The Unreliable Intuitions Objection Against Reflective Equilibrium

Norbert Paulo

Received: 7 March 2019 / Accepted: 7 January 2020 / Published online: 28 February 2020
© The Author(s) 2020

Abstract
Reflective equilibrium (RE) has been criticized for various reasons ever since the publication of Rawls’ A Theory of Justice. Recent empirical research into moral decision-making poses new challenges to RE because it questions the reliability of moral intuitions. This research might discredit moral intuitionism in general and RE in particular insofar as it ascribes epistemic value