

INTUITIONS, BIASES, AND EXTRA-WIDE REFLECTIVE EQUILIBRIUM

SAMUEL DIRECTOR

It seems that intuitions are indispensable in philosophical theorizing. Yet, there is evidence that our intuitions are heavily influenced by biases. This generates a puzzle: we must use our intuitions, but we seemingly cannot fully trust those very intuitions. This paper develops a methodology for philosophical theorizing that attempts to avoid this puzzle. Specifically, it develops and defends a methodology it calls Extra-Wide Reflective Equilibrium. It argues that this method allows us to use intuitions, while also providing a mechanism to check the influence of bias on our intuitions. In section 1, it defends the claim that intuitions are indispensable in philosophical theorizing. In section 2, it outlines recent arguments against the reliability of intuitions. In section 3, it explains and defends its account of Extra-Wide Reflective Equilibrium.

Keywords: bias, intuitions, philosophical methodology, reflective equilibrium

It seems that intuitions are indispensable in philosophical theorizing. Yet, there is evidence that our intuitions are heavily influenced by biases. This generates a puzzle: we must use our intuitions, but we seemingly cannot fully trust those very intuitions. In this paper, I develop a methodology for philosophical theorizing that attempts to avoid this puzzle. Specifically, I develop and defend a methodology that I call Extra-Wide Reflective Equilibrium. I argue that this method allows us to use intuitions, while also providing a mechanism to check the influence of bias on our intuitions.

Before continuing, I should clarify that I do not argue that my proposed methodology *solves* the aforementioned puzzle; rather, I argue that my method can help to *lessen* the effects of biases on our intuitions, without eliminating them altogether. In that regard, my claim is fairly modest. At the very least, I intend this paper to be mainly an exploration of a possible way of doing philosophy that reduces the influence of bias. I certainly do not think that this methodology is a foolproof solution to this problem; it is merely an interesting proposal for how it might be mitigated.

In section 1, defend the claim that intuitions are indispensable in philosophical theorizing. In section 2, I outline recent arguments against the reliability of intuitions. In section 3, I explain and defend my account of Extra-Wide Reflective Equilibrium.

1. The Indispensability of Intuitions in Philosophy

Historically, there are two views about the nature of intuitions. On the classical view, an intuition is a proposition that is self-evident or basic and is known by reflection.¹ On the contemporary view, intuitions are pre-reflective experiences, or snap judgments. In line with this view, Huemer defines an intuition in the following way: “The way things seem prior to reasoning . . . [is] an ‘initial appearance.’ An initial, intellectual appearance is an ‘intuition.’ That is, an intuition that p is a state of its seeming to me that p that is not dependent on inference from other beliefs and that results from thinking about p” (2005, 102). On this contemporary conception, I have an intuition that P when P seems to me to be true. In this paper, I discuss only the contemporary conception of intuitions. It is this understanding of intuitions that I defend, and it is this understanding that dominates the field today.

In the rest of this section, I defend the claim that intuitions (as defined above) are indispensable in philosophical theorizing. In other words, we cannot do philosophy without using our intuitions. My argument for the indispensability of intuitions is based on the claim that any denial of intuitionism is ultimately self-defeating. The following argument establishes this conclusion.²

P1: intuitionism holds that when it seems to S that P, S has a prima facie reason to believe P.

P2: if S is to deny intuitionism, then S must do so on the basis of how things seem to her.

P3: if S forms her belief in the denial of intuitionism on the basis of how things seem to her, then S must believe that seemings or intuitions confer justification.

P4: if S believes that seemings or intuitions confer justification, then S believes that intuitionism is true.

C1: thus, in order to deny intuitionism, S must first affirm intuitionism (from P2, P3, P4).

C2: thus, S’s denial of intuitionism is self-defeating and should be rejected (from C1).

Let me now defend each premise individually, beginning with P1.

P1: intuitionism holds that when it seems to S that P, S has a prima facie reason to believe.

As Huemer puts it, intuitionism is defined by the following principle, which he calls the Principle of Phenomenal Conservatism: “If it seems to S that p, then in the absence of defeaters, S thereby has at least some degree of justification for believing that p” (2007, 30).

P2: if S is to deny intuitionism, then S must do so on the basis of how things seem to her.</>

Suppose that S is an anti-intuitionist philosopher who thinks that intuitions do not justify beliefs. As such, she seeks to deny the truth of intuitionism. Premise 2 claims, however, that, in order to deny intuitionism, S must rely on the way that things seem to her. This is true because, in general, we form the vast majority, if not all, of our beliefs on the basis of the way that things seem to us.

