus on main contenders. Here we take our clue from Dale Dorsey (2016: 174) who observes that "[A] number of recent (and not so recent) works in the metaethics of practical rationality have suggested that features of a person's character, commitments, and projects have important normative upshots."[14] Following Dorsey, let us call these theories different versions of the Normative Significance of the Self account (NSS). One motive of NSS theorists is to give morality (what they perceive to be) its due place in the realm of practical reason while simultaneously acknowledging individual discretion and peculiarities.

As with the previous two steps, we do not offer a knock-down argument against the diverse versions of NSS. The versions differ too much and too manifold are the issues that we would have to address. However, we shall outline an argument that we hope will suffice to motivate enough scepticism about NSS to conclude that RU is still a candidate with prospects good enough to warrant investigating Q.

The claim that features of a person's character, commitments, and projects have important normative upshots can be spelled out in different ways: either these aspects of the self (or some of them) themselves ground reasons or they impact – condition or modify in some way – reasons that are grounded in considerations other than the self. Let us consider each possibility in turn.

Let us first consider the sources of reasons, beginning with single-person cases. The grounding possibility, we submit, loses appeal when compared to considerations of well-being. On the one hand, it seems plausible that many cases in which a person's character, commitments, or projects seem to provide reasons can equally well be accounted for in terms

---

[14] See, for example, Chang (2013), Manne (2013), Korsgaard (1997), Williams (1976), Buss (2006), Portmore (2007), Chalhoun (2009), Dorsey (2016). It is irrelevant for our purposes if all these authors would agree with being grouped together or with Dorsey's characterization of their views.

of well-being. Surely, character, commitments, and projects are often sources or even constitutive parts of well-being. On the other hand, it seems that in most single-person cases in which reasons (allegedly) associated with the self come apart from reasons grounded in well-being, the latter seem to prevail. Picture an action that would fall out of character or be at odds with the agent's commitments or projects but assume that the action would be, all things considered, best for the agent – and recall that nobody else is concerned. It seems hard to believe that the agent in this constellation could be rationally permitted not to perform the action. But this suggests that, in single-person cases, considerations of well-being are decisive, leaving no significant role to aspects of the self as sources of reasons.

Of course, a lot depends on how we understand 'well-being'. We do not deny, to take an example from the literature, that Freud had decisive reason not to take a painkiller that would have made it impossible for him to continue with his studies.[15] We only say that *if* taking the painkiller would have been on balance better for Freud according to the best theory of well-being and assuming, hypothetically, that nobody else would have been affected by Freud's decision, then Freud would have had decisive reason to take the painkiller.

If we now turn to multiple-person cases, why should this suddenly change? Of course, aspects of the self will be invoked by proponents of NSS to bear on the relative weight of self-regarding and other-regarding reasons. But currently we are considering the proposal that aspects of the self *ground* reasons, rather than impacting them. If all the reasons in single-person cases eventually seem to be grounded in the agent's well-being, then it is natural to assume that all reasons in multiple-person cases are also grounded in well-being.

The upshot is Welfarism About Reasons (WAR): the claim that all reasons are provided by considerations of well-being. But, of course, these reasons could be impacted by aspects of the self. This brings us to the second variant of NSS.

---

[15] Griffin (1986: 8).

Consider a single-person case with two options, one of them being on balance better for the agent than the other. It seems uncredible that aspects of the self could nonetheless somehow make it the case that the agent has more reason to perform the second action. Of course, at this point proponents of NSS will disagree and say that even if morality does not play any role in single-person cases, the agent could still have reasons that are neither grounded in well-being nor in their character, commitments or projects. Aesthetic, religious, financial reasons and possibly even professional and legal reasons come to mind.

We admit that one can speak of aesthetic, religious, financial and other reasons as considerations that, from certain standpoints, speak for or against actions.[16] Of course, such reasons may matter indirectly, insofar as they are relevantly related to well-being. To rebut WAR in single-person cases, however, proponents of NSS would have to claim that such reasons are *normative* reasons, i.e., reasons that normatively matter. They might either matter by themselves or aspects of the self could make them matter.

We find it hard to believe that an agent in a single-person case can ever have decisive reason – aesthetic, religious, or otherwise – not to perform an action that is, on balance, best for the agent. But this leaves such reasons quite irrelevant, at best, and it is natural to suspect that such reasons, by themselves, do not matter at all. Can aspects of the self make it the case that aesthetic, religious etc. reasons matter? Again, we must be careful not to confuse such aspects either with considerations of well-being or with the agent's desires. Our discussion of Instrumentalism suggests that desires prima facie, though not *secunda facie*, do a good job at explaining reasons. But if desires and well-being are off the table, we see little reason to believe that a person's character, commitments, or projects can make aesthetic, religious etc. reasons matter. If an agent is committed to a cause or pursues aesthetic projects, say, but would be

---

[16] Dorsey (2016: ch. 1).

better off, on balance, leading a different life, and others are not affected, then the agent seems best advised to change their life.

