

Reflective Equilibrium and Counterexamples

Without specifying what exactly they mean by it, philosophers in all areas of research frequently use the term ‘reflective equilibrium’ when they mention the methods and aims of their inquiry. It has been suggested that ‘reflective equilibrium’ is trivial.¹ However, reflective equilibrium has been extensively discussed *as a method* (MRE in what follows) in normative ethics and political philosophy. The aim of this paper is to explore whether MRE is as trivial as it seems. I mention how Rawls and other proponents of MRE conceived of MRE as non-trivial. My focus lies on comparing the method to our practice of debating thought experiments as counterexamples to philosophical theories in non-normative (non-moral and non-political) philosophy. I argue that MRE does not correspond with our practice, which shows that the method is not trivial.

1 An Account of Reflective Equilibrium

The key idea of MRE as introduced by John Rawls ([8], [9]) is that we test our sets of moral judgments and moral principles against each other and revise and refine both when they are inconsistent with other principles or judgments.² I will focus on Daniels’ ([1], [2], [3]) adaptation of Rawls’ original account of MRE. Daniels’ account can be given in several steps. In step 1, we choose a set of considered judgments amongst our initial moral judgments. In step 2, we determine sets of moral principles that fit best with the set of considered judgments. In step 3, we determine alternative sets of principles that are not directly supported by the considered judgments. In step 4, we look for background theories that deliver arguments for or against the competing sets of principles. In step 5, we finally use background theories and considered judgments to figure out the best and most coherent set of principles by adjusting judgments, principles, and background theories.

I will argue that step 5 contains the two following methodological claims.

¹E.g., DePaul [4], Foley [5, p. 117-131], Scanlon [10, p. 149], Williamson [12, p. 244-246].

²In the literature, two versions of MRE have been distinguished, *narrow* and *wide* MRE (Norman Daniels [1] finds this distinction implicit in Rawls [8] and explicit in Rawls [9]). I focus on wide MRE, and I will use ‘MRE’ for wide MRE only.

Equal Weight. The subject engaged in MRE assigns equal epistemic weight to intuitions and principles.

Coherence First. In the case of a conflict, we regain coherence either by dropping our intuition or by adjusting our principles.

2 Applying Reflective Equilibrium

I will apply Daniels' account of MRE to the Gettier Cases and compare the method with how the cases have in fact been discussed.

Step 1. Sophie is an epistemologist who has been thinking about knowledge and justification for a long time. In the year 1963, she gets to read Edmund Gettier's paper 'Is Justified True Belief Knowledge?' as well as Michael Clark's response 'Knowledge and Grounds: A comment on Mr. Gettier's Paper'. Sophie has the intuition that Smith does not have knowledge in one of the Gettier Cases. According to the theory of knowledge that Sophie holds, the JTB theory, Smith has knowledge. Hence, the content of Sophie's intuition is inconsistent with what follows from the JTB theory. Sophie strives to resolve it with the help of MRE.

Step 2 and Step 3. Sophie spends a long time thinking about the cases and about Clark's no-false-lemma reply to Gettier. Sophie also comes up with her own theories, of which the first is a defeasibility analysis of knowledge that is similar to what Lehrer and Paxon suggested much later.³ A further theory Sophie comes up with contains a simple reliability condition.⁴ Sophie decides that Clark's adjusted JTB theory fits best with her intuitions in the Gettier Cases (step 2). Both the defeasibility account and the reliability condition are *more general* than necessary to cover the Gettier Cases. The JTB theory, the defeasibility account and the reliability account are alternative theories that Sophie has to consider (step 3).

Step 4. Relevant background theories concerning a theory of knowledge obviously include theories of justification, theories about the accessibility of our knowledge and of the truths about the world. Let us stipulate that Sophie has internalist intuitions and theories about justification and knowledge.

Step 5. Sophie has to remove the inconsistency by applying Equal Weight and Coherence First to choose a theory. Sophie comes to the conclusion that Clark's analysis fits her intuitions best, but it is not systematic enough because it does not cover cases that are very similar to the ones presented by Gettier. The defeasibility analysis Sophie comes up with is more general, but she has to admit that is also extremely complicated. The reliability account clearly contradicts Sophie's internalist intuitions. Sophie decides

³Lehrer and Paxon [7, p. 230].

⁴See, e.g., Goldman [6].

that the alternatives to the JTB theory are unsatisfactory. According to Coherence First, Sophie ought to drop her intuition and endorse the JTB theory. She concludes that the JTB theory is, after all, the best choice.

Most epistemologists did not think that holding on to the JTB theory was a good idea.⁵ Surely, Sophie's case is fictional and could have been told differently, but I will show that there is an underlying pattern that explains why a philosopher who applies MRE is more likely to stick with the JTB theory than one who does not apply MRE.

3 Our Current Practice

I will show that in our current practice, we take the fact that we have an intuition to be the relevant *prima facie* evidence for its content and against our theory. We then move on to further investigations into the fact that we have an intuition and into the truth or falsity of the content of the intuition. We do not simply disregard the content of the intuition in the sense that it is treated as false in our theory. We rather give an explanation of the fact that we have an intuition by appeal to something different from the truth of the content. Only if we have an alternative explanation, we legitimately disregard its content. As methodological advice, the following two claims seem adequate, given our current practice. We ought to

1. search for the best-fitting theory that makes the content of the intuition true and
2. search for an explanation as to why we have the intuition that is not based on the truth of the content of the intuition.

If it is correct that our current practice does not correspond with MRE, it shows that MRE is not trivial.

References

- [1] DANIELS, N. Wide reflective equilibrium and theory acceptance in ethics. *The Journal of Philosophy* 76, 5 (1979), 256–282.
- [2] DANIELS, N. On some methods of ethics and linguistics. *Philosophical Studies* 37, 1 (1980), 21–36.
- [3] DANIELS, N. Reflective equilibrium and archimedean points. *Canadian Journal of Philosophy* 10, 1 (1980), 83–103.

⁵Weatherson argues that because of these qualities, the JTB theory is still the best trade-off and hence 'is still a live possibility, even considering Gettier Cases'. Weatherson [11, p. 11].

- [4] DEPAUL, M. R. Why bother with reflective equilibrium? In *Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry*. Rowman and Littlefield Publishers, Inc., 1998, pp. 293–309.
- [5] FOLEY, R. *Working Without a Net: A Study of Egocentric Epistemology*. Oxford University Press, New York, 1993.
- [6] GOLDMAN, A. What is justified belief? In *Justification and Knowledge: New Studies in Epistemology*. Reidel, Boston, Dordrecht, London, 1979, pp. 1–23.
- [7] LEHRER, K., AND PAXON, T. Knowledge: Undefeated justified true belief. *Journal of Philosophy* 66 (1969), 225–237.
- [8] RAWLS, J. *A Theory of Justice*. Cambridge University Press, Harvard, 1971.
- [9] RAWLS, J. The independence of moral theory. *Proceedings and Addresses of the American Philosophical Association XLVII* (1974/75), 5–22.
- [10] SCANLON, T. M. Rawls on justification. In *The Cambridge Companion to Rawls*, S. Freeman, Ed. Cambridge University Press, Cambridge, 2006, ch. 3, pp. 139–167.
- [11] WEATHERSON, B. What good are counterexamples. *Philosophical Studies* 115 (2003), 1–31.
- [12] WILLIAMSON, T. *The Philosophy of Philosophy*. Oxford University Press, 2007.