P3: if S forms her belief in the denial of intuitionism on the basis of how things seem to her, then S must believe that seemings or intuitions confer justification.</>

Simply put, if S bases her denial of intuitionism on how things seem to her, then S must believe that the way things seem to her (her intuitions) can justify her beliefs.

P4: if S believes that seemings or intuitions confer justification, then S believes that intuitionism is true.</>

Intuitionism just is the claim that seemings confer justification on beliefs; given this, by her own actions, S believes that intuitionism is true.

C1: thus, in order to deny intuitionism, S must first affirm intuitionism (from P2, P3, P4).</>

Given the argument so far, if S is to deny the truth of intuitionism, she must do so on the basis of how things seem to her, which means that she takes seemings as being able to justify

beliefs, which means that she affirms intuitionism.

C2: thus, S's denial of intuitionism is self-defeating and should be rejected (from C1).

In short, S can only deny intuitionism by also believing in intuitionism; since this is incoherent and self-defeating, the denial of intuitionism ought to be rejected. Of course, objections can be made to this argument, but due to considerations of space I will have to leave the matter as is. From this, I conclude that intuitions, understood as intellectual seemings, are indispensable for philosophical theorizing.³

2. Challenges to the Use of Intuitions

Philosophers and psychologists have recently argued that our intuitions may be unreliable and shaped by bias. This challenge has come from a wide variety of sources. Here, I briefly highlight a few well-known strands of this argument.

First, the literature in experimental philosophy has shown that, when nonphilosophers are asked about classical philosophical thought experiments, they give widely divergent answers, suggesting that intuitions are sometimes tracking something other than the truth and may be swayed by irrelevant factors. For example, Buckwalter and Stich (2014) show that, in introductory philosophy classes, women and men give widely divergent answers when asked to respond to certain classical philosophical thought experiments. Furthermore, the experimental philosophy literature has shown that folk intuitions are subject to both framing and ordering effects.⁴

Second, the literature on implicit biases suggests that our intuitions may be swayed by biases without us even knowing it.⁵

Third, philosophically inclined psychologists and neuroscientists have argued that many of our intuitions are biased by our evolutionary history. For example, Greene has developed an entire research program that seeks to show that certain kinds of moral intuitions (namely, deontological intuitions) “reflect the influence of morally irrelevant factors and are therefore unlikely to track the moral truth” (2008, 69–70). Haidt (2001) has reached a similar conclusion about the use of intuitions in moral reasoning. In reference to both Haidt and Greene's work,

Singer claims that “recent research in neuroscience gives us new and powerful reasons for taking a critical stance toward common intuitions” (2005, 332).

All of this evidence suggests that our intuitions are biased by forces that threaten to render them epistemically unreliable. In short, the way that the world seems to us may not be the way that the world actually is. This leads to a puzzle: if I am correct, then intuitions are indispensable for philosophical theorizing; however, we have reason to distrust our intuitions. Thus, we are stuck in a trap: we cannot trust the very source of our knowledge about philosophy.

3. A Partial Solution: Extra-Wide Reflective Equilibrium

In this section, I develop a methodology that I think goes *part of the way* toward solving the aforementioned puzzle. This method, which I call Extra-Wide Reflective Equilibrium, allows us to use our intuitions but also contains mechanisms for reducing (although not eliminating) the influence of bias in our intuitions. Let me explain this method, which I argue helps to resolve this puzzle.

3.1. Wide and Narrow Reflective Equilibrium

Reflective Equilibrium (RE) is a method in which one finds coherence between one’s intuitions about concrete cases and one’s general principles, weighing them against each other and revising them to be coherent with each other. Philosophers distinguish between Wide and Narrow RE. As Daniels says, “The method of wide reflective equilibrium is an attempt to produce coherence in an ordered triple of sets of beliefs held by a particular person, namely, (a) a set of considered moral judgments, (b) a set of moral principles, and (c) a set of relevant background theories” (1979, 258). Logistically, this involves an “agent working back and forth, making adjustments to his considered judgments, his moral principles, and his background theories,” ultimately finding coherence between all three (259). In contrast, Narrow RE involves only finding coherence between parts (a) and (b), without including (c). In that manner, Wide RE encompasses more than Narrow RE.

Both Wide RE and Narrow RE are methods that involve a single individual carrying out this process of finding coherence between her thoughts. My preferred version of RE takes the

method of Wide RE and applies it into a group setting. In the rest of this section, I develop and defend this method, which I call Extra-Wide RE.

