Undergraduate Philosophy Journal of Australasia www.upja.online
Vol. 3, December 2021

The Method of Wide Reflective Equilibrium and Self-evidence

Tiago Carneiro da Silva[1]

Federal University of Rio de Janeiro

Abstract

In this essay, my ultimate aim is to show that the method of wide reflective equilibrium (MWRE) can be improved in a way that allows us to detect self-evident propositions in a reasonably effective way. In order to do this, I first argue that appealing to self-evidence does not have to be considered a dogmatic approach in the search for moral justification. I do this while describing characteristics of self-evidence that are worth considering in devising a moral methodology. This allows us to see how the search for self-evident propositions may be compatible with the MWRE. I then defend that the method is not as radically opposed to the appeal to self-evidence as it has commonly assumed. When doing this, I argue that the MWRE is more effective in leading us to find self-evident beliefs than one might initially expect. Finally, based on some features self-evident beliefs have, I propose that, in addition to following the steps that the MWRE requires us to follow, we should meet two further requirements in order to detect self-evident propositions in a more effective way. Furthermore, the resulting methodological proposal, I argue, can be desirable even if there happens to be no self-evident propositions.

1. Introduction

Is there a correct, or best, method for us to acquire justified moral beliefs? If so, what method would that be? A very popular candidate is called "the method of wide reflective equilibrium" (henceforth, "the MWRE"), as it was named by John Rawls in his *The Independence of Moral Theory*.[2] This method essentially tells us to reflect on our beliefs about issues of a given area of enquiry while considering relevant theories

[1] *Tiago Carneiro has recently completed his undergraduate degree in philosophy at the Federal University of Rio de Janeiro, Brazil. He is currently looking to pursue a master's degree and carry out research on moral intuitions.*
[2] Rawls 1974.

ISSN: 2653-3146

and arguments in a way that leads us to form a coherent set of beliefs.[3] This end point of reflection, achieved after revising our beliefs in light of these relevant considerations, is what is called "wide reflective equilibrium".[4]

The method is so popular that some of the defenders of the method even hold that there is no other reasonable alternative to it.[5] They argue that one of the advantages of the method in relation to other possible alternatives is that it is not dogmatic. This is because the method takes every belief to be revisable. In this sense, its proponents generally hold that it is opposed to the idea that we can acquire justified moral beliefs by appeal to self-evidence apart from the application of the method. Appealing to the method commits one to the idea that there are unrevisable beliefs — and this, they claim, is a dogmatic approach to morality.[6] According to them, the only way we could appeal to self-evidence non-dogmatically is if we reach this conclusion *after* applying the MWRE, as it takes every belief to be revisable (even the belief that there are self-evident propositions), and those that accept this possibility believe that this is unlikely to happen.[7]

Rejection of self-evidence seems to come at a cost, though. First, to reject it is to abandon a firm ground for moral justification. It does not seem, for example, that one can be justified in having the belief that it is right to enslave black people even if this belief were part of a completely coherent set of beliefs or if it were acquired after reflecting on all of one's beliefs and alternatives to them. Self-evidence, on the other hand, seems to provide such a firm ground. Second, denying self-evidence requires one to explain why some specific beliefs seem to be justified independently of coherence or of what merely seem true to us. Taking an example from Cuneo and Shafer-Landau, it seems that one who adequately understands the sentence that says "for beings like us, in worlds like ours, the interests of others can be morally weightier than our own" and believes it on the basis of this understanding would be justified in believing it.[8] This belief does not have to be part of a coherent set of

[3] Here, I shall focus on methodological issues in the moral domain.
[4] The "wide" in "wide reflective equilibrium" indicates that a wide range of types of proposition is taken into consideration when one applies this version of the method.
[5] Smith 1994, pp. 40– 41; Scanlon 2003, p. 149; DePaul 2006, p. 618; and Walden 2013, p. 254. DePaul, for example, argues that we are not destined to have firmer grounds for "very much of what we believe " than the intuitions we have after having carefully taken all relevant views into consideration, which is what the MWRE tells us to do. Therefore, we cannot reasonably appeal to, say, self-evidence independently of the application of the method. Indeed, we should do our best to reflect rigorously about morality, but the most we can do is to follow the steps the method requires us to follow. In addition, other less rigorous alternatives would be too permissive.
[6] Daniels 1979, p. 264–267; Brink 1989, p. 8–9; Ebertz 1993, p. 213; Rawls 1999, p. 19.
[7] Rawls 1974; Daniels 2016.
[8] Cuneo & Shafer-Landau 2014.

ISSN: 2653-3146

beliefs, nor is it justified only because it seems true to us. Rather, this seems to be a special type of belief because its content is probably self-evident. [9]

Having these *prima facie* reasons to accept the existence of self-evident propositions in mind, we have at least some good reasons to attempt to devise a method that effectively detects them. However, although I do think that applying the MWRE is not the best way to detect them, I believe the method is a good starting point for us to devise a better procedure in that respect. To show why this is so, I will argue in favour of three main theses:

(i) The appeal to self-evidence is not intrinsically dogmatic. This, if correct, would show that dogmatism does not have to be a problem for those who hold that self-evident propositions exist in ethics. Additionally, this will show applying the MWRE and appealing to self-evidence are compatible in an important respect, as both approaches would be shown not to be dogmatic.

(ii) There are good reasons for preserving the steps the MWRE requires us to follow so that we can acquire self-evident moral beliefs and be self-evidentially justified, which is important for my third thesis.

(iii) The likelihood for us to acquire self-evident moral beliefs is increased if we, in addition to following the steps required for us to apply the MWRE, meet further requirements that are based on the conditions for being self-evidentially justified. Doing this, I will argue, is better than only applying the MWRE as it is traditionally conceived.