Things appear to be different, at first glance, in multi-person cases. If WAR is true, as our discussion of Instrumentalism and NSS so far suggests, and if the best theory of practical reasons is person-neutral, as our discussion of Rational Egoism (conjoined with our discussion of Instrumentalism) suggests, then the upshot seems to be RU (or at least some closely related position). But the result would be very demanding and not leave much discretion to the agent. It is, therefore, tempting to say, with NSS theorists, that person-neutrality is really a feature of morality and that moral reasons can somehow be attenuated by aspects of the self.

We do not deny that NSS can be so construed as to leave us with a picture of the relevance of other-regarding reasons that coheres with the pre-theoretical intuitions of many people in terms of demandingness. But at this point, this move would be ad hoc. Our discussion so far suggested that RU is the best theory of practical reason, combining both WAR and person-neutrality. We doubt that pre-theoretical intuitions that suggest discounting the relevance of other individuals' well-being for the agent justify the rejection of RU at this dialectical point on behalf of NSS. On the one hand, NSS, given the rejection of Instrumentalism and the elaborations in the foregoing paragraphs, seems to have little going for it apart from delivering the desired result regarding demandingness. On the other hand, it is not farfetched to suspect that the pre-theoretical intuition NSS seems to accommodate is rooted, at least in part, in self-serving biases and philosophical positions that, as our previous discussion suggests, seem inferior to RU, most notably Instrumentalism and Rational Egoism.

With this third step our argument for RU is complete. We admit that the argument was painted with broad strokes. Still, we hope that we have said enough to show that RU is a promising theory of practical reason. Let us then continue by addressing Q.

3. The Rational Authority of Morality

One might think that RU trivially leads to MU because RU somehow conveys rational authority to MU. How could this idea be spelled out in detail? Three suggestions come to mind. But they all fail.

The first suggestion is to appeal to the position that moral requirements are requirements of practical reason in order to derive MU from RU. Those who hold the position accept:

> *Moral Rationalism*: If a person has a moral obligation to φ, then the person has decisive reason to φ.[17]

However, Moral Rationalism does not help in our case. This is because moral rationalism states a necessary condition for moral requirements, but what would be needed is a *sufficient* condition. After all, what we are asking is if a rational requirement – an action's being more strongly supported by reasons, according to RU, than any of its alternatives – thereby also qualifies as a moral requirement. (Apart from this problem, it is worth mentioning that Moral Rationalism is a controversial position – we will return to this point later.)

To derive MU from RU, one would need to claim that rational requirements are or provide moral requirements. This brings us to the second suggestion, which appeals to:

> *Inverse Moral Rationalism*: If a person has decisive reason to φ, then the person has a moral obligation to φ.

---

[17] For discussions see Brink (1997), Darwell (2006), Dorsey (2016), Portmore (2011, 2021), Stroud (1998).

Inverse Moral Rationalism would do the trick: actions that are rationally required – in the sense of being more strongly supported by reasons than any alternative according to rational utilitarianism – would also be morally required. However, Inverse Moral Rationalism is problematic. Ordinary moral discourse recognises that a person may have decisive prudential reason to choose watching Star Trek over watching Star Wars without having any moral obligation to this effect. Put differently, according to common-sense morality, not all rational requirements are also moral requirements.

Inverse Moral Rationalism is implausible for a second reason. Assume that Rational Egoism is the best theory of practical reason. Then Inverse Moral Rationalism would lead to Ethical Egoism, i.e., the position that it is always morally required of you to maximize your own well-being. But this is highly problematic, because the interests of others seem to be relevant from a moral point of view.

Let us turn to the final suggestion. According to some leading moral error theorists, the argument from queerness is best understood as targeting the categoricity or irreducible normativity of morality.[18] This suggests that if there are moral reasons, then the following is true of them:

Authority. If a person has moral reason to φ, then the person has categorical reason to φ.

Roughly, a reason is categorical in the relevant sense if two conditions hold. First, for a reason to be categorical, it must be a normative reason, a reason that matters from the standpoint of practical reason. Second, in addition, the reason must be irreducible. Irreducible in what sense? Error theorists elaborating on this question mention in particular that the reason must be such

---

[18] Garner (1990), Joyce (2001), Olson (2011, 2014).

that it can neither be explained in terms of the agent's desires nor in terms of institutional arrangements.