3.2. *Extra-Wide RE*

Extra-Wide RE (EWRE) includes the same process of finding coherence between intuitions, principles, and background theories, but it adds into the mix that this coherence occurs in a group of individuals that satisfies certain conditions. Before I describe these conditions, it is important to clarify that these conditions may require actions that, given the current demography of the profession, are nearly impossible to achieve. And it is also important to note that these conditions are not meant to be institutional *requirements* but are rather meant as *recommendations* or best practices. Both of these caveats should be kept in mind as I describe the conditions of EWRE.

Let me discuss the conditions under which this group RE should be conducted. First, EWRE should include a representative number of philosophers from marginalized communities. This group should include individuals who have been historically marginalized in the profession as a whole and individuals who have been marginalized and excluded from the specific topic or question under discussion. Second, EWRE should include as much intellectual and viewpoint diversity as possible. Third, EWRE should include a diverse array of individual methods. For example, some of the philosophers should be classical intuition pumpers, some should be methodological naturalists, and so on. To the greatest extent possible, the most diverse range of individual methods should be included in EWRE. Fourth, EWRE should follow an egalitarian procedure. There should be a strictly enforced question, objection, rejoinder policy, in which individuals put their names in a queue and specify if they have a question, objection, or rejoinder. Each different kind of comment should have a specific time limit attached to it, so that no one individual can monopolize the conversation. In addition, nobody will be allowed to interrupt anyone; rather, everyone must obey the queue. Fifth, EWRE should include as many people as is feasibly possible to include in a group philosophy setting. Sixth, EWRE should be sensitive to the boundaries of specialization. For example, there should be some topic-specific groups that follow the first five conditions but within a group of philosophers of the same specialty, subtopic, and so forth. For example, there might be an EWRE group centered on a discipline as broad as feminist philosophy and a group as specific as one devoted entirely to

certain niche topics in bioethics. In addition, EWRE should also involve groups that intentionally bring philosophers together across the boundaries of specialization. For example, perhaps there could be an EWRE group that studies rationality and includes epistemologists, philosophers of social science, and so on.

With these conditions outlined, I will offer a brief sketch of what EWRE would actually look like. First, the group must be formed (according to the conditions described above). Second, the group must meet, ideally in person. In the meetings, the group must first determine the agenda of questions and issues under discussion. Second, the group must deliberate under the conditions described above.

It is important to note that EWRE is not meant to function as a democratic voting body. There need not be a group result that everyone converges on. If this were the case, then it would certainly be a good result. But, consensus is not the intention of this method. Rather, I intend EWRE to function as a means of bias reduction in each of the members of the group. This is compatible with each member of the group still retaining a different answer to the question at hand. In this manner, EWRE is different from discourse theories of ethics, because it does not require consensus. Rather, the point is to help the individual members of the group know which of their intuitions are biasing them.

3.3. How EWRE Reduces Bias

With EWRE described, let me now offer some reasons for thinking that this method helps to reduce the influence of biased intuitions.

First, EWRE helps to root out biased intuitions, because it includes the perspectives of underrepresented and marginalized groups in philosophy. These individuals, in virtue of their social position, have access to pools of evidence and experience different from those of philosophers who are not from marginalized groups. As a result, if one philosopher from a dominant group defends an intuitional claim that is being swayed by an unconscious bias or by a lack of experience, the philosophers from marginalized group can point this out. Assuming that the egalitarian conditions of EWRE obtain, this exchange would be amicable and well received. For example, in *A Theory of Justice* (1971) Rawls assumes that the individual in the Original Position should be a head of household, and he possibly envisioned this person as being a male

head of household. Okin (1989) rightly points out that this condition leaves issues related to the family outside the scope of justice. Presumably, Rawls did not make this claim out of any overt malice; it very likely just did not occur to him that this would be a problem. In other words, given his own biases, he likely did not see this implication of his view. Now, suppose that Okin and Rawls had both been engaged in the process of EWRE together; Okin would have been able to raise this concern, and Rawls would have had a chance to respond. Presumably, he would have revised his position, acknowledging that he had not thought of this issue, or he would have given Okin a response to defend his position. The basic idea working in this discussion is that individuals from marginalized groups have access to different evidence and can use this evidence to point out biased claims made by other philosophers.⁶ A similar point applies to viewpoint and intellectual diversity. If EWRE involves individuals of diverse worldviews and methodologies, there is a higher chance of avoiding groupthink and ideological piling on.

But, we need not take my a priori word for it. There is a wealth of empirical evidence to support the following two claims: (1) group deliberation helps to reduce biases and (2) group interaction helps to reduce biases. If both of these claims are true, then EWRE does indeed offer a mechanism for reducing bias among philosophers.