If these three theses are correct, the ultimate goal of this essay can be established: we can improve the MWRE as a method for acquiring justified beliefs (and moral knowledge). If there are self-evident propositions (as we have at least some reasons to think so) and we apply this resulting method, we will more effectively detect moral truths and gain moral knowledge. Increasing our chances of doing this is a

[9] There are other arguments against self-evidence other than the one that says that appealing to it is to commit oneself to some sort of dogmatism. One such argument is that of Brink, which assumes that being justified in believing a proposition *p* requires one to have a higher-order belief that expresses the proposition that *p* is a type of belief *T* and beliefs of type *T* are justified (Brink 1989, pp. 116-122). If this is so, it would not be possible to be non-inferentially justified, as we would be if we were self-evidentially justified. However, as Shafer-Landau later persuasively replied, this assumption is problematic for at least two reasons (Shafer-Landau 2003, pp. 252-253). First, there does seem to be beliefs that do not require higher order beliefs to be justified, such as simple perceptual beliefs. Second, it threatens to start a vicious infinite regress, as being justified in having a second-order belief would require one to have a third-order belief, and so on. Since we do not have an infinite number of meta-beliefs of the required kind, none of our beliefs would be justified, including the belief that Brink's assumption is correct.

result we can expect from a better method, since increasing our chances of acquiring moral knowledge is desirable for a method of moral justification. However, even if there happens to be no self-evident moral propositions, the proposal I offer for enhancing the MWRE based on features of self-evidence may, I will argue, still be worth accepting.

2. Self-evidence and Dogmatism

Before dealing with the compatibility between the MWRE and self-evidence, it is useful to show that appealing to self-evidence does not have the disadvantage that proponents of the method generally hold that such appeal has — namely, that of being dogmatic. This will pave the way for the discussion of how the MWRE can be applied in order to lead us to be self-evidentially justified. This shall be clear soon enough.

I take Robert Audi's account of self-evidence to provide a very plausible epistemology.[10] It captures the essence that is thought to be present in classical examples of propositions that are candidates for self-evidence — namely, that they are true propositions which can be known *a priori* and non-inferentially (a knowledge that, as Ross says, requires a "certain mental maturity" to be achieved). Also, it provides a detailed explanation of the role that reason has in apprehending those truths: that of leading us to understand them adequately.[11] This way, he proposes that self-evident propositions would be those true propositions that (i) if adequately understood by *S*, *S* is justified in believing them on the basis of that understanding, and (ii) if believed in such a way, S knows them. Based on this, three common misconceptions about self-evidence may be dispelled. First, notice that these features of self-evidence do not entail that *S will believe* a self-evident proposition *p* once *S* adequately understands it. If *S* understands *p*, even if adequately, this does not mean that *S* will immediately see that *p* is true.[12] Second, adequately understanding *p* does not mean that *S* will necessarily take it to be self-evident (*S* may not even know what self-evidence is or may even deny that self-evident propositions exist). Third, this account of self-evidence does not entail that there is no way of inferring *p* from premises.

Contrary to classical intuitionist philosophers such as H.A. Prichard, W. D. Ross and G.E. Moore, this account of self-evidence allows us to defend that self-evident

[10] Audi 2015, pp.65-66; Audi 2018, pp. 1-3.
[11] This explains why disbelieving a strong candidate for self-evidence may very well be a product of inadequate understanding. Take, for example, the oft-cited proposition that all bachelors are unmarried. If someone denies it, we have strong reasons to think that the person in question either does not understand at least one of the concepts in the proposition or does not understand the conceptual relation between them.
[12] This may happen, but not necessarily so.

propositions "need not be unprovable, need not be obvious, and need not be rationally beyond dispute".[13] This provides a reply to criticisms against classical intuitionists that charge them for being dogmatic due to their defence of self-evidence. Consider Audi's example that 'if *p* entails *q*, and *q* entails *r*, and *r* entails *s*, and *s* is false, then *p* is false'. This may be obvious to some people (maybe to those who have experience in dealing with logic), but it is certainly not obvious to

most people. To see that such a proposition is true might require one to draw an inference from '*p* entails *q*, and *q* entails *r*, and *r* entails *s*' to the conclusion that '*p* entails *s*', which may lead one to see more clearly that if *s* is false, then *p* is false.[14] Nevertheless, although this inference might be a way to adequately understand it, it is *this understanding* that grounds the self-evidential justification. Also, these steps might be a way to show to another person who does not adequately understand it that the long proposition is true. However, if one is able to show these inferential steps to another person, one may still justifiably believe the long proposition solely on the basis of adequately understanding it. [15]

In fact, self-evident propositions might even sound counterintuitive. Another example of Audi's is that "a child can be born by its grandmother". At first sight, at least, it might seem to *S* that such a proposition is not true. It may take *S* to think of a specific case in which a mother has a child with her son for *S* to understand it appropriately and see its truth. One who knows the story of Oedepus Rex, who had children with his mother unknowingly, may see this more easily than a person who does not. In any case, it is possible that, when one is not thinking about such cases, one does not find that proposition intuitive even after accepting its truth.

Given these considerations, some possible explanations for why people may deny self-evident propositions may be offered. People tend not to accept counterintuitive propositions. This way, people may be less prone to accept self-evident propositions that sound counterintuitive at first. Self-evident propositions that involve complex relations among the concepts that figure in them do not tend to be obvious and may require imagination or application of inferences for one to be able to understand them adequately. Moral propositions are arguably more likely to be harder to be adequately apprehended, as they involve more complex concepts and relations

[13] Audi 2015, p. 66.
[14] Notice that this is an *internal* inference; that is, this inference has a premise that is supplied by the very proposition that is the conclusion of the inference (the long self-evident proposition). That is why this inference is only one way of adequately understanding the long proposition, rather than a conclusion of an argument that has as premises propositions other than those that are supplied by it. In any case, it is possible to infer self-evident propositions from "external premises". However, one is not required to draw *this kind* of inference in order to be justified in having self-evident beliefs.
[15] For one to adequately understand it, one needs to understand it to a degree that is possible to see that it is true, and being able to explain a self-evident proposition is evidence that that person has this degree of understanding. Hence, it is plausible to defend that adequately understanding a proposition involves being able to explicate one's understanding of it.