How does this help when it comes to deriving MU from RU? Notice that if we assume that categoricity is a *distinctive* feature of morality, as error theorists suggest, then we seem justified in turning Authority on its head:

> ***Reverse Authority***. If a person has categorical reason to φ, then the person has moral reason to φ

Reverse Authority could be used to get from RU to MU. The reasons involved by RU are categorical: you have reason to maximize well-being, according to RU, that can neither be reduced to your desires nor to institutional arrangements. Therefore, according to Reverse Authority, RU qualifies as a moral theory.

Our response to this proposal is based on the elaborations in section 2. There we suggested, following Parfit, that prudence involves categorical reasons: you have reasons for doing what is good for you, and these reasons seem to be irreducible – in particular, they cannot be reduced to your desires, as our discussion of Instrumentalism showed. Therefore, we reject the suggestion that categorical reasons are distinctive of morality.[19]

The upshot is that, even if moral requirements involve categorical reasons or are rationally authoritative, this does not enable us to derive MU from RU.

---

[19] Error theorists could respond that *if* prudential reasons are categorical, then they should be tossed out of the window together with moral reasons because both kinds of reasons would be 'queer'. But here we are not interested in the argument from queerness; our aim was to see if the reasons stated by RU should be classified as moral due to their categoricity.

4. From Rational to Moral Utilitarianism

Can we do better? We propose the following argument:

(1) [RU:] For any alternative acts *a* and *b,* there is more reason to perform *a* than *b* – and, accordingly, less reason to perform *b* than *a* – to the degree that, and only because, the outcome of *a* contains more well-being than the outcome of *b*.

(2) If RU is true, then all reasons for action are moral reasons.

Therefore,

(3) All reasons for action are moral reasons.

(4) If RU is true and all reasons for action are moral reasons, then an act is morally right iff, and because, none of its alternatives is more strongly favoured by moral reasons.

Therefore,

(5) An act is morally right iff, and because, none of its alternatives is more strongly favoured by moral reasons.

(6) An act is morally wrong iff it is not morally right. An act is morally obligatory iff none of its alternatives is morally right.

Therefore,

(7) [MU:] An act is morally right iff, and because, none of its alternatives leads to an outcome with more well-being. An act is morally wrong iff it is not morally right. An act is morally obligatory iff none of its alternatives is morally right.

Premise (1) is assumed for the sake of argument, (3) follows from (1) and (2), (5) follows from (1), (3) and (4), (6) is a conceptual truth, and (7) follows from (1), (5) and (6). The problematic

premises, which we are going to defend next, are (2) and (4). We start by making prima facie cases for (2) and (4) in this section and turn to objections in the next two sections.

Why believe (2), i.e., the proposition that if rational utilitarianism is true, then all reasons for action are moral reasons? In the previous section, we have dismissed Reverse Authority, which would have given us (2). Here, we want to suggest that three widely perceived general features of morality cohere with RU in a way that lends credence to (2). These features constitute what we call *the moral guise of RU*.

First, impartiality. Morality is commonly perceived as being essentially impartial in some way. RU accommodates the impartiality of morality because every person's (or even every sentient being's) well-being is equally important on RU.

Second, benevolence and beneficence. Benevolence and beneficence are widely considered to be important parts of morality. Since a person who follows RU does what best promotes the well-being of individuals affected by them, the followers of RU perform acts of beneficence and their attitudes are best described as attitudes of generalized benevolence.[20]

Third, goodness. Although this is dispensable, utilitarians – classically and typically – accept an axiology according to which one outcome is *better* than another to the degree that the first contains more well-being than the second.[21] This impartial axiology, which evaluates outcomes as better simpliciter rather than better-for, seems to be a moral axiology.

So, we submit that RU has a moral guise in virtue of features related to impartiality, benevolence and beneficence, and goodness. Notice that the lack of this kind of feature explains why the attempt to derive Ethical Egoism from Rational Egoism seems desperate. Rational Egoism just lacks a moral guise. It is also noteworthy that our ability to differentiate between

---

[20] cf. Smart 1977. Foot (1985) has famously argued that utilitarianism fails to recognize that benevolence is only one virtue among others. For our purposes, however, it is enough to concede that benevolence *is* plausibly considered a virtue and can therefore serve as an one element in the moral guise of RU.

[21] Why think that an axiology is dispensable for utilitarianism? Well, notice that the very definitions of RU and MU that we employ in this paper do not contain references to axiological concepts like *better* or *good* but simply refer to the well-being contained in outcomes.

the cases of RU and Rational Egoism makes our content-based (and more piecemeal) answer to Q more attractive than the attempts that we rejected in section 3. Note, finally, that our focus on the moral guise of theories of practical reason suggests general standards by which to answer questions of the Q-kind.