First, there is compelling evidence from psychology that the process of group interaction reduces biases. This claim, which is referred to as the Contact Hypothesis, holds that “intergroup contact under certain prerequisite conditions promotes the development of more harmonious intergroup relations” (Gaertner et al. 1994, 225). These conditions include “equal status between the groups . . . cooperative intergroup interaction, opportunities for personal acquaintance between outgroup members, and norms within and outside of the contact setting that support egalitarian intergroup interaction” (225). In short, the Contact Hypothesis claims that when groups interact under these conditions, their biases toward each other are lowered. As Gaertner argues, “Research within laboratory and field settings generally supports the efficacy of the list of prerequisite conditions for achieving improved intergroup relations” (225). Specifically, Gaertner argues that these improved relations include reduced biases; as he says, “we have direct evidence that interaction . . . can reduce bias” (1999b, 397). For example, in a study on more than a thousand students at a multicultural high school, Gaertner found that “the analyses involving students’ perceptions of the conditions of intergroup contact in this study’s multicultural setting offer continued support for the utility of the contact hypothesis as a strategic

framework for reducing intergroup bias” (Gaertner et al. 1994, 225). In total, there is psychological evidence that when groups interact under the conditions of the Contact Hypothesis, they experience less bias toward each other. The conditions under which EWRE is to be performed are nearly identical to the conditions of the Contact Hypothesis; thus, EWRE will potentially reduce the biases of its members. Furthermore, as Sunstein argues, “groups are especially likely to outperform the average individual when members are subject to ‘egocentric biases,’” which are biases that involve giving one’s own answer a higher probability than the answers of others (2005, 993). As Sunstein says, “in groups with diverse views, individual members learn that their own position is not universally held, and hence the bias is reduced” (Sunstein 2005, 993). In my experience, philosophers (myself included) are particularly susceptible to egocentric biases; thus, Sunstein’s argument suggests that diverse group interaction is particularly relevant for philosophers.

Second, there is evidence from psychology that the process of group deliberation can reduce the biases of the groups’ members. Gaertner claims that there is “substantial documentation that intergroup cooperative interaction reduces bias” (1999a, 692). Essentially, when put in deliberative interactions, humans have reduced biases toward each other. EWRE involves putting philosophers in these very kinds of interactions. Thus, EWRE has the potential to reduce (although not eliminate) biases.

In addition to psychological evidence, there is evidence from democratic theory that group deliberation is more likely to arrive at the truth than individual deliberation. The rationale for this position comes from the Condorcet Jury Theorem, which holds the following: “If each member of a jury is more likely to be right than wrong, then the majority of the jury, too, is more likely to be right than wrong; and the probability that the right outcome is supported by a majority of the jury is a (swiftly) increasing function of the size of the jury, converging to 1 as the size of the jury tends to infinity” (List and Goodin 2001, 283). In other words, for any group of people, if the probability of each individual member being correct in her judgment is more than 0.5, then adding more and more people increases the likelihood that the group will choose the correct answer. Conversely, if the probability that each person is correct is less than 0.5, then adding more people increases the likelihood that the group will choose the wrong answer. I take it to be plausible that individual philosophers, *when asked about a question that is within their research area*, have at least a 0.51 chance of giving the correct answer. Perhaps one might

disagree with this, but for the sake of argument, I will assume this.⁷ If this is true, then adding more philosophers to a group deliberation that is within their area will increase the chance of getting a true result. And since truth and bias are incompatible, it follows that adding more people will likely reduce bias. Because EWRE includes as large a group as is feasibly possible, it has a higher probability of arriving at true conclusions and of avoiding biases.

Conclusion

Intuitions are indispensable in philosophy, but they may be heavily influenced by biases. In order to partially remedy this problem, I have proposed and defended EWRE, which I believe has the resources to reduce (although not eliminate) the influence of biases on our intuitions. Finally, it is important to note that my account of EWRE is meant to outline a possible methodological ideal; this ideal may be, at present, not possible to realize without broader changes in the profession. With that in mind, I am not claiming that EWRE is a method that we could implement today; my claim is only that if we could use this method, it would help to reduce our biases.