among them than the examples I gave.[16] Self-evident propositions are truths that can be apprehended *a priori*, but that does not mean that they have to be analytical.[17] They are propositions we can see to be true by reflection on the relations among their constituents alone, even though we may need sense experience to learn the concepts involved.[18] What we do before we "see" that a self-evident proposition is true is to correctly apprehend the relations among the concepts that are part of it. If these relations are complex, as they seem to be in moral propositions, the harder it is to understand them adequately — and the less likely it is for them to be obvious.[19] A candidate for self-evidence which might be an example of a complex moral truth is the proposition expressed by Sidgwick's principle of justice:

> It cannot be right for *A* to treat *B* in a manner in which it would be wrong for *B* to treat *A*, merely on the ground that they are two different individuals, and without there being any difference between the natures or circumstances of the two which can be stated as a reasonable ground for difference of treatment.[20]

It may be hard to understand many propositions adequately, as it involves more than mere semantic comprehension. A bilingual person may be able to translate the sentence 'a child can be born by its grandmother' correctly after comprehending it

---

[16] The examples I gave might seem to indicate that self-evident propositions are only analytical or logical truths. However, they are only straightforward ways to demonstrate how self-evident propositions may be provable or counterintuitive. If these kinds of self-evident propositions may be so, there is good reason to think that self-evident moral propositions may also be (that is, if there really are self-evident propositions — be they moral or non-moral).

[17] One example of a proposition that is probably self-evident and not analytical is the proposition that nothing can be red and green all over at the same time. "Not green" is not plausibly part of the *meaning* of "red". Therefore, this proposition would not be true in virtue of its meaning. The same would be true for at least most moral propositions.

[18] It is important to note that the dependence on experience is limited to the acquisition of concepts. The truth of the proposition that faulty electrical plug sockets tend to cause fires, for example, can only be justifiably believed empirically. If this proposition is what is meant by the sentence "faulty electrical plug sockets can cause fires", then this sentence is not self-evident. If, on the other hand, what is meant by it is that it is conceptually possible that faulty electrical plug sockets can cause fires, then this might very well be self-evident. However, this does not seem to be what we mean when we say such a sentence, as when we say it we have the intention of warning against having faulty electrical plug sockets, as they are *likely* to cause fires. The example that a child can be born by its grandmother, on the other hand, does not express tendencies that can only be empirically established; it is only a proposition that contradicts something that seems impossible (that is why it is probably a self-evident proposition which is more interesting than the one that expresses that it is conceptually possible that faulty electrical plug sockets can cause fires).I thank an anonymous referee for pressing me on this point.

[19] Hence, a method that allows us to identify self-evident propositions accurately needs to take into account how adequately we understand them. This is what my proposal in section 4 does.

[20] Sidgwick 1907, p. 380. Notice that this is a *candidate* for self-evidence, and it may even be a false proposition. However, given that this seems to be another way of stating that moral properties supervene non-moral ones, which has been defended by philosophers of several different backgrounds, this seems to be a good candidate. If it is indeed self-evident, denial of this proposition might be explained by

semantically. But for this person to understand it adequately, more reflection will be necessary. In fact, there may be many "comprehensional variables", which may account for at least part of the denial of self-evident propositions.[21] These "variables" are the factors that determine how well a person understands a concept or proposition. One such variable concerns the ability to translate a sentence from one language to another. Another concerns the ability to identify instances of propositions, such as identifying that the wrongness of John killing Mary for fun instantiates the proposition that torturing people for fun is wrong.[22]

Having these factors that may lead us to deny self-evident propositions in mind, it should be clear by now that we are not infallible in detecting self-evident beliefs. Moreover, these considerations are only part of the story: not only can we deny self-evident propositions, but we can also accept as self-evident what is *not* self-evident. We may believe that a proposition is self-evident due to cultural, or even genetic, influences.[23] As self-evident propositions are objectively true, neither our culture nor our genes determine what propositions are self-evident.[24] Apart from this, cognitive biases also affect our moral judgments and are possibly a cause for our acceptance (or rejection) of some moral propositions which contradict self-evident propositions.[25] Thus, it seems to be desirable to find ways of diminishing these influences in our moral thinking.

It is important to note that, although we may not believe a self-evident proposition *p* even after appropriately understanding it, understanding *p* adequately does give a rational person a disposition to believe it. Thus, we should expect rational people to have such disposition after we conclude that they have appropriately understood *p*. This does not mean, though, that people who adequately understand *p* will necessarily be thinking irrationally if they deny *p*. One can rationally be committed to moral scepticism or to a theory that denies a self-evident proposition. Such a scepticism or theory may provide *some* justification for denying it. To the extent that

[21] For a more detailed account of what "adequate understanding" amounts to, see Audi's 2018 article.
[22] I will later mention some other comprehensional variables which might be useful to take into consideration when trying to detect self-evident beliefs in a more effective way.
[23] There are arguments that appeal to evolution to criticise moral realism, which indirectly attack the classical appeal to self-evidence (see, for example, Street 2006 and Joyce 2006). It is possible, though, to limit the range of the target of that kind of argument to only a certain type of moral belief. One example of how this can be done can be seen in Lazari-Radek & Singer 2014, pp. 174-199. In particular, they target only those intuitions that merely increased the likelihood of reproductive success, which, they argue, are intuitions that are not impartial (egoistic intuitions and those that place more importance on our kin or group for their own sake). Analogously, debunking arguments of such a limited range may be based on considerations regarding cultural differences.
[24] This way, self-evident propositions cannot clash or contradict each other. Hence the importance of appealing to a method that tell us to aim for coherence, as does the MWRE.
[25] See, for example, studies which show evidence of bias in moral thinking, such as Tversky & Kahneman 1981, p. 453; Petrinovich & O'Neill 1996; and Haidt & Baron 1996. Although the first of these studies is not specifically about bias in moral thinking, it does show evidence of that type of bias.

they provide such a degree of justification, those who deny it based on such reasons are not irrationally denying it.