Before we turn to (4) and then to more general objections to (2) and (4) in the next sections, let us briefly address two worries concerning our argument for the prima facie plausibility of (2). The first worry is that, even though the moral guise of RU suggests that RU *can* rationally accept the view that all reasons for action are moral reasons, it is unclear why they *should* accept that view. Wouldn't it be better, as Crisp seems to suggest, to avoid moral concepts altogether? We don't think so. Recall that, according to the methodology outlined in section 1, it counts in favour of normative theories if they can explain our considered normative judgements. But moral discourse is certainly a major part of normative discourse, and if rational utilitarians cannot only explain our considered judgements about what one has reason simpliciter to do but also what one has moral reason to do, then this counts in favour of RU. Avoiding moral concepts altogether, hence, is just to forego free-lunch theoretical benefits. The upshot is that advocates of RU should accept (2) because this allows them to incur the theoretical benefit of explaining not only reasons simpliciter but also moral reasons. [22]

The second worry is that the moral guise of RU is not best accommodated by inferring that, as (2) has it, all practical reasons are moral reasons, but by a kind of fragmentation: Why not say that, given RU, only those reasons that concern actions affecting others are moral reasons, whereas those reasons that concern actions affecting only the agent are not moral, but prudential reasons?

---

[22] Notice that these thoughts provide us with reason to reject scalar utilitarianism. Scalar utilitarianism operates with axiological concepts but eschews the binary deontic concepts of moral obligation, rightness, wrongness etc., cf. Norcross XXXX. On the wide reflective equilibrium account adopted in this paper, scalar utilitarianism has the demerit of not being able to explain moral obligations.

In response, we agree that rational utilitarians should accept the view that reasons concerning actions that affect only the agent are prudential reasons. For by accepting this view, rational utilitarians become able to provide a neat account of prudence and incur even further benefits in terms of explanatory power and scope. We already alluded to this in section 2.[23] However, the worry suggested, in addition, that prudential reasons are not (also) moral reasons. *This* claim, though, should be rejected by rational utilitarians. Of course, it could be argued that moral and prudential reasons are mutually exclusive according to common-sense morality, and some ethical theories suggest such a view. But, however that may be, the above-mentioned considerations that constitute the moral guise of RU – *viz.,* impartiality, benevolence and beneficence, and goodness – are equally present in situations where the actions of a utilitarian agent affect only that agent. Prudence here coincides with impartiality, the best outcome is the one best for the agent, and the agent, by acting prudentially, actualizes benevolence and beneficence towards herself. This suggests that, given RU, we should drop the view that moral and prudential reasons are mutually exclusive and instead accept the claim that *all* reasons for action are moral reasons. Hence, the idea of fragmentation fails.[24]

Let us now turn to (4), i.e., the proposition that, if RU is true and all reasons for action are moral reasons, then an act is morally right iff, and because, none of its alternatives is more strongly favoured by moral reasons. If we remove the focus on RU, (4) becomes a more general claim:

(8) If all the reasons that constitute a requirement of practical reason are moral reasons, then the requirement is also a requirement of morality.

---

[23] See also Nagel, Parfit…

[24] We return to the issue that utilitarianism does not cohere with common-sense morality's distinction between moral and prudential reasons when addressing objections in section 5. And we briefly revisit the utilitarian account of prudence in section 8.

We take it that (8) is prima facie plausible. It is hard to see how a requirement that is constituted by moral reasons could fail to be a moral requirement. Since (4) is just an instance of (8), it should now be clear why (4) is prima facie plausible.

Are we too quick? (8) suggests that rational utilitarians *can* rationally accept the view that an act is morally right iff, and because, none of its alternatives is more strongly favoured by moral reasons. But why *should* they accept that view? This is the same kind of worry that we addressed above regarding our case for (2) based on the moral guise of RU. And our response is similar. By embracing (4), rational utilitarians become able to explain our considered judgements about right actions. This makes RU more appealing than it otherwise would be.

Let us stress again that the arguments for (2) and (4) presented in this section are just meant to lend prima facie credibility to these propositions. For a full justification, we need to address objections, which so far we have postponed. Let us now turn to these objections.

5. Discrepancies

The objections that will be addressed in sections 6 and 7 can be spelled out as being directed against propositions (2) or (4) of our argument (or against both propositions). The starting point of both objections is the observation that there are significant discrepancies between MU and common-sense morality. The objections differ in their claims about the relevance of these discrepancies.

What are the discrepancies between MU and common-sense morality? The most salient discrepancies include:

First, over-demandingness and over-permissibility. Textbook objections to MU point out that MU does not allow for supererogation and cannot accommodate special obligations,

promises, individual rights, and other phenomena in ways that cohere with common-sense morality. [25]

Second, over-moralization. MU implies that imprudent actions that affect only the agent herself are morally wrong.[26] My decision to watch Star Wars rather than Star Trek might be morally wrong according to MU.

