Reflective equilibrium and moral objectivity

Sem de Maagt

To cite this article: Sem de Maagt (2017) Reflective equilibrium and moral objectivity, Inquiry, 60:5, 443-465, DOI: 10.1080/0020174X.2016.1175377

To link to this article: https://doi.org/10.1080/0020174X.2016.1175377
Reflective equilibrium and moral objectivity

Sem de Maagt
Department of Philosophy and Religious Studies, Utrecht University, Utrecht, The Netherlands

ABSTRACT
Ever since the introduction of reflective equilibrium in ethics, it has been argued that reflective equilibrium either leads to moral relativism, or that it turns out to be a form of intuitionism in disguise. Despite these criticisms, reflective equilibrium remains the most dominant method of moral justification in ethics. In this paper, I therefore critically examine the most recent attempts to defend the method of reflective equilibrium against these objections. Defenders of reflective equilibrium typically respond to the objections by saying that either reflective equilibrium can in fact safeguard moral objectivity or alternatively, even if it cannot, that there simply are no reasonable alternatives. In this paper, I take issue with both responses. First, I argue that given the non-foundationalist aspirations of reflective equilibrium, moral objectivity cannot be maintained. Second, I argue that reflective equilibrium is not the only game in town once intuitionism has been discarded. I argue that given their own normative ambitions, combined with their rejection of intuitionism, proponents of reflective equilibrium have reason to take alternative methods of moral justification, and more specifically transcendental arguments, more seriously than they have done so far. I end by sketching the outlines of what this alternative methodology might look like.

ARTICLE HISTORY
Received 27 July 2015; Accepted 29 March 2016
KEYWORDS
Reflective equilibrium; moral objectivity; coherentism; foundationalism; considered judgments; transcendental arguments

CONTACT
Sem de Maagt • s.demaagt@uu.nl

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.
1. Introduction

Reflective equilibrium is the most dominant method of moral justification in both normative and applied ethics.\(^1\) In addition, reflective equilibrium is widely accepted by both moral realists and moral constructivists as their preferred moral epistemology.\(^2\) One of the main appeals of reflective equilibrium as a moral methodology lies in its ambition to provide an account of moral objectivity, without having to bear the controversial epistemological and metaphysical burdens of alternative, ‘foundationalist’ methods of moral justification, such as rational intuitionism, which claim that certain moral beliefs can somehow be known non-inferentially (see, e.g. Daniels 1979, 1980; Rawls 1999a; Walden 2013; Scanlon 2014).\(^3\)

However, ever since the introduction of reflective equilibrium in ethics, it has been argued that reflective equilibrium cannot have it both ways (Hare 1973; Singer 1974, 2005; Brandt 1979). Either one accepts reflective equilibrium as a distinctive, non-foundationalist method of moral justification, in which case one has to give up the objectivist aspirations of ethics. Alternatively, one may want to preserve moral objectivity, in which case one has to be prepared to commit oneself to the kind of foundationalism proponents of reflective equilibrium reject. That is, the objection is that reflective equilibrium is either committed to moral relativism, or turns out to be a form of intuitionism in disguise.\(^4\)

Defenders of reflective equilibrium typically respond to these kinds of objections in two ways. First, they try to show that reflective equilibrium does in fact have the resources to safeguard moral objectivity, without relying on foundationalist premises (e.g. Daniels 1979; Holmgren 1987; Brink 1989; Scanlon 2014). Second, they argue that even if reflective equilibrium

---

\(^1\)The method of reflective equilibrium has its origin in Goodman's (1955) defense of the rules of inductive inference. Rawls (1999a) introduced reflective equilibrium in ethics as an alternative to rational intuitionism. For an overview of the historical predecessors of reflective equilibrium in ethics see Brink (2014). For an overview of the use of reflective equilibrium in applied ethics see Arras (2007).

\(^2\)For references see Section 2, of this paper.

\(^3\)Due to limitations of space, I will assume throughout the paper that proponents of reflective equilibrium are correct in dismissing intuitionism on epistemological and/or metaphysical grounds.

\(^4\)I take moral relativism to be the view that the correctness of moral judgments is in some way dependent on contingent standards, such as an individual’s motivational set, a specific cultural context or a specific conceptual scheme (for recent defenses of moral relativism see, e.g. Street [2006, 2009, 2012] and Velleman [2013]). Due to limitations of space I cannot discuss the relative merits of accounts of moral objectivity vis-à-vis various forms of moral relativism. Instead, this paper takes for granted the idea of moral objectivity that proponents of reflective equilibrium themselves accept (see also Section 3 of this paper). Let me stress that I agree with Street (2009) and Velleman (2013) that moral relativism is a more respectable position than is often assumed. However, the case for moral relativism rests on a rejection of moral objectivity (see, e.g. Street 2012, 24; Velleman 2013, 94). Moral relativism, as even defenders of moral relativism seem to acknowledge, is therefore only a second best theory. Moral relativism can therefore only be defended once we conclude that accounts of moral objectivity indeed fail.
ultimately leads to moral relativism, this constitutes no objection to the methodology because there simply exists no reasonable alternative to reflective equilibrium (e.g. Scanlon 2003, 149; DePaul 2006, 618; Walden 2013, 254; Floyd, forthcoming, 12).

In this paper, I take issue with both responses. First, I critically evaluate the most recent defenses of reflective equilibrium and argue that they fail to show that reflective equilibrium can lead to objective moral beliefs, at least insofar as reflective equilibrium is to remain a distinctively non-foundationalist methodology. I focus mainly on Scanlon’s (2014) defense of reflective equilibrium in his latest book Being Realistic about Reasons. Second, I argue that reflective equilibrium is not the only game in town once intuitionism has been discarded.

The aim of this paper is not to provide an ultimate rejection of reflective equilibrium, but to clarify the (normative) costs involved in accepting reflective equilibrium. In addition, I aim to argue that given their own normative ambitions, combined with their rejection of intuitionism, proponents of reflective equilibrium have reason to take alternative methods of moral justification, and more specifically transcendental arguments, more seriously than they have done so far.