Accepting that one can rationally deny self-evident beliefs and that one is not infallible in detecting self-evident propositions is a stance that can hardly be considered dogmatic. If one accepts that there are self-evident propositions, one can hold such a view and does not have to take oneself to be infallible in detecting them. Moreover, having the factors that may lead one to fail to detect self-evident moral propositions in mind will help us to try to get around the problems we might face in the search for self-evidence. In what follows, I will explore how the well-known MWRE may help us to detect self-evident propositions and how its basic structure can lead us to devise a better method.

3. The Method of Wide Reflective Equilibrium and Self-evidence

The MWRE tends to be presented as a method that conflicts with the idea that there are self-evident propositions. Some say that, after applying the method, what does (or most likely does) the justificatory work in the final system of beliefs is coherence. [26] Other philosophers say that only other forms of "weaker" foundations have justificatory power in it.[27] Both these positions oppose the idea that applying the method would lead us to find strong non-inferential foundations for moral justification or knowledge, such as self-evident propositions. [28]

In contrast to these views, I believe that the MWRE can not only be compatible with self-evidence (which would mean that coherence is not the only factor that does the justificatory work) but also be more effective in detecting self-evident propositions than one might expect. To argue in favour of this, I will first present a general characterisation of the method, as it is to be applied in moral philosophy, and then attempt to show how it could lead us to be self-evidentially justified.[29]

---

[26] See, for example, Brink 1989, Rawls 1974 and Daniels 2016.

[27] For example, although Ebertz and McMahan defend that what does the justificatory work in systems of moral beliefs is not coherence, they do think that the method should be understood as a "modest foundationalist model" of justification which does not appeal to self-evidence (Ebertz 1993; McMahan 2000).

[28] Recall what I said in the introduction regarding Rawls and Daniels's positions on the improbability of us acquiring beliefs that serve as self-evident foundations for all other beliefs of the relevant domain. This would be improbable because when we apply the method, we reflect on our beliefs in light of all other relevant beliefs. We also revise them in light of them. This, they argue, increases the likelihood that we will only acquire inferential beliefs.

[29] The method can be applied in many domains, such as epistemology, logic and normativity in general. Indeed, one of the first characterizations of the method was described by Goodman, who wanted to apply it in the domain of logic (Goodman 1953). After Goodman, Rawls offered an influential characterization of the method, defending that we should employ it in order to theorize about justice (Rawls 1971). Daniels, then, proposed an even more detailed description of the method and is still recognized for having offered the most complete account of the method (Daniels 1979).

The MWRE can be divided into four or five stages, depending on what account of the method we are talking about.[30] Below, I describe it as having five stages:

(i) We first identify our initial moral beliefs – the ones we have when starting to apply the method. These beliefs may be either particular or general. That is, they may be moral beliefs about particular cases (e.g., that it was wrong for John to lie) or very general abstract beliefs (e.g., that we should aim at the good on the whole).

(ii) Then we select from that set of initial moral beliefs some beliefs which are simultaneously (a) those in which we are confident; (b) those that are not likely to be influenced by egoism or strong emotions (e.g., desires regarding only one's own welfare, fear or anger); (c) those that are likely to be stable in the face of scrutiny; and, at least if we follow Rawls's remarks about considered judgements in a 1951 paper,[31] (d) those that have not been derived consciously from the application of principles. Those beliefs are called "considered moral judgements".

(iii) After having selected the considered moral judgements, we try to find principles that account for, or systematize, those judgements. These principles should be ones which would lead the person who is applying the method to accept, by only using these principles as premises in conjunction with propositions that describe relevant situations, the same considered moral judgements previously selected. [32]

(iv) Instead of just making adjustments in our system of beliefs in order to match principles with considered judgements (which, if successful, would lead to a coherent system of beliefs called "narrow reflective equilibrium"), we also take into account moral and non-moral theories (as well as arguments in favour and against them) that might be relevant to the moral domain.

(v) By reflecting on everything that has been taken into consideration so far, we revise our initial considered judgements, principles, and moral and morally

[30] The second of the steps described is omitted by some philosophers. There are at least two: Goodman 1953 and Lewis 1983.
[31] Rawls 1951.
[32] For example, for 'it is wrong to eat meat' (call it 'P') to be a principle of this kind, P in conjunction with sentences describing relevant facts such as 'Peter ate meat' should lead to a conclusion that is itself a considered judgement, such as 'it was wrong for Peter to eat meat'. Notice that this is a condition for a statement to be considered the type of principle that *figures in this step of the method*. Remember that it is possible for a considered judgement to be just as general as, or even more general than, the principle that I presented as an example.

> relevant non-moral theories in order to achieve a final coherent system of beliefs. This endpoint of reflection is called "wide reflective equilibrium".

The MWRE is the version of the method that is considered to lead us to have justified moral beliefs. On the other hand, the method of *narrow* reflective equilibrium (*i.e.*, the one which does not tell us to take into consideration things such as moral theories, arguments in the literature, etc.) is not taken to be relevant to the debate of moral justification[33]. Indeed, this simpler method just tells us to systematise beliefs we had before applying it — and this does not seem to be a very promising way of acquiring systems of beliefs that are systematically justified. This is why in this essay I am focusing on the method of wide reflective equilibrium.

To point out that the method *itself* does not deny that there are self-evident propositions, it is important to notice that the description of the method does not tell us what makes beliefs justified. Therefore, it does not tell us whether there are propositions such that one would be justified in believing them solely on the basis of adequately understanding them. However, I do not only think that the method is neutral as to the existence of self-evident propositions; I also believe that we can acquire *self-evidential* justification when applying the method — that is, that we can acquire self-evident beliefs on the basis of adequately understanding them. Nevertheless, one could counter this view with the following argument:

> *P1* – We are only self-evidentially justified if we believe a self-evident proposition based on our appropriate understanding of it.
> *P2* - The method of reflective equilibrium only leads us to have beliefs in a purely inferential way — and not based on our understanding of them.
> *C* – The method does not lead us to be self-evidentially justified. [34]

I do not think that *P1* is problematic.[35] The problem with the argument, though, is

[33] Rawls 2001; Lazari-Radek & Singer 2014; Daniels 2016. Kamm defends that a version of the method of narrow reflective equilibrium reveals some underlying psychological structure of our moral thinking (Kamm 1993, p. 8). It might seem at first that Kamm is an exception to this trend. However, as Kamm herself recognizes, attempting to reveal this structure is a *descriptive* task; not a *justificatory* one.