Third, responsibility. Common-sense morality suggests that only wrong actions are blameworthy and that only right actions are praiseworthy. Moral utilitarians typically deny these claims.[27]

Given our methodological assumptions (section 1), it is clear that the discrepancies between MU and common-sense morality *prima facie* count against MU. Even though the discrepancies do not pose a problem for MU *if* RU is correct and implies MU, we should briefly consider the standard defences employed by moral utilitarians. Even though it is beyond the scope of this paper to evaluate how successful these standard defences are without the assumption of RU, it is important to notice that such defences are available. For the availability of the defences bears on our responses to the objections considered in sections 6 and 7.

There are two basic approaches for moral utilitarians to respond to the discrepancies. First, moral utilitarians can attack common-sense morality and argue that, upon closer examination, morality turns out not to be as suggested by common-sense morality. To this end, moral utilitarians can argue, e.g., that common-sense morality is incoherent in some respects.[28] Second, moral utilitarians have suggested that MU is *less* demanding than is often claimed and that moral utilitarianism can accommodate such phenomena as supererogation, moral rights,

---

[25] Bvkvist (XXXX), Mulgan (XXXX), Shaw (XXXX). The overdemandingness-objection can be spelled out in terms of moral authority rather than referring to the content of morality (supererogation). However, given our elaborations in section 2, the interpretation focusing on authority is not relevant.

[26] REF (many rule-utilitarians…)

[27] REF.

[28] Sidgwick, Parfit, Kagan… A different kind of attack on common-sense morality consists in debunking intuitions that seem to support common-sense morality; REF.

responsibility, and the distinction between prudence and morality.[29] And of course, the two approaches can be combined. It can be argued that common-sense morality gets it wrong to some extent and that MU can accommodate what is left of common-sense morality after removing the misguided contents.

If RU is true, this strengthens the prospects of defending MU in light of the discrepancies. However, it is important to notice that moral utilitarians have suggested defences that are independent of the truth of RU. Such independent defences are important when we evaluate, in the following sections, objections to our argument for an affirmative answer to Q.

6. The Error-Theoretical Objection

How can the discrepancies between MU and common-sense morality block the move from RU to MU? The discrepancies might make it seem problematic that we introduce the concepts of moral reason and moral rightness in propositions (2) and (4) of our argument. Our critic might want to maintain what we call the

> ***Error-Theoretical Objection***: Given the discrepancies between MU and common-sense morality, the correctness of RU does not imply MU but a moral error theory.

Typically, arguments for moral error theories have two steps: one first shows that moral discourse is centrally committed to some thesis X and then shows that X is false.[30] In our case, X would need to be some (possibly conjunctive) thesis that features the discrepancies between MU and common-sense morality, e.g.: that there are supererogatory actions, that we have

---

[29] REF…
[30] REF…

special obligations, or that one is to be blamed only for morally wrong actions. The second step, which we will accept for the sake of argument, is that, given RU, *X* is false.

Let us focus on the first step: How could one establish that moral discourse is centrally committed to *X*? Certainly, for moral discourse to be centrally committed to a thesis, it is not sufficient that the thesis is contained in common-sense morality. Otherwise, moral revisions would not be possible. But what, then, is sufficient? A general criterion for what it takes for moral discourse to be centrally committed to a thesis is hard to find. As Joyce puts it:

> Even if the error theorist can articulate a clear and determinate problematic feature of morality, the dispute over whether this quality should count as a 'non-negotiable component' of morality has a tendency to lead quickly to impasse, for there is no accepted methodology for deciding when a discourse is 'centrally committed' to a given thesis. [31]

However, even without a general criterion, it is plausible to hold that MU does not contradict a central commitment of moral discourse, but is only highly revisionary. This is plausible for four reasons.

First, compare the Error-Theoretical Argument discussed here with what is widely perceived as the strongest argument for moral error theory: the argument from queerness. As mentioned in section 3, this argument is often understood as focusing on categoricity in the sense of irreducible normativity. Now, for the argument from queerness to support moral error theory, moral discourse would have to be committed to the claim that moral facts are categorical in this sense. Importantly, if morality involved categorical reasons, then this feature of morality would be global: all of morality would be pervaded by categorical reasons. By

---

[31] Joyce SEP

contrast, the discrepancies between MU and common-sense morality concern only certain parts of moral discourse. Of course, we should not belittle the discrepancies. But even staunch critics of MU should admit that the extent to which MU 'gets it wrong' (according to the critics) concerns only certain parts of morality and aspects that are less basic than the issue of categorical reasons. The upshot is that the proponent of the Error-Theoretical Argument has to shoulder a heavy burden if he is to argue that moral discourse is centrally committed to some thesis that is denied by MU, considerably heavier than the burden resting on the shoulders of error theorists who base their views on the argument from queerness.