The structure of the paper is as follows. In Section 2, I briefly introduce the method of reflective equilibrium. In Section 3, I provide an analysis of moral objectivity and provide some prima facie reasons to question the ability of reflective equilibrium to deliver the kind of moral objectivity it aspires to deliver. In what follows, I discuss recent attempts by proponents of reflective equilibrium to face these challenges and conclude that reflective equilibrium cannot live up to its objectivist promise, at least as long as one wants to retain reflective equilibrium as a distinctive, autonomous and practicable moral methodology (Sections 4 and 5). In the final section of the paper (Section 6), I argue that reflective equilibrium is not the only game in town once intuitionism has been discarded. I end by sketching the outlines of an alternative methodology in the form of a transcendental argument in ethics.

2. Reflective equilibrium

The distinctive claim of reflective equilibrium is that moral justification does not depend on an ultimate moral foundation, but on the coherence between all moral and non-moral beliefs that are relevant to the issue at
hand. Reflective equilibrium thus offers a ‘non-foundationalist’ or ‘coherentist’ account of moral justification. Reflective equilibrium is non-foundationalist because it claims that there exist no beliefs that have a special justificatory standing independent from engaging in the process of pursuing reflective equilibrium. Reflective equilibrium is a form of coherentism because, according to this methodology, justification, as Rawls (1999a) famously writes, ‘is a matter of the mutual support of many considerations, of everything fitting together into one coherent view’ (507).

As a method of moral justification, reflective equilibrium is, at least to some extent, neutral with respect to the metaphysics of morality. In fact, it is an influential method of moral justification for both moral realists and moral constructivists. Brink, for instance combines reflective equilibrium as a method of moral justification with his naturalist moral realism. That is, he claims that ‘coherence with, among other things, considered moral beliefs provides evidence of objective moral truth’ (Brink 1989, 143; for similar views see Holmgren 1987; Ebertz 1993; Scanlon 2012, 2014). Brink thus embraces reflective equilibrium as a moral epistemology and at the same time he holds that moral facts are ‘mind-independent’, so that there is a conceptual gap between moral justification and moral truth. This implies that a moral belief might be ultimately false, even though the belief is justified in the light of all possibly relevant considerations we have in favor of the acceptance of the belief.

Rawls (1980), on the other hand, defends both the epistemological and a metaphysical interpretation of reflective equilibrium, claiming that his

---

5This obviously raises the question of what exactly counts as the appropriate type of coherence and the criteria that should guide the process of reaching reflective equilibrium. Rawls does not go beyond the metaphorical idea that we should go ‘back and forth’ to describe the criteria that should guide the process of going back and forth between different (moral) beliefs. Others have tried to provide a list of epistemic desiderata that should guide this process, including consistency, systematicity, generality and simplicity (Kappel 2006; Campbell 2014).

6Some authors claim that reflective equilibrium is compatible with – or even most properly understood as – a form of modest foundationalism according to which certain beliefs, i.e. considered judgments, have a privileged epistemic status independent of their coherence with other beliefs even though these ‘foundational’ beliefs are not immune to revision (DePaul 1986; Holmgren 1987; Ebertz 1993; McMahan 2000). In Section 4, I will argue that a modest foundationalist interpretation of reflective equilibrium collapses into a form of intuitionism and that we should therefore reject a modest foundationalist interpretation of reflective equilibrium and understand reflective equilibrium as a distinctively coherentist methodology (see also Brink 2014; Brun 2014).

7The idea that reflective equilibrium is neutral with respect to the metaphysics of morality should be qualified in at least two ways. First, although it might be true that reflective equilibrium is compatible with both moral realism and moral constructivism it might not be compatible with all possible metaphysical positions. For instance, it seems that for some forms of expressivism, the question of moral justification as it appears in this paper wouldn’t arise in the first place. In addition, it might be possible that there are specific challenges and problems for reflective equilibrium depending on whether one is a moral realist or a constructivist. For instance, the realist would have to explain why reflective equilibrium has any contact with the metaphysical reality of moral truths, whereas a constructivist would have to explain why we are morally committed to whatever follows from reflective equilibrium.
‘Kantian constructivism holds that moral objectivity is to be understood in terms of a suitably constructed social point of view that all can accept. Apart from the procedure of constructing principles of justice, there are no moral facts’ (519). In Rawls’ constructivist interpretation of reflective equilibrium, truth thus collapses into justification: moral truth simply is whatever we arrive at through a suitable procedure of constructing justice.

In order to understand the attraction of reflective equilibrium to both realist and constructivist, it is crucial to note that reflective equilibrium is introduced in ethics, first and foremost, as an alternative to various forms of epistemological intuitionism, i.e. the position which holds that fundamental moral truths can be justified non-inferentially, either because they are self-evident or through a special faculty of intuition. The introduction of reflective equilibrium as an alternative to intuitionism is most clearly articulated in Rawls’ (1999a) *A Theory of Justice* where he explicitly rejects the idea that his principles of justice ‘are necessary truths or derivable from such truths. A conception of justice cannot be deduced from self-evident premises or conditions on principles’ (19; for similar views see Daniels 1979, 264–267; Rawls 1980, 559; Brink 1989, 8–9; Scanlon 2014, 69–72). Instead, Rawls claims that the aim of his theory is to set up ‘an Archimedean point for assessing the social system without invoking a priori considerations’ (231). The introduction of reflective equilibrium as a method of moral justification is thus meant to disprove the idea that ‘there is no alternative but to judge institutions in the light of an ideal conception of the person arrived at on perfectionist or on a priori grounds’ (230). This goal of searching for an alternative moral epistemology to intuitionism is not limited to moral constructivists such as Rawls, but concerns a substantial number of moral realists as well who aim to disentangle the metaphysics of moral realism from the epistemology of intuitionism.8