[34] Some clarifications might be appropriate. First, notice that it allows that it is possible to be self-evidentially justified *simpliciter*. It allows that one could be self-evidentially justified by other means. For example, one could believe a self-evident proposition on the basis of adequately understanding it when one is applying the method of narrow reflective equilibrium. Second, it concedes that one can acquire self-evident beliefs when applying the method. That is because one can have a self-evident belief p even if one does not believe p based on one's understanding of p. For example, one can believe it solely on the basis of argumentation from other premises present in the system of belief. Being self-evidentially justified is different from having self-evident beliefs.

[35] Indeed, it is part of the very definition of self-evidential justification that one is self-evidentially justified only if one believes a self-evident proposition based on an adequate understanding of it.

P2. Although the MWRE is a dialectical procedure, this does not mean that every belief in reflective equilibrium is acquired in a purely inferential way. First, as I mentioned before, one may draw inferences in order to *adequately understand* a proposition, and this does not mean that one accepts it on a purely inferential basis. It is the *understanding* that grounds justification; not the inference that helped one to understand the proposition in question. Second, it is possible for a belief to be a *product* of an inference without it being an *inferential* belief, even if such an inference was not necessary for an adequate understanding of it. Consider the case in which I do not believe the self-evident proposition *p*, even after obtaining an adequate understanding of it, because I believe the following inference:[36]

P1 - If *r*, then ¬*p*
P2 – *r*
C – ¬*p*

This way, I do not believe *p* even though *p* is self-evident. However, my reason for believing ¬*p* may be undermined if afterwards I believe ¬*r* based on an inference such as:

P1 – If *w*, then ¬*r*
P2 – *w*
C - ¬*r*

As *p* is self-evident, I tend to believe *p* based on my adequate understanding of it. Furthermore, I do not see a reason to doubt it anymore, as I no longer believe *r*.[37] I then come to believe *p* without *p* being a conclusion of an argument. All that has happened is that I do not believe that there is a defeater for *p* anymore, as the last inference made me change my mind about the truth of *r*. Therefore, my belief that *p* was a product of an inference; however, such a belief is not inferential. This kind of reasoning may occur quite regularly when one applies the MWRE. Especially when we consider philosophical arguments, such as the ones we may encounter when applying the method, we may come to believe certain sceptical theses which lead us to deny propositions that are self-evident. However, these sceptical theses may be undermined as well, which increases the likelihood that we will believe the self-evident propositions these theses denied.

[36] Recall that adequately understanding a proposition does not mean that one *will* believe it. There is only a *tendency* to believe it, which may be overridden by some other opposing belief that happens to be more appealing to the person in question.
[37] Recall what I said about self-evidential justification: if one adequately understands a self-evident proposition, one tends to believe it.

Given what has been argued above, I believe we can conclude that the MWRE not only is compatible with the search for self-evident propositions, but also compatible with self-evidential justification. It may be argued, though, that it is unreasonable to endorse the method in order to acquire self-evidential justification, as the MWRE is not effective in leading us to detect self-evidence, nor in leading us to adequately understand self-evident propositions.

Regarding this, I do concede that it is not the best method for acquiring self-evident beliefs and for being self-evidentially justified.[38] Nevertheless, I do think that it is more effective in leading us to find self-evident propositions than one might expect, and we could make use of the features of the method in order to detect them effectively. The first reason for this is that it is plausible that at least some of our considered judgements are self-evident or have as content propositions that are similar to some self-evident propositions.[39] Thus, as the MWRE tells us to aim for coherence, applying the MWRE will probably lead us to filter out the judgements that most obviously do not cohere with self-evident propositions.[40] If so, this could at least diminish the probability that we will hold too many beliefs that more clearly deny self-evident propositions, such as beliefs that state that torturing for fun is morally good.[41]

Another reason to think that the MWRE's ability to lead us to find self-evident propositions is not so low is related to the mistaken assumption that they must be very general. This assumption seems to be based on the idea that self-evident propositions are unprovable. However, self-evident beliefs *can* be provable. This means that there may be general propositions (some of which may also be self-evident) that can serve as premises to prove them. Consider some examples offered by Terence Cuneo and Russ Shafer-Landau that might be good candidates for self-evidence:[42]

a) For beings like us, in worlds like ours, it is *pro tanto* wrong to engage in the recreational slaughter of a fellow person.
b) For beings like us, in worlds like ours, it is *pro tanto* wrong to humiliate others simply for pleasure.

[38] That is exactly why I am proposing that we add further requirements than the ones the MWRE tells us to meet.
[39] Recall the characteristics of considered judgements from page 10. As we tend to believe self-evident propositions after having adequately understood them, they may very well be selected as considered judgements.
[40] Rawls denies that considered judgements are self-evident, and I agree that, as such they are not (Rawls 1999, p. 507). That does not mean, though, that some of them may be.
[41] To be sure, even if this is how it happens to be, it would certainly still be possible for one to hold unreasonable beliefs.
[42] Cuneo & Shafer-Landau 2014, p.7.

c) For beings like us, in worlds like ours, it is *pro tanto* wrong to torture others just because they have inconvenienced you.

Part of ethical theorizing might be to find out how propositions such as these can be proved. This, however, allows that one can be justified in believing them on the basis of adequately understanding them (in case they are indeed self-evident). Maybe one of the premises that may be used to prove them can be something like Sidgwick's proposal that 'the good of any one person is no more important from the point of view (if I may put it like this) of the universe than the good of any other' (call it "the principle of benevolence").[43] If this is the case, it can be said that there are more self-evident propositions than several philosophers believe there are, as they think that the view that defends self-evidence must assume that self-evident propositions are only unprovable axiomatic propositions.[44] As there would be more self-evident propositions to be discovered, it is more likely that we will discover them by employing the MWRE than it would be if there were only axiomatic self-evident propositions.