Second, notice that when anti-utilitarian moral philosophers discuss the discrepancies between MU and common-sense morality, they typically do not go further than to claim that MU is (clearly) false or (highly) implausible and they do typically concede that moral utilitarianism gets *something* right about morality and even has *some* prima facie plausibility as a moral theory.[32] Compare this to the reaction of moral philosophers who discuss Ethical Egoism: Ethical Egoism is often and plausibly considered a *non-starter* as a moral theory.[33] These different reactions by the experts – to MU on the one hand and Ethical Egoism on the other – suggest that MU has a stronger claim to be considered a (perhaps, highly-revisionary) moral theory than Ethical Egoism, with the latter being in such head-on conflict with basic moral intuitions that it could not even be considered a moral theory.

Our third reason for thinking that MU does not contradict a central commitment of moral discourse is that the moral guise of utilitarianism reinforces the first two reasons. The difference in reactions of experts to MU on the one hand and Ethical Egoism on the other is easily explained by the considerations underlying what we called the moral guise of RU (impartiality, benevolence and beneficence, and goodness). These considerations are missing

---

[32] Examples…
[33] Examples…

in the case of Rational Egoism. From the perspective of common-sense morality, Rational Egoism seems to have a *prudential* guise. And the moral guise of utilitarianism illustrates that MU coheres to a considerable degree with common-sense morality, which makes it very hard for the proponent of the Error-Theoretical Argument to show that the discrepancies between MU and common-sense morality amount to the denial of a thesis to which moral discourse is centrally committed.

Finally, as we have seen in section 5, moral utilitarians have proposed defenses regarding the discrepancies between their position and common-sense morality. To the extent that these defences are successful, we have *additional* reason for rejecting the Error-Theoretical Objection.

7. The Alternative-Theory Objection

Having set aside the Error-Theoretical Objection, we can now turn a perhaps even more forceful challenge. It is the following:

> ***The Alternative-Theory Objection***: Given the discrepancies between MU and common-sense morality, some theory other than MU is the best moral theory. Standard methods of selecting moral theories – such as the method of reflective equilibrium – strongly suggest that MU, given its discrepancies with common-sense morality, is inferior to other candidate moral theories.

In response, notice, first of all, that, if moral rationalism is true, then the objection most likely fails. For it would be hard to come up with a moral theory that is different from MU but nonetheless yields only obligations which agents have decisive reason to perform (in accordance with moral rationalism) given RU, i.e., given that agents always have decisive

reason to promote well-being impartially. A direct rebuttal of the Alternative-Theory Objection would thus consist in arguing for moral rationalism. However, since moral rationalism is controversial and we do not want to be committed to its truth, this is not our preferred way of dealing with the objection.[34] However, it is important for our criticism of the Alternative-Theory Objection that even those meta-ethicists who reject moral rationalism typically argue that it is rationally permissible for most agents and in most situations to fulfil their moral duties.[35]

Our argument against the Alternative-Theory Objection resembles pliers (and is somewhat similar to a dilemma). We first delimit the possibilities of what the allegedly best alternative theory could look like and argue that proponent of the Alternative-Theory Objection would not want the alternative theory to sit on an edge, i.e., be close to an extreme. (This resembles the horns of a dilemma.) Then we close the pliers, so to speak, by arguing that the positions between the extremes would not be comfortable either.

Let's begin. Assume that the proponent of the objection singles out some moral theory, M, (with M being different from MU) as the best moral theory. The more similar M is to common-sense morality, the more different M will be from MU. (So, common-sense morality and MU are the extremes.) If M is *very* different from MU, then – given the assumption of RU – you will hardly ever have reason to comply with M. A proponent of M will, therefore, not only have to deny moral rationalism, but will be committed to the position that you hardly ever have reason to comply with M. But this position is hard to believe. As mentioned before, even those philosophers who deny moral rationalism find it plausible that you often have reason to comply with morality. Accordingly, M will invite error-theoretical objections.

---

[34] Critics of moral rationalism include Foot, Brink, Railton.…
[35] Foot, Brink, Railton…

Notice that these error-theoretical objections will have considerable force. True, being much more similar to common-sense morality than MU, *M* will, unlike MU, not give rise to objections based on discrepancies between its contents and common-sense morality. But we have seen in the previous section that the discrepancies of MU likely fail to justify a moral error theory in the case of RU anyhow. In contrast, the error-theoretical objections to *M*, if *M* is very similar to common-sense morality and given RU, are more similar to the argument from queerness, in the sense of affecting (almost) all of morality.