Although proponents of reflective equilibrium reject the foundationalist aspirations of intuitionism on the meta-ethical level, they share the objectivist aspirations of intuitionism on the first-order normative level.9 That is, the idea among proponents of reflective equilibrium is that moral beliefs can still be objective, even though they cannot be derived from foundational principles. The ambition to combine reflective equilibrium with an objective ethics is most clearly expressed by Rawls (1999a) when he claims that

[the] principles [of justice] are objective. They are the principles that we would want everyone (including ourselves) to follow were we to take up together the

---

8 See above for references.
9 An important exception of this rule is Street’s (2006, 110) use of reflective equilibrium as part of her ‘Humean constructivism’.
appropriate general point of view […] We do not look at the social order from our situation but take up a point of view that everyone can adopt on an equal footing. In this sense we look at our society and our place in it objectively: we share a common standpoint along with others and do not make our judgments from a personal slant. Thus our moral principles and convictions are objective to the extent that they have been arrived at and tested by assuming this general standpoint and by assessing the arguments for them by the restrictions expressed by the conception of the original position. (453; see also Daniels 1979, 273–281; Brink 2014, 677f26; Scanlon 2014, 93–94)

Rawls thus suggests that the principles of justice are objective in the sense that they should be accepted by any individual independent from her personal preferences or desires. The method through which Rawls hopes to arrive at these objective moral beliefs is the method of reflective equilibrium.

If reflective equilibrium succeeds in fulfilling this ambition the method seems to be able to combine the best of both worlds, because it would be able to deliver objective moral beliefs while steering clear of the epistemological and metaphysical controversies surrounding intuitionist accounts of moral justification. However, in what follows, I will argue that proponents of reflective equilibrium cannot have their cake and eat it too.

3. Moral objectivity

In order to evaluate whether or not the method of reflective equilibrium is indeed able to generate objective moral beliefs, we should first be more precise about what the idea of moral objectivity exactly refers to. The point here is not to argue for a specific conception of moral objectivity, but to get a better understanding of the idea of moral objectivity that proponents of reflective equilibrium themselves accept.

For a start, we can understand the concept of objectivity as referring to the idea that the truth or correctness of a certain belief is in some way independent of us. In other words, a moral belief is objective when its truth or correctness is independent of whether one thinks something is true or correct. This intuitive understanding of objective in terms of independence is, however, open to different interpretations.

On a minimal interpretation of objectivity, a moral belief is objective when the correctness of a moral belief is independent of whether or not someone actually accepts a particular moral belief. That is, on this understanding of objectivity the fact that one holds a particular moral belief – e.g. the belief that redistributive taxation is a violation of property rights – is not what justifies the moral belief. Minimal objectivity thus allows for the possibility
that one can be wrong about a particular moral belief, because it separates the acceptance of a particular belief from its being justified.

Obviously, reflective equilibrium can account for this minimal type of objectivity. In reflective equilibrium, a moral belief is not simply justified when an individual actually holds a specific moral belief but only when this belief coheres, in the appropriate way, with all possibly relevant considerations to the issue at hand. One’s belief that redistributive taxation is a violation of property rights might, for instance, be undermined if it is shown to be inconsistent with the beliefs one has about the justification of private property and/or the relation between self-ownership and democratic decision-making. In reflective equilibrium, a particular moral belief might thus be disproven when it is shown to be inconsistent with other beliefs the individual holds, assuming that it is more rational to give up the particular belief under consideration and not the background beliefs with which the belief is shown to be inconsistent.\(^\text{10}\)

However, although the possibility that one might be wrong about a particular belief captures an important feature of moral objectivity, moral beliefs are typically thought to be objective in a more robust sense than is demanded by minimal objectivity. For instance, on a widely accepted view of morality, moral beliefs are not just those beliefs that cohere with one’s other beliefs, but moral beliefs are those beliefs one should accept independent of the contingent beliefs or desires one happens to have, whether or not these beliefs and desires are internally coherent. On a maximal interpretation of objectivity, the correctness of a moral belief is thus independent of someone’s contingently held belief set.

There are different ways to understand this maximal interpretation of objectivity. According to an ontological or realist conception of maximal objectivity, moral beliefs are correct when they correspond with moral reality, i.e. mind-independent moral facts of the matter. On this realist conception of objectivity, the correctness of a moral belief is thus understood in terms of its truth, which in its turn is understood in terms of the correspondence between moral beliefs and moral facts. On a practical or constructivist conception of (maximal) objectivity, on the other hand, the correctness of moral beliefs is to be determined by reference to the practical point of view an individual has to accept independent of her specific desires, aims or beliefs. On this constructivist conception of objectivity, no reference is made to the idea of moral truth since the correctness of a moral belief is not to be determined by reference to moral reality. Instead, the idea of moral

\(^\text{10}\)Inconsistency is a very minimal interpretation of coherence and some authors argue that reflective equilibrium needs a more substantive idea of coherence in order to be plausible (see, e.g. Arras 2007).
objectivity refers to those beliefs that an individual is necessarily committed to accept in virtue of being an agent. Rawls’ intersubjective conception of objectivity – the ‘common standpoint’ we share with others – is an example of this kind of practical objectivity.

Despite the important differences between these different conceptions of maximal objectivity they have in common, the idea that maximal objectivity demands that an individual could be wrong about a moral belief even though the belief coheres with the other relevant moral and non-moral beliefs she happens to hold. And even though realists and constructivists ultimately have different conceptions of moral objectivity – mind-independent moral truth vs. a practical standpoint every rational agent should adopt – there are several reasons to be skeptical about the idea that reflective equilibrium can deliver maximal objectivity, however one exactly understands it.

First, it seems that in order to generate objective moral beliefs, we should have reasons to believe that there will be (sufficient) convergence on a set of moral beliefs in reflective equilibrium. This kind of convergence seems to be essential to a method of moral justification because otherwise there might be as many possibly conflicting, justified moral belief sets as there are people engaging in the method of reflective equilibrium. Sufficient intersubjective agreement is the most important indicator of moral objectivity in the method of reflective equilibrium because without this kind of agreement there might be a plurality of incompatible reflective equilibria for which reflective equilibrium offers no resources to privilege one of them (see also Rawls 1999b, 515–516).