To be sure, as it is likely that there are more self-evident propositions than it was assumed in the past, this makes virtually *every* method more likely to detect them. What reasons do we have to pick the MWRE instead of any other method to try to detect them? One reason for choosing the method is that it tells us to consider many moral views that have been accepted by many reflective people. As I mentioned before, these reflective people probably thought carefully about the propositions they accepted, and thus were more likely to adequately understand them and, consequently, find them intuitive and defend them. A second reason is that it tells us to consider many particular intuitions, which may help us to adequately understand self-evident propositions that we may encounter, and maybe even serve as evidence in favour or against certain self-evident principles.[45] Third, it tells us to consider arguments which may count in favour or against the reliability of certain intuitions. Fourth, although it not merely tells us to form coherent systems of beliefs, the fact

[43] To be sure, maybe even none of them is self-evident — these are only some plausible examples that I am using to show how there may be less general self-evident propositions that can be proved by more general ones. In any case, both this proposition and the less general ones seem to be propositions that one would be justified in believing solely on the basis of adequately understanding them. To be sure, understanding them in such a way does not mean that one *will* believe them; however, *if* one believes them on such a basis, it is plausible to say that one will be justified. That is why they are candidates for self-evidence.

[44] Influential intuitionists who have explicitly defended such a view are Moore, H.A. Pritchard and Ross (Moore 1903, p. x; H.A. Prichard 1912, p. 29-30; and Ross 1930, p. 29). Rawls and Daniels seem to have defended it as well, and this may be a reason why they thought it was so unlikely that one would find beliefs that could be self-evident when applying the MWRE (Rawls 1974, p. 8; and Daniels 2016). Brink, a defender of the MWRE, holds more explicitly that self-evident propositions are unprovable (Brink 1989 p.112-113).

[45] For example, if these more particular intuitions tend to be unconscious applications of self-evident principles.

that it tells us to pursue coherence is an important feature which may make it more likely for us to find self-evident propositions.[46]

For these reasons, I do not think that we should abandon the general steps of the procedure. However, for us to increase its effectiveness in leading us to find self-evident beliefs, I do think that there ought to be additional steps. In the next section, I will make a general outline of what kind of steps we can add.

4. A Better Method for Detecting Self-evident Beliefs

The procedure I will propose in this section preserves all the steps one has to take in order to apply the MWRE. However, I believe that a better method should have more requirements than the MWRE (as it is traditionally defended) imposes. One of them is related to adequately understanding candidates for self-evidence. After having considered and reflected on many moral judgements when one applies the MWRE, one is likely to have found reasonable candidates for self-evidence and may be even self-evidentially justified regarding some propositions. If some of these candidates are self-evident, and one adequately understands them, one will tend to believe them.[47] This gives us reason to try to identify the propositions we think we and others adequately understand and tend to believe them on the basis of that understanding. These will be candidates for self-evidence. Once these candidates for self-evidence have been identified, one could apply an additional procedure that filters out candidates that do not satisfy the characteristics we expect self-evident propositions to have. The procedure I have in mind is based on the following reasoning:

> *P1* – If *S* adequately understands proposition *p* and *p* is self-evident, *S* tends to believe it on the basis of adequately understanding it.

Now, imagine we empirically confirm the following:

> *P2* - *S* adequately understands *p* and does *not* tend to believe it on the basis of adequately understanding it.

In that case, we could conclude, then, that *p* is not self-evident. This reasoning allows us to exclude some candidates for self-evidence, and based on this I propose the first

[46] If we tend to believe self-evident propositions on the basis of understanding them, and if we try to make our system of beliefs more coherent, coherence may lead us to find other self-evident beliefs that are coherent with the ones we hold.
[47] And it will be more likely for us to be self-evidentially justified.

additional methodological requirement to be followed after following the steps the MWRE tells us to follow:

Additional methodological requirement 1- One should either empirically test or look for studies that empirically test whether people adequately understand candidates for self-evidence and believe them on the basis of that understanding.[48] The test may yield different results and we should adopt different positions accordingly. It may be found that:

(i) Subjects who adequately understand a candidate for self-evidence satisfy either (a) or (b) below:

(a) They do *not* tend to believe the candidate for self-evidence on the basis of adequately understanding it. In that case, we should not consider it self-evident (or at least consider it a bad candidate, as the empirical experiment may not be perfect).
(b) They tend to believe it on the basis of adequately understanding it. In that case, other things being equal, the proposition in question should remain a candidate for self-evidence and our confidence that it is self-evident may be increased for having passed the test.

It may also be found that:

(i) Subjects who do not adequately understand a candidate for self-evidence satisfy either (a) or (b) below:

(a) They do *not* tend to believe a candidate for self-evidence on the basis of adequately understanding it. In that case, we have a reason to think why they deny it even though it may be self-evident (*i.e.*, they may deny it because they do not adequately understand it).
(b) They tend to believe it. In that case, we have reason to think that it may be a proposition that is not self-evident even though people do tend to believe it.

It could be tested, based on this requirement, for example, whether people who adequately understand Sidgwick's principle of benevolence tend to believe it on the

[48] What I propose is essentially an application of experimental philosophy that helps us to discard candidates for self-evidence that we would not otherwise discard. It might also provide reasons for accepting candidates for self-evidence that have passed this test. This way, my position fits within both the so-called "negative" and the "positive" programmes, as this type of empirical studies may have as effect the rejection of some candidates for self-evidence as well as the vindication of other candidates.

basis of adequately understanding it.[49] If they do not, this indicates that it is probably not self-evident.[50] If they do, other things being equal, we may still take it to be a candidate for self-evidence. If no one, or if no significant number of people, adequately understand it, more studies designed to make subjects better understand the principle of benevolence and/or more studies with a higher number of subjects should be conducted.[51]

Notice that we cannot ascertain without a doubt that such a candidate for self-evidence is *indeed* self-evident. It is possible that it is not true (which would imply that it is *not* self-evident) even if it passes the test. However, if the test has worked and a proposition has not passed the test, it can be safely concluded that the candidate for self-evidence is *not* self-evident.[52] Discarding candidates for self-evidence this way makes it more likely that we will hold self-evident beliefs, as there will be less wrong candidates for us to accidentally hold.