On the other hand, the more similar *M* is to MU, the more discrepancies there will be between *M* and common-sense morality. If *M* is *very* similar to MU (like, say, a satisficing variant of MU), then *M* loses almost entirely its advantages over MU in terms of resembling common-sense morality. Moreover, it is precisely in those areas in which *M* keeps these advantages that *M* lacks rational authority. It is hard to see, then, why, given the assumption of RU, one would want to accept *M* rather than MU.

Of course, one could respond with regard to the last point that it is not a big problem for *M* to lack rational authority where the practical implications of *M* differ from MU because, by hypothesis, standard methods of moral theory selection will imply that *M* is more likely to be correct than MU. E.g., while one might not have reasons to favour one's own interests over those of others in certain situations, it is likely to be true that one has a *moral* prerogative to favour one's own interests in these situations. However, this response is unconvincing because if *M* is very similar to MU but lacks rational authority precisely where it differs from MU, then this gives us reason for thinking that we made an epistemic mistake – i.e., we did not apply our methods of theory selection correctly – and that MU is after all superior to *M*. Put differently, the constellation under consideration seems to be incoherent because our best methods of theory selection will recognise that rational authority is *one* factor that can favour one theory

over another, and when two theories are very similar as far as their contents are concerned, then it is plausible to think that this factor is decisive.

It is now time to close the pliers. So far, we have argued that, given RU, neither theories that are very similar to MU nor theories that are very similar to common-sense morality would be superior to MU. But what about theories that lie between these extremes? In terms of its contents, any allegedly superior theory *M* that is neither very similar to common-sense morality nor very similar to MU would necessarily incur costs in terms of rational authority *and* in terms of coherence with common-sense morality. Moreover, the costs in either category could only be reduced by trading them off against more costs in the other category. It is therefore hard to see how such a theory could be superior to MU.

Like our response to the Error-Theoretical Argument in the last section, the response to the Alternative-Theory Objection in this section does not amount to a 'proof' that the objection fails. However, it seems to us that, when considering the multiple considerations adduced in our responses, the proponent of the Alternative-Theory faces a tough challenge.

Notice, finally, that our discussion so far does not take into account the defences proposed by moral utilitarians regarding the discrepancies with common-sense morality (see section 5). The availability of these defences makes it even more likely that the Alternative-Theory Objection fails.

8. From RU to MU…And Beyond

This concludes our journey from RU to MU. Let us now go beyond by considering the overall picture. Recall that not only MU but also, as we suggested in sections 2 and 4, a plausible utilitarian theory of prudence, which is very similar to Rational Egoism, can be derived from RU. We suggest formulating the utilitarian theory of prudence along the lines of Rational Egoism but restrict its scope to situations that affect only the agent:

Prudential Utilitarianism (PU): For any alternative acts *a* and *b* that affect only the agent*,* the agent has more prudential reason to perform *a* than *b* – and, accordingly, less prudential reason to perform *b* than *a* – to the degree that, and because, the outcome of *a* is overall better for the agent than the outcome of *b*.

PU cannot conflict with MU because PU does not apply in multi-person cases. If PU did apply, the agent's well-being would be counted double, which would contradict RU.

There are two salient alternatives to this proposal, which both would have PU apply also to multi-person cases. The first is to have PU lack rational authority in multi-person cases. The second is to modify *MU* such that it does not take into account the agent's well-being and have MU and PU compete in ways that on balance are in line with RU. While we think our proposal is more attractive, there is no need for discussion. The important point is that RU cannot only accommodate morality but also prudence.

We have explained (when discussing propositions (2) and (4) of our argument in section 4) that rational utilitarians should embrace MU. The same holds for PU, and for similar reasons. RU gains explanatory scope via PU just like it does via MU.

Finally, notice that we have not yet exhausted the limits of RU. Utilitarian theories have also been suggested as political and legal normative theories and in aesthetics.[36] It seems likely that such theories can be inferred from RU, at least when RU is carefully expanded. For example, we can just apply RU to multiple agents, as a theory of what agents have reason to do *together*.[37] This would give us an account of political action. Or we might add a utilitarian theory of value to RU.[38] Then RU will yield evaluations of political and social institutions.[39] If we are able to distinguish between aesthetic and non-aesthetic pleasures or preferences, we can

---

[36] REF
[37] Woodard
[38] See section 3.
[39] Hardin

develop a utilitarian theory of aesthetic reasons (which would need to be restricted in scope, even more than PU).