But why should we believe that convergence on a reflective equilibrium is likely to happen? Not only is there persistent disagreement about specific moral judgments in specific cases, but also with respect to background theories. For instance, there exists a plurality of incompatible metaphysical theories about personhood that have their own normative implications. In addition, there is disagreement about which background theories are morally relevant in the first place (Cohen 2003; Estlund 2011; de Maagt 2014). It is unclear how bringing different beliefs into coherence could increase our confidence in the possibility of intersubjective agreement on morality. It is therefore not surprising that the method of reflective equilibrium is used to support a wide range of normative (political) theories and positions, including Utilitarianism (Brink 1989); Rawlsian political theory (Rawls 1999a); Left-libertarianism (Vallentyne, Steiner, and Otsuka 2005); Cosmopolitanism (Caney 2006); and Nationalism (Miller 2012). It is important to note that the problem here is not merely that there might be a plurality of possibly conflicting reflective equilibria, but more importantly that the methodology
offers no resources to adjudicate between different equilibria, which means no more and no less than that it is normatively arbitrary which equilibrium one ends up accepting.

Second, there is a deeper concern about reflective equilibrium as a coherentist method of justification in relation to its claim to moral objectivity. Even if we assume that convergence on moral beliefs is to be expected, one might wonder why the very fact of coherence provides us with a reason to accept some moral conception, i.e. whether this is a reason to believe that it tracks independent moral truth or that it expresses a practical standpoint that every individual should adopt. This challenge is known as the ‘garbage in, garbage out’ objection (Jones 2005, 74). Brandt (1979) formulates this challenge as follows:

There is a problem here which is quite similar to that which faces the traditional coherence theory of justification of belief: that the theory claims that a more coherent system of beliefs is better justified than a less coherent one, but there is no reason to think that this claim is true unless some of the beliefs are initially credible – for some reason other than coherence, say, because they state facts of observation. (20)

Imagine that someone starts from an initial judgment that entails moral disapproval of homosexual relationships and consequently formulates moral principles that account for this immorality (e.g. the principle that only sexual relationships between men and women are morally acceptable). Checking this principle against relevant background theories might show that the principle in fact coheres with certain religious or evolutionary background theories. Although we obviously think that this judgment and the subsequent choices made in reflecting on this judgment that ultimately lead to a reflective equilibrium are unreasonable, it is unclear whether reflective equilibrium has the theoretical resources in order to rule out these kinds of equilibria, which are obviously unreasonable from our point of view.

The question is thus why we should believe that coherent homophobia is better justified than incoherent homophobia, and in fact, why coherent homophobia is justified at all. It seems that the method of reflective equilibrium cannot provide the resources to exclude these kinds of moral prejudices being brought into coherence in reflective equilibrium. Thus, as long as the input into the method of reflective equilibrium is not initially credible, the equilibrium arrived at after due reflection ‘may be no more than a reshuffling of moral prejudices’ (Brandt 1979, 22). This seems to threaten

---

11Although the use of background theories might also uncover that one’s disproval of homosexual relationships to some extent expresses a (contingent) cultural idea about sexual relationships, this does not have to lead to a revision of one’s initial judgment (are there initial moral judgments which do not, in one way or another, express certain cultural ideas?).
the idea that reflective equilibrium is an appropriate method to generate objective moral beliefs.

These two objections provide prima facie reasons to doubt reflective equilibrium’s ability to generate objective moral beliefs. In the next sections, I will discuss recent attempts of proponents of reflective equilibrium to respond to these kinds of objections, and argue that they have not succeeded in formulating a satisfactory answer – at least not one that is able to guarantee reflective equilibrium’s objectivist aspirations while at the same time remaining a distinctive, coherentist moral methodology.

4. Considered judgments

Proponents of reflective equilibrium typically point out that their moral methodology does have the theoretical resources to exclude unreasonable equilibria and to make convergence on a reflective equilibrium more likely than one might expect. Whereas the challenges mentioned in the previous section presuppose that *any* input is allowed into the method, most (if not all) versions of reflective equilibrium in fact limit the set of judgments that are allowed into the process of reaching a reflective equilibrium. Rawls pursues this first strategy when he limits the relevant set of judgments to, what he calls, *considered* judgments.

According to Rawls (1999a),

considered judgments […] enter as those judgments in which our moral capacities are most likely to be displayed without distortion […] For example, we can discard those judgments made with hesitation, or in which we have little confidence. Similarly, those given when we are upset or frightened, or when we stand to gain one way or the other can be left aside. All these judgments are likely to be erroneous or to be influenced by an excessive attention to our own interests. (42)

Rawls suggests that only those judgments should be allowed into the method of reflective equilibrium that are made under conditions that are favorable for making correct judgments. Rawls mentions two types of constraints on judgments that should guarantee that judgments have some initial credibility. First, we should be sufficiently certain about a judgment. This constraint excludes judgments in which we have little confidence. Second, Rawls mentions several psychological states – distress, fear, etc. – that are unfavorable to making reasonable judgments. Thus, only those judgments we make when we are ‘in the right state of mind’ are allowed to enter reflective equilibrium.

The idea here is thus that reflective equilibrium is able to guarantee objective moral beliefs because the starting points should have at least
some initial credibility. In order to prevent the outcome of reflective equilibrium being unreasonable or merely a set of moral prejudices, the set of judgments that is allowed in reflective equilibrium is limited. In addition, by limiting the set of judgments that is admitted into the process of reaching equilibrium, it might also become more likely that there is convergence on a set of principles, because a significant amount of moral disagreement is ruled out from the very start.