Admittedly, this is a very complex procedure. And this is not even the whole story: there are many factors we must consider in order to check whether someone adequately understands a candidate for self-evidence. Audi, for example, argues that for one to understand a proposition adequately, one must satisfy *nine* "comprehensional variables". Recall that adequately understanding a proposition does not merely involve semantic understanding. It also involves, for instance, being able to respond adequately to questions related to what follows or does not follow from the proposition in question, being able to explicate one's understanding of it, being able to identify instances of what the proposition states *etc.* All these factors, even though they are many, should be taken into account if we want to identify self-evident beliefs in a more effective way.

The fact that it is too complex may lead one to reject such a procedure. It might be argued, for instance, that people are justified in having moral beliefs without having to apply such a procedure. This, the objection goes, is sufficient for us to just dismiss the procedure. I think this would be a bad objection, though. The point is not that people will *only* be justified if they follow this procedure.[53] Indeed, if we accept that there are self-evident propositions, one *is* justified in believing them solely on the basis of adequately understanding them. There are even other ways one can be

[49] Recall that this principle states that the good of any one person is no more important "from the point of view of the universe" than the good of any other.
[50] That is, if it were self-evident, people *would* tend to believe it.
[51] The relevant result is seen only when subjects adequately understand the candidates for self-evidence.
[52] If a candidate for self-evidence passes this test, though, this may be a further reason for accepting it as self-evident, because it has not been discarded as self-evident *and* because it is still a candidate for self-evidence on reflection.
[53] Here, I am trying a different approach than philosophers such as Rawls, Brink and Daniels seem to have tried, in that the method I am proposing is not necessary for justification, whereas they seemed to defend that the MWRE *is* necessary.

justified in having a belief. However, if we can diminish the likelihood that we will take certain propositions to be self-evident when they are not or diminish the likelihood that we will *not* take self-evident propositions to be self-evident, I believe it is reasonable to do so.

Science, for example, has developed very complex methods to find out things about the world. Yet, we should not dismiss them because they are complex. Accepting the methods of science also does not commit us to not accepting the conclusions we arrive via less complex methods. I am justified in believing there is a computer in front of me even though such a belief is grounded merely on my seeing it. I do not have to wait for science to tell me whether I really am seeing it in order to be *prima facie* justified.[54] Like science, though, my proposal may lead us to increase the degree of justification of our systems of moral beliefs.

A further question that may arise is the following: why use an empirical procedure to find out whether people tend to believe a proposition instead of trying to discover our tendencies to believe it *a priori*? The answer is that the statistical apparatus that can be used in empirical studies are more powerful than our introspection when it comes to detecting patterns. The former would be more sensitive to detect the relevant factors that lead us to believe certain propositions than the latter. By appealing to the type of experimental data I am defending we should pursue, we could detect tendencies of belief or disbelief that are influenced by factors other than an appropriate understanding of propositions and select candidates for self-evidence more reliably.

This last point leads us to another consideration: cognitive biases may affect people's judgement and, consequently, lead them to either tend to believe propositions that are *not* self-evident non-inferentially or tend *not* to believe propositions that *are* self-evident even when one understands them adequately. This way, the studies we should appeal to in order to satisfy this additional requirement I am proposing must control for cognitive biases. They should consider whether, for example, the instances of self-evident principles subjects identify when they are being tested for adequacy of understanding are triggering biases. Controlling for, say, order effects would be desirable in such studies.[55]

Controlling for biases such as these in types of studies that are different from the ones I have mentioned may also be useful for the search of self-evident propositions. A subject may deny a self-evident proposition because he or she thinks that an

[54] Science may increase my degree of justification by confirming, say, that I am not hallucinating, but my being justified is not determined only by appealing to science.
[55] "Order effects" occur when the mere difference in the order of sentences subjects consider lead to differences in the acceptance of such propositions.

intuition he believes speaks against it, and this intuition may be biased. This allows us to propose another additional requirement.

Additional requirement 2: One should either test or look for studies that test whether people are biased when they have a particular intuition that contradicts a particular candidate for self-evidence. The test may yield different results and we should adopt different positions accordingly:

(i) If subjects are biased, our confidence that the candidate for self-evidence is self-evident should be maintained.

(ii) If subjects are not biased, our confidence that the candidate for self-evidence is likely to be self-evident should, other things being equal, diminish.

Consider Sidgwick's principle of benevolence again, which may be denied because of, say, the intuition that the good of our kin and of other people from the group we are part of is intrinsically more important than the good of other people. In that case, we should test or look for studies that test whether this intuition (or some other intuition) that leads people to deny Sidgwick's principle is a product of bias.[56] If it is biased, then, other things being equal, the principle should remain a candidate for self-evidence.

To be sure, we should reflect on the findings and conclusions drawn from each kind of study and aim at coherence.[57] In any case, even if there are no self-evident propositions, a method that meets these two requirements I proposed still seems better than the MWRE as traditionally defended. A method that requires us to empirically see whether the moral beliefs we hold non-inferentially tend to be accepted once they are adequately understood may help us to see whether we are accepting moral beliefs without understanding them properly. The way the MWRE is traditionally conceived does not require this and is, therefore, less likely to yield this desirable result. And a method that requires us to see whether we tend to accept certain moral beliefs for reasons we do not find morally relevant may lead us, as individuals, to be more cautious in accepting them. This also seems to be an advantage in and of itself, regardless of whether you are a moral realist.

[56] Maybe we could find out that this intuition is a product of evolutionary forces that do not tend to select beings for their ability to apprehend *a priori* truths. However, this is only one way such intuition could be biased. If there are other possibilities of bias, empirical studies should try to detect them. If all reasonable possibilities of bias have been investigated, our confidence that Sidgwick's candidate for self-evidence is indeed self-evident should, other things being equal, be diminished. This is indeed a difficult task, as detecting bias is difficult — let alone different types of bias. However, if this can be, it should.