References

- Brennan, Jason. 2010. "Scepticism About Philosophy." *Ratio* XXIII, no. 1:1–16.
- Brownstein, Michael. 2019. "Implicit Bias." *Stanford Encyclopedia of Philosophy*, Edition, edited by Edward Zalta. <https://plato.stanford.edu/entries/implicit-bias/>
- Buckwalter, Wesley, and Stephen Stich. 2014. "Gender and Philosophical Intuition." In *Experimental Philosophy*, edited by Joshua Knobe and Shaun Nichols, 307–46. Oxford: Oxford University Press.
- Climenhaga, Nevin. 2017. "Intuitions are Used as Evidence in Philosophy." *Mind* 127, no. 505:69–104.
- Daniels, Norman. 1979. "Wide Reflective Equilibrium and Theory Acceptance in Ethics." *Journal of Philosophy* 76, no. 5:256–82.
- Director, Samuel. 2018. "A Dilemma for Saulish Skepticism: Either Self-Defeating or Not Even Skepticism." *Disputatio* 10, no. 48:43–55.

- Gaertner, Samuel, Mary Rust, John Dovidio, Betty Bachman, and Phyllis Anastasio. 1994. "The Contact Hypothesis: The Role of a Common Ingroup Identity on Reducing Intergroup Bias." *Small Group Research* 25, no. 2:224–49.
- Gaertner, Samuel, Jeffrey Mann, John Dovidio, Audrey Murell, and Marina Pomare. 1999a. "How Does Cooperation Reduce Intergroup Bias?" *Journal of Personality and Social Psychology* 59, no. 4:692–704.
- Gaertner, Samuel, John Dovidio, Mary Russt, Jason Nier, Brenda Banker, Christine Ward, Gary Mottola, and Missy Houlette.. 1999b. "Reducing Intergroup Bias: Elements of Intergroup Cooperation." *Journal of Personality and Social Psychology* 76, no. 3:388–402.
- Greene, Joshua. 2008. "The Secret Joke of Kant's Soul." In *Moral Psychology*, vol. 3, edited by Walter Sinnott-Armstrong, 35–77. Cambridge, Mass.: MIT Press.
- Haidt, Jonathan. 2001. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108, no. 4:814–34.
- Huemer, Michael. 2005. *Ethical Intuitionism*. New York: Palgrave Macmillan.
- . 2007. "Compassionate Phenomenal Conservatism." *Philosophy and Phenomenological Research* 24, no. 1:30–55.
- Jaggar, Alison. 1995. "Toward a Feminist Conception of Moral Reasoning." In *Morality and Social Justice Point/Counterpoint*, edited by James P. Sterba, 115–46. Lanham, Md.: Rowman and Littlefield.
- List, Christian, and Robert Goodin. 2001. "Epistemic Democracy: Generalizing the Condorcet Jury Theorem." *Journal of Political Philosophy* 9, no. 3:277–306.
- Okin, Susan Moller. 1989. *Justice, Gender, and the Family*. New York: Basic Books.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- Saul, Jennifer. 2013. "Scepticism and Implicit Bias." *Disputatio* 37:243–63.
- Singer, Peter. 2005. "Ethics and Intuitions." *Journal of Ethics* 9, nos. 3–4:331–52.
- Sinnott-Armstrong, Walter. 2008. "Framing Moral Intuitions." In *Moral Psychology*, vol. 2: *The Cognitive Science of Morality: Intuition and Diversity*, edited by Walter Sinnott-Armstrong, 47–76. Cambridge, Mass.: MIT Press.
- Sunstein, Cass R. 2005. "Group Judgments: Statistical Means, Deliberation, and Information Markets." *NYU Law Review* 80:962–1049.

¹ Sidgwick and Ross are proponents of this view and had this understanding in mind when they viewed themselves as intuitionists.

² This argument draws heavily from a similar argument made in Huemer 2007.

³ Not only does this Huemer-inspired argument support the theoretical claim that philosophers should use intuitions, there is good reason to believe that philosophers do, in fact, use intuitions widely. For an argument to this effect, see Climenhaga 2017.

⁴ For a discussion of this point, see Sinnott-Armstrong 2008.

⁵ For a general discussion of implicit bias and its connection with epistemology, see Brownstein 2019. For an argument that implicit biases yield a form of skepticism see Saul 2013. For a response to Saul, see Director 2018.

⁶ Jaggar, in the process of developing a feminist method of moral reasoning, makes a similar point; she claims that “it is precisely those women [who have been excluded or marginalized] who may be best able to provide moral insight and even moral inspiration, especially in certain areas” (1995, 127).

⁷ There is perhaps strong reason to doubt that a philosopher taken at random is likely to be reliable in this fashion. See Brennan 2010 for an argument in defense of this view. It’s worth noting, however, that if philosophers are less than 0.5 likely to get the right answer, then their answers are no better than chance. This seems implausible and would suggest that we should not do philosophy at all.