The point is that the increase in explanatory scope that RU gains by implying MU und PU make RU even more attractive than the arguments of sections 2. By promising further implications, of utilitarian theories in politics and aesthetics, RU scores even higher on the method of wide reflective equilibrium, leaving us with a beautiful utilitarian landscape.

9. Conclusion

Our paper began by presenting reasons for believing that rational utilitarianism is the best theory of practical reason. Then we argued that rational utilitarianism (broadly) implies moral utilitarianism. We also argued that rational utilitarianism implies prudential utilitarianism and promises to imply utilitarian theories in other normative realms. The most important upshot of all this is that rational utilitarianism is attractive even apart from the considerations that favour if over alternative theories of practical reason: By implying utilitarian theories in other normative realms, like morality and prudence, rational utilitarianism displays tremendous explanatory scope and explanatory power.

The focus of the paper was on the question of whether philosophers who argue that rational utilitarianism is the best theory of practical reason should also claim that utilitarianism is the best moral theory. We think that they should. Despite what one might think, the step from rational to moral utilitarianism is not trivial. In particular, even if moral requirements are requirements of practical reason, this would not by itself bring us from rational to moral utilitarianism. However, considerations regarding the contents of rational utilitarianism and constitute its moral guise support the step from rational to moral utilitarianism. This step is not made incredible by the discrepancies between common-sense morality and moral utilitarianism. The discrepancies merely show that moral utilitarianism is – perhaps highly –

revisionary. But they fail to show that moral utilitarianism is not the best moral theory. If rational utilitarianism is the best theory of practical reason, then moral utilitarianism is the best moral theory.

It would be interesting to compare rational utilitarianism even more to other theories of practical reason, in addition to our comparisons of section 2 and our elaborations on why Rational Egoism does not plausibly imply Ethical Egoism. We challenge Kantians and others to show that their theories do not fare worse on the method of wide reflective equilibrium.

References

Brink, D. O. (1997). "Kantian Rationalism: Inescapability, Authority, and Supremacy." In G. Cullity and B. Gaut (eds.), Ethics and Practical Reason, pp. 255–291. New York: Oxford University Press.

Buss, Sarah (2006), "Needs (Someone Else's), Projects (One's Own) and Reasons." Journal of Philosophy, 103, 373–402.

Bykvist

Calhoun, Cheshire (2009), "What Good is Commitment." Ethics, 119, 613–41.

Chang, Ruth (2013), "Commitments, Reasons, and the Will" in R. Shafer-Landau, ed., Oxford Studies in Metaethics, vol. 8. Oxford: Oxford University Press, 74–113.

Crisp, Roger (2006): *Reasons and the Good*. Oxford University Press.

Darwall (2006). "Morality and Practical Reason: A Kantian Approach." In D. Copp (ed.), The Oxford Handbook of Ethical Theory, pp. 282–320. Oxford: Oxford University Press.

de Lazari-Radek and Singer

Dorsey, Dale (2016). *The Limits of Moral Authority*. Oxford: Oxford University Press.

Foot

Garner, R. (1990): "On the genuine queerness of moral properties and facts", *Australasian Journal of Philosophy* 68: 137–146.

Griffin, James (1986), Well-Being: Its Meaning, Measurement and Moral Importance, Oxford: Clarendon Press.

Hardin

Joyce, R. (2001): *The Myth of Morality.* Cambridge University Press.

Joyce SEP

Kagan

Kahane, Guy (2011): "Evolutionary Debunking Arguments," *Noûs* 45: 103–25.

Korsgaard, Christine (1997), The Sources of Normativity. Cambridge: Cambridge University Press.

Mackie, J. L. (1977): *Ethics: Inventing Right and Wrong,* Pelican.

Manne, Kate (2013), "On Being Social in Metaethics" in R. Shafer-Landau, ed., Oxford Studies in Metaethics, vol. 8. Oxford: Oxford University Press, 50–73.

Mill

Mulgan

Olson, J. (2011): "In Defense of Moral Error Theory", *New Waves in Metaethics*: 62–84.

Olson, J. (2014): *Moral Error Theory: History, Critique, Defence*. Oxford University Press.

Parfit (1984)

Parfit (2011)

Portmore, Douglas (2007), "Welfare, Achievement, and Self-Sacrifice." Journal of Ethics and Social Philosophy, 3, 1–28.

Portmore (2011, 2021),

Railton

Shaw 1999

Sidgwick

Smart 1956

Smart 1977

Stroud, S. (1998). "Moral Overridingness and Moral Theory." Pacific Philosophical Quarterly 79: 170–189.

Williams, Bernard (1976), 'Persons, character and morality' in A. O. Rorty (ed.), The identities ofpersons (Berkeley: University of California Press).

Woodard