[57] Studies that meet the two requirements at once may even be more effective in detecting self-evident propositions and, ultimately, leading us to have moral knowledge.

5. Conclusion

In this essay, I first tried to show that appealing to self-evidence is not a dogmatic move. This opens one door to taking self-evidence more seriously and considering the MWRE as a way to find self-evident propositions. Then, I argued that the method is more effective in leading us to acquire self-evidential justification than we might initially expect.

Being able to acquire self-evidential justification by applying the steps required by the MWRE, as traditionally conceived, makes us capable of selecting candidates for self-evidence. I argued, then, that we can filter out candidates for self-evidence by meeting the additional requirements I proposed. This could help us to decide which propositions to favour when we reflect on our considered judgements, principles, and moral and morally relevant non-moral theories.[58]

In the end, even if self-evident propositions do not exist, these additional requirements still seem to be desirable in the search for moral justification. We do not have to be moral realists to defend this. Not accepting these additional requirements might, then, not only run the risk of being further away from possible foundations of morality, but also lead us away from having more justified moral beliefs even according to a non-realist perspective. [59]

[58] Admittedly, many more details could be added to these requirements. For example, details of what kind of tests could be applied in the empirical studies in order to identify whether subjects adequately understand the propositions in question.

[59] I would like to thank, for their invaluable comments, the two referees and the member of UPJA's editorial team who have reviewed my essay. Also, thanks to my supervisor, Antonio Saturnino, for all the discussions we had during the time I did my research on the method of reflective equilibrium and intuitionism; and to my colleague, Pedro Freire, for the very helpful comments and conversations on an earlier version of this essay.

References

Audi, R. (2015). 'Intuition and Its Place in Ethics.' *Journal of the American Philosophical Association.* Vol. 1. pp. 57–77. Cambridge University Press.

Audi, R. (2018). 'Understanding, Self-evidence and Justification.' Philosophy and Phenomenological Research, vol. 99, Issue 2, pp. 358-381. United States of America: Wiley-Blackwell.

Brink, D.O. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Cuneo and R. Shafer-Landau (2014) 'The Moral Fixed Points: New Directions for. Moral Nonnaturalism.' *Philosophical Studies.* Vol. 171(3). pp. 399-443

Daniels, N. (1979). 'Wide Reflective Equilibrium and Theory Acceptance in Ethics.' *The Journal of Philosophy*, vol. 76, no. 5, pp. 256-282. New York: Journal of Philosophy, Inc.

Daniels, N. (2016). 'Reflective Equilibrium.' *The Stanford Encyclopedia of Philosophy.* Retrieved February 28, 2021 from https://plato.stanford.edu/entries/reflective-equilibrium/

De Lazari-Radek, K.; Singer, P. (2014). *The Point of View of the Universe: Sidgwick and Contemporary Ethics.* Oxford: Oxford University Press

DePaul, Michael R. (2006). "Intuitions in Moral Inquiry." *The Oxford Handbook of Ethical Theory*, edited by David Copp, 595–623. Oxford: Oxford University Press.

Ebertz, R. (1993). 'Is Reflective Equilibrium a Coherentist Model?' *Canadian Journal of Philosophy*, vol. 23, no. 2, pp. 193-214, Cambridge University Press.

Goodman, N. (1955). *Fact, Fiction, and Forecast*. Cambridge, MA: Harvard University Press.

Haidt, J., & Baron, J. (1996). Social roles and the moral judgement of acts and omissions. *European Journal of Social Psychology*, Vol. 26(2). pp. 201–218

Joyce, R. (2006). *The Evolution of Morality.* Cambridge: MIT Press.

Kamm, F. M. (1993). Morality, Mortality, Vol. I: Death and Whom to Save From It. New York: Oxford University Press.

Lewis, D. (1983). *Philosophical Papers*, Vol. I. Oxford: Oxford University Press.

McMahan, J. (2000). "Moral Intuition." *The Blackwell Guide to Ethical Theory.* Edited by LaFollette Hugh, 92–110. Oxford: Blackwell.

McMahan, Jeff. 2000. "Moral Intuition." *The Blackwell Guide to Ethical Theory*, edited by LaFollette Hugh, pp. 92–110. Oxford: Blackwell.

Moore, G.E. (1903). *Principia Ethica*. London: Cambridge University Press.

Petrinovich, L., & O'Neill, P. (1996). Influence of wording and framing effects on moral intuitions. *Ethology & Sociobiology*, Vol. 17(3). pp. 145–171.

Prichard, H.A. (1912). 'Does moral philosophy rest on a mistake?'. *Moral Obligation*. Oxford: Oxford University Press.

Rawls, J. (1974-5). 'The Independence of Moral Theory'. *Proceedings and Addresses of the American Philosophical Association*, vol. 48, pp. 5–22. American Philosophical Association.

Rawls, J. (2001). *Justice as Fairness: A Restatement*. Cambridge, Mass.: Belknap Press of Harvard University Press

Rawls, J., (1999). *A Theory of Justice*, 2nd edn. Cambridge, Mass.: Belknap Press of Harvard University Press.

Ross W.D. (1930). *The Right and the Good.* Oxford: Oxford University Press.

Scanlon, T. M. (2003). 'Rawls on Justification.' In The Cambridge Companion to Rawls, edited by S. Freeman (ed.), pp. 139-167. Cambridge: Cambridge University Press.

Sidgwick, H. (1907). *The Methods of Ethics*, 7th edn. London: Macmillan.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell Publishers.

Street, S. (2006). 'A Darwinian Dilemma for Realist Theories of Value.' *An International Journal for Philosophy in the Analytic Tradition.* Vol. 127, No. 1. pp. 109-116. Springer.

Tversky, A.; Kahneman, D. (1981). 'The Framing of Decisions and the Psychology of Choice'. *Science*, Vol. 211. pp. 453–8.

Walden, Kenneth. (2013). 'In Defense of Reflective Equilibrium.' Philosophical Studies, vol. 166, no. 2, pp. 243–256.

ISSN: 2653-3146