However, the use of considered judgments raises two problems. First, one might wonder whether limiting the input to considered judgments is actually necessary and sufficient for making reasonable judgments. That is, judgments made under unfavorable conditions do not necessarily seem to be unreasonable, and judgments made under favourable conditions do not necessarily seem to be reasonable judgments (cf Kelly and McGrath 2010, 349). It seems unlikely, for instance, that all individuals who agreed on the moral permissibility of slavery were not ‘in the right state of mind’. In addition, it might be possible that despite being uncertain or upset one can make a reasonable judgment – social and political activism, for example, often has its roots in the fact that people are upset. And it might well be that in times when homosexuality was commonly regarded as a sin, someone who judged the opposite was in fact hesitant about this judgment. Surely, we do not want to exclude these kind of progressive moral judgments from our moral methodology simply because people were upset or uncertain about their judgments.

In a recent article, Kelly and McGrath (2010) therefore propose to replace Rawls’ procedural notion of ‘consideredness’ with a more substantive characterization of which moral judgments are reasonable. Thus, according to them, we should first justify a set of reasonable moral judgments before we can use them as input in reflective equilibrium. In a similar way, Brink (1989) proposes to include not just moral beliefs that have been formed ‘under conditions of general cognitive reliability but also on the basis of an impartial and imaginative consideration of the interests of the relevant parties’ (132). It seems that by relying on a more substantive characterization of reasonable moral judgments, reflective equilibrium can indeed respond to the objections mentioned above, because only those moral beliefs are justified that cohere with other moral beliefs that have a certain epistemic authority independent of reflective equilibrium.

However, limiting the admissible input to either reasonable or impartial judgments has significant consequences for reflective equilibrium as a distinctive method of moral justification because arguably
the most interesting part of the story concerns not the pursuit of equilibrium itself, but rather what makes it the case that certain starting points are more reasonable than others, and how we manage to recognize or grasp such facts. (Kelly and McGrath 2010, 353–354)

This brings us to the second problem with the introduction of considered or reasonable judgments.

Even if one does not accept the conclusion that one needs more than merely procedural criteria in order to generate a reasonable equilibrium, any qualification of the starting points of reflective equilibrium – e.g. by referring to ‘considered’ or ‘reasonable’ judgments – that is substantive enough to guarantee objectivity needs a justification that cannot be provided by reflective equilibrium itself. With the introduction of limits on the input in reflective equilibrium, the method of reflective equilibrium thus becomes parasitic on some other method of justification, such as intuitionism.

Although this need for justification is most obvious in variants of the methodology which depend on a substantive characterization of the admissible input, Rawls’ procedural notion of consideredness is in similar need of justification. For one thing, Rawls’ criteria for considered judgments simply exclude egoism as a possible normative position. It therefore seems that his notion of consideredness implies certain substantive normative commitments that should be justified, instead of remaining implicit in his moral methodology.

Saving reflective equilibrium by qualifying the starting points thus undermines reflective equilibrium as a distinctive coherentist moral methodology, because it is no longer the coherence of all possibly relevant considerations but the reasonableness or consideredness of the input that bears most of the justificatory weight. Or at least, it makes it hard to see to what extent reflective equilibrium is still an alternative to other moral methodologies, because the idea that given certain starting points our belief set should be coherent seems to be compatible with any other method of moral justification. Introducing limits on the input in reflective equilibrium thus comes at the cost of undermining one of the most important features of the method of reflective equilibrium: the idea that there are no moral beliefs independent

12Rawls (1999a) seems to believe that the specific conditions he introduces are fairly uncontroversial. He claims, for instance, that ‘the criteria that identify these judgments are not arbitrary. They are, in fact, similar to those that single out considered judgments of any kind. And once we regard the sense of justice as a mental capacity, as involving the exercise of thought, the relevant judgments are those given under conditions favorable for deliberation and judgment in general’ (42). However, either the criteria are in fact uncontroversial, but in that case they are insufficient to do any interesting normative work. Or the criteria are substantial enough to guarantee objectivity, but in that case they are in need of justification.

13I agree with Kelly and McGrath (2010) when they conclude that ‘of course, once that move is made, one might very well wonder whether the picture of inquiry that emerges still deserves the name “the method of reflective equilibrium”’ (353).
of the procedure of reaching a reflective equilibrium that can determine its reasonableness.

Some proponents of reflective equilibrium, however, argue that the reliance on an epistemically privileged set of moral beliefs, e.g. considered judgments, does not undermine reflective equilibrium as a distinctive method of moral justification. Reflective equilibrium, so they argue, is best understood not as a coherentist methodology but as a ‘modest foundationalist’ method of moral justification (DePaul 1986; Holmgren 1987; Ebertz 1993; McMahan 2000). According to this interpretation of reflective equilibrium, certain beliefs have a privileged epistemic status independent of their coherence with other beliefs. This foundationalism is, however, ‘modest’ because these foundational beliefs are not immune to revision. For instance, if, on due reflection, one would discover that one’s considered judgments have been formed under conditions that are likely to import moral biases into one’s belief set, this would constitute a reason to revise the set of considered judgments.

Although a modest foundationalist interpretation of reflective equilibrium might be able to deliver objective moral beliefs, it comes at the cost of reintroducing intuitionism through the backdoor. Whereas reflective equilibrium was originally introduced in ethics as an alternative to intuitionism, a modest foundationalist interpretation of reflective equilibrium can be best understood as a form of intuitionism, because it accepts the metaphysical and epistemological commitments of intuitionism and only adds the qualification that our intuitions should be revisable. In other words, in the modest foundationalist interpretation, the most interesting part of the justificatory story is not the critical reflection on all relevant considerations, but the question of why certain beliefs have a special epistemic status.

Of course, it might turn out that a modest foundationalist interpretation of reflective equilibrium is to be preferred over a coherentist interpretation but in that case the proponents of reflective equilibrium will have to engage in the kind of epistemological and metaphysical questions they typically want to avoid. Thus, in so far as we want to understand reflective equilibrium as a truly distinctive methodology and as an alternative to intuitionism, we should focus on reflective equilibrium as a coherentist moral methodology. And insofar as it wants to remain a coherentist methodology it cannot deliver objective moral beliefs. Or so I have argued.

14 A qualification recent versions of intuitionism would be happy to accept as well (see, e.g. Audi 2005, 29).
15 An additional reason to reject a modest foundationalist interpretation of reflective equilibrium is that this modest foundationalism is clearly in conflict with constructivist interpretations of reflective equilibrium.
5. Scanlon on considered judgments

It seems that the method of reflective equilibrium can only guarantee objectivity by undermining the basic methodological commitments that set it apart from foundationalist alternatives such as intuitionism. But this conclusion comes too quickly. Recently, Scanlon (2014) has responded to the critique that the use of considered judgments makes reflective equilibrium parasitic on another moral methodology, and he argues that the inclusion of considered judgment is in fact not in tension with reflective equilibrium once we properly understand the way considered judgments are selected.

Scanlon agrees that we need to make a substantive evaluation about which judgments qualify as considered judgments in order for a reflective equilibrium to have any justificatory credentials. According to Scanlon, mere coherence is not enough to confer justificatory status on a set of beliefs, and there is thus no reason to accept homophobia simply by virtue of the fact that it could be part of a coherent belief set. Instead, Scanlon (2014) claims that

the justificatory force of the fact that we have arrived at certain judgments in reflective equilibrium depends on the substantive merits of the judgments we make along the way, in beginning with certain considered judgments and in modifying these judgments and others as we progress. (82)

Thus, he claims that ‘in deciding whether to count a belief among our considered judgments, the question we ask is whether it is something that it is reasonable to believe’ (81). In order to determine whether or not something is reasonable to believe we have to establish whether ‘it seems reasonable to believe this’ (81).

It might appear as if Scanlon, by relying on the reasonableness of judgments, reintroduces intuitionism through the backdoor. This suspicion is further strengthened by his claim that ‘it seems that we can discover normative truths and mathematical truths simply by thinking about these subjects in the right way’ (2014, 80). Drawing the analogy between normative truths and mathematical truths suggests that certain normative truths can be known a priori, through thinking about them in ‘the right way’. In addition, he claims that considered judgments are judgments ‘that seem clearly to be correct’, (77) and ‘that seem to me to be clearly true when I am thinking about the matter under good conditions for arriving at judgments of the kind in question’ (82). The question whether or not a judgment is a considered judgment then depends ‘on the truth of claims about which things are or are not reasons for action’ (Scanlon 2012, 236). These kinds of claims look
very similar to the kind of claims that intuitionists make about the way we arrive at true moral judgments.\textsuperscript{16}

Scanlon (2014), however, makes clear that reflective equilibrium ‘does not privilege judgments of any particular type – those about particular cases, for example – as having special justificatory standing’ (77). More specifically he claims that the reasonableness of judgments should not be determined \textit{independent from} engaging in reflective equilibrium, but that this question is itself part of the process of reaching reflective equilibrium: ‘[it] is not something separate from the method of reflective equilibrium but, as I have emphasized, a crucial part of carrying out that method’ (84–85). In other words, the question whether or not certain judgments are considered judgments is just another question that should be taken into consideration when engaging in reflective equilibrium. It is thus a mistake to think that reflective equilibrium’s reliance on considered judgments in any way undermines reflective equilibrium as a distinctive and autonomous moral methodology, at least once we properly understand the way considered judgments are to be selected.

Scanlon’s expansion of the scope of reflective equilibrium, to include questions about the reasonability of judgments, thus does not seem to be susceptible to the objections raised in the previous section, because it explicitly rejects the idea that we need a determinate set of considered judgments \textit{independent from} engaging in reflective equilibrium. This proposal does, however, have some problems of its own.

First, it is unclear how considered judgments can lead to objective moral beliefs if the question of what counts as a considered judgment is included in the search for reflective equilibrium. The reason for this is that there seems to be significant (normative) disagreement about the conditions that are conducive to making good moral judgments. For instance, for certain crude forms of utilitarianism, the conditions under which good judgments about morality are made are conditions of detached calculative rationality. For virtue ethicists, on the other hand, good judgments can only be made under conditions in which one is thoroughly involved in the practices about which a normative claim is being made. For a contractualist like Scanlon, the appropriate conditions are conditions of impartiality. It is therefore unclear how adding the question of what counts as a considered judgment to the elements that should be taken into consideration, can make it more likely that there is convergence on a reflective equilibrium, or that reflective equilibria that are obviously unreasonable from \textit{our} specific normative perspective can

\textsuperscript{16}There is, for instance, a striking similarity between Scanlon’s terminology and George Bealer’s (1996) theory of intuitions as a priori, intellectual seemings.
be ruled out. Of course there might be less disagreement about what counts as the conditions under which we make good judgments compared to disagreement concerning first-order normative claims, but it seems unlikely that this kind of reduction of normative disagreement on the level of what counts as a considered judgment is sufficient to safeguard maximal objectivity.

Second, Scanlon’s defense of reflective equilibrium is an instance of a more common expansionist tendency among defenders of reflective equilibrium to include just about any kind of disagreement – be it normative, meta-ethical or otherwise – into their methodology. When someone asks the proponent of reflective equilibrium to give an account of when judgments are reliable or reasonable, how one has to make choices when a conflict between different beliefs arises or which background theories are relevant when reflecting on the truth of a moral belief, the reply often is that these questions should themselves be settled in reflective equilibrium.

This kind of expansionism is problematic for several reasons. One problem is that by including just about any possible disagreement related to the justification of our moral beliefs into its methodology, reflective equilibrium runs the risk of becoming vacuous as a method of moral justification, because ultimately reflective equilibrium will simply be reduced to reasoning about ethics in general. That is, if any kind of disagreement is included in the search for reflective equilibrium it is not evident that it can still function as a method of moral justification.

A second problem is that this kind of expansionism undermines the method’s practicability (cf. Arras 2007, 55–56). One of the reasons for the popularity of reflective equilibrium, especially in the context of normative and applied ethics, is that it is a practicable methodology. Reflective equilibrium is not just a meta-ethical position, but it can actually be employed to make substantive normative judgments and to apply these judgments to actual practical problems. However, given the almost infinite number of considerations that should be included in reaching reflective equilibrium, it seems that the method will never be able to reach a determinate normative judgment. In other words, the more considerations one has to include the less practicable the method becomes.

6. The only game in town?

I have argued that the method of reflective equilibrium cannot deliver maximal objectivity, at least while remaining a distinctive, autonomous

17Another context in which this happens is in the context of possible biases that influence our normative judgments (see, e.g. Tersman 2008; Brink 2014, 55–56).
and practicable methodology. One might wonder whether the fact that a meta-ethical theory, e.g. reflective equilibrium, cannot deliver certain normative outcomes, e.g. moral objectivity, is by itself a reason to reject this methodology. I do not think that this follows, because it would assume that the truth of some moral claim (e.g. the existence of certain objective moral beliefs) is an argument for the appropriate method to justify moral beliefs. This is to put the cart before the horse. Here, I am therefore in agreement with Scanlon (2003) who claims that ‘the fact that the method of reflective equilibrium could lead to a result that called into question the objectivity of our moral beliefs is not an objection to that method’ (153).

However, one implication of the current discussion is that we should be more aware of the limited normative potential of the methodology: moral objectivity might be beyond reach and the outcome of its procedure might be no more than a reshuffling of moral prejudices. It is important to note that this does not mean that ‘anything goes’ when adopting the method of reflective equilibrium, because, depending on the exact way one spells out the requirements of coherence, moral beliefs can be shown to be unacceptable in light of one’s larger belief set. The method of reflective equilibrium thus does allow for the idea that someone can be mistaken about a moral belief and offers at least some resources for the critical revision of one’s belief.

However, as I have argued above, on a widely accepted view of morality (and one that is shared by most proponents of reflective equilibrium), moral beliefs are those beliefs one should accept independent of the contingent beliefs or desires one happens to have (whether or not these beliefs and desires are internally coherent). It is this type of maximal objectivity that reflective equilibrium cannot account for. This suggests that proponents of reflective equilibrium cannot have it both ways: reflective equilibrium either leads to moral relativism, or it turns out to be a form of intuitionism in disguise.

Ultimately, proponents of reflective equilibrium typically suggest that we should adopt reflective equilibrium, even if this means giving up on moral objectivity. For instance, in a recent paper, Michael DePaul (2006) claims that there simply is no reasonable alternative to the method of reflective equilibrium, claiming that it is natural to want certainty or something else that is more secure than what seems true upon reflection. But we are not destined to have such things, at least not for very much of what we believe. The best we can do is think things through and trust the conclusions we reach. The defender of reflective equilibrium calls

---

18As I have noted above there are different ways to flesh out the idea of maximal objectivity (depending on whether or not one is a moral realist) but this should not concern us here.
it like it is and says what is good enough. No more hankering after what cannot be. (618)

In a similar way, Walden (2013) recently claimed that

it is the only conception of normativity that we are left with once we discover that a kind of absolutism is impracticable. It is the answer we are left with when we find that we cannot divine the future or use some inner sense to tell right from wrong (254).

Scanlon (2003) claims that ‘it is the only defensible method: apparent alternatives to it are illusory’ (149). Finally, Jonathan Floyd (forthcoming) claims that ‘the ultimate defence of Rawls’ method […] is that unless we can construct an alternative, together with a convincing argument regarding its superiority, we should just “keep calm and carry on” (12).

The implicit assumption of proponents of reflective equilibrium is thus that we have to choose between either foundationalism, more specifically intuitionism or the coherentism of reflective equilibrium. Given the known problems with intuitionism we should adopt reflective equilibrium, even if this means giving up on moral objectivity. Or so the argument seems to go. This argument, however, overlooks an important alternative moral epistemology.¹⁹ Besides reflective equilibrium and intuitionism, a third possible method of moral justification is a so-called transcendental argument for objective moral principles. In contrast to intuitionism, transcendental arguments are seldom discussed by proponents of reflective equilibrium.²⁰ At the same time there is a growing interest in transcendental arguments in meta-ethics, but this debate unfortunately does not (yet) have any implications for the methodology of normative ethics.

A transcendental argument is an argument that starts from an inescapable or unquestionable feature of our self-understanding and consequently explores the necessary conditions of possibility of this specific feature of our self-understanding. More specifically, a transcendental argument is an argument that tries to show that a commitment to X is a necessary condition

¹⁹Another problem with the argument is that it does not provide a positive reason to adopt reflective equilibrium. That is, it is unclear why we should adopt reflective equilibrium and not say global skepticism about morality, which denies that there are any justified moral beliefs at all.

²⁰Recent examples of transcendental arguments in ethics include Gewirth (1978), O’Neill (1986), Habermas (1990), Korsgaard (1996), Illies (2003) Darwall (2006). In addition, there is a growing literature on constitutivism in meta-ethics, which could be understood as one particular form a transcendental argument in ethics might take. For a good overview see Tubert (2010).

²¹There are a few exceptions: Rawls (1999b) briefly discusses the alleged failure of Kant’s transcendental deduction of the categorical imperative. More recently, Scanlon (2012) claims that he is ‘not convinced by any argument I have seen for the claim that we must see requirements as binding on us insofar as we see ourselves as acting at all’ (238). However, both fail to systematically engage with recent attempts to provide transcendental arguments in ethics, nor do they discuss the respective strengths and weaknesses of transcendental arguments and the method of reflective equilibrium.
for the possibility of Y – where, given that Y is inescapable, it logically follows that one is necessarily committed to X.\textsuperscript{22}

In the context of ethics, the starting point of transcendental arguments (Y) is typically a certain conception of action or interaction, and the transcendental commitment (X) includes a universal and categorical principle.\textsuperscript{23} Transcendental arguments in ethics aim to show that we are necessarily committed to accepting certain moral principles insofar as we understand ourselves as agents or insofar as we are engaged in certain types of interaction. In other words, instead of looking for standards of moral evaluation independent from ourselves (intuitionism), or for the coherence between our contingently held moral and non-moral beliefs (reflective equilibrium), transcendental arguments aim to make explicit those moral principles that we necessarily have to accept given certain inescapable aspects of our self-understanding. A successful transcendental arguments in ethics would thus vindicate a practical conception of moral objectivity, because transcendental commitments do not depend on the contingent preferences or projects of individuals, but simply on their self-understanding as agents or their being engaged in certain forms of interaction.

Christine Korsgaard (1996), for instance, argues that, given the fact that we have to understand ourselves inescapably as reflective beings with reasons for action, we necessarily have to value our own humanity and (in a subsequent step of her argument) also the humanity of others (for a similar view see Gewirth 1978). Arguing from a slightly different, communicative conception of agency, O’Neill (1986) argues that agents are necessarily committed to the principles presupposed by the public use of reasoning, including principles of toleration, non-injury and non-coercion (for similar views see Habermas 1990; Darwall 2006).

Transcendental arguments can neither be reduced to the coherence-ism of reflective equilibrium nor to foundationalism, or at least in the way these approaches are standardly understood in moral epistemology.\textsuperscript{24} Transcendental arguments are not foundationalist because, unlike intuitionism, they do not rely on self-evident, i.e. non-inferentially justified, moral beliefs. That is, transcendental approaches to morality reject the idea of self-evidence and instead try to provide a reason for accepting moral principles.

\textsuperscript{22}This is a slightly revised version of the definition of transcendental arguments put forward by Stern (2013).

\textsuperscript{23}In the context of ethics, the transcendental commitment is a practical commitment (i.e. a principle of action) and not a doxastic commitment (e.g. a justified belief, or a true belief). Transcendental arguments in ethics are therefore not susceptible to Barry Stroud’s (1968) objections to transcendental arguments in ethics. For a discussion of the difference between transcendental arguments in theoretical philosophy and in practical philosophy see Stern (2013).

\textsuperscript{24}I am thankful to an anonymous reviewer for pressing me on this point.
Although transcendental arguments could be understood as a form of coherentism, it is important to note that they diverge significantly from the specific type of coherentism of reflective equilibrium. Reflective equilibrium focuses on the coherence of our contingently held belief set, including, for instance, judgments, principles and background theories. Transcendental arguments, by contrast, try to show that there are certain moral commitments that any agent, simply by virtue of being an agent, necessarily has to accept, independent from the contingent content of his or her specific beliefs or desires. A transcendental argument focuses on the principles that one cannot deny on pain of contradicting that one is an agent or that one is engaged in certain forms of interaction. The specific criterion of coherence in transcendental arguments is thus the absence of self-contradiction, and not coherence simpliciter.

Transcendental arguments constitute an attractive alternative to both intuitionism and reflective equilibrium, because transcendental arguments promise to justify objective moral principles (pace reflective equilibrium), without relying on foundationalist premises (pace intuitionism). Because transcendental arguments are part of the larger coherentist family of which reflective equilibrium is also a member, proponents of reflective equilibrium have no a priori reason to reject transcendental arguments. In addition, because transcendental arguments aim to justify objective moral principles, proponents of reflective equilibrium have all the more reason to take this alternative moral methodology serious before concluding that reflective equilibrium is the only game in town.

7. Concluding remarks

In this paper, I have critically evaluated the most recent attempts to defend the method of reflective equilibrium against common objections to the methodology. I have argued that proponents of reflective equilibrium still fail to show that their methodology can safeguard moral objectivity. In addition, I have argued that reflective equilibrium is not the only game in town once intuitionism has been discarded. I have concluded that given their own normative ambitions, combined with their rejection of intuitionism, proponents of reflective equilibrium have reason to take transcendental arguments more seriously than they have done so far.

25Walden (2013, 252–254) fails to make this distinction between reflective equilibrium and transcendental arguments when he discusses Korsgaard’s method of reflective endorsement as a variant of reflective equilibrium. Korsgaard makes clear that she does think that moral justification is constituted by the reflective endorsement of our contingent beliefs (our ‘practical identities’). Instead, she argues moral justification is constituted by whatever one necessarily has to accept, independent from the specific content of one’s practical identities (1996, 120–123).
Of course, these transcendental lines of argumentation are themselves far from uncontroversial and should be fully developed in order to constitute a plausible alternative to both intuitionism and reflective equilibrium. But transcendental arguments should at least be considered as a possible alternative moral methodology. In fact, transcendental arguments might be our best hope of securing moral objectivity.

Acknowledgments

Earlier versions of this paper were presented at the MANCEPT Seminar Series (2014), the study group on Meta-Ethics of the Dutch Research School of Philosophy (2014), the annual conference of the Dutch Research School of Philosophy (2014), the PhD Forum on Practical Philosophy at the Norwegian University of Science and Technology in Trondheim (2015) and the Political Philosophy Seminar at the London School of Economics (2015). I am grateful for the comments and questions I received on these occasions. Special thanks to Rutger Claassen, Marcus Düwell, Fleur Jongepier, Ingrid Robeyns and Miriam Ronzoni for their helpful comments on various earlier drafts of this paper. I would also like to thank an anonymous reviewer for constructive comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the author.

Funding

This work was supported by the Netherlands Organisation for Scientific Research [317-20-010].

References


26Recent criticisms of transcendental arguments in ethics question both the inescapability of the starting point of the argument (i.e. action or interaction) and the potential of constructing substantive moral commitments on the basis of a transcendental argument. For the inescapability objection see Enoch (2006, 2011). For the emptiness objection see Street (2012) and Tiffany (2012). I think these objections can be answered but I cannot go into these issues here as they are beyond the scope of this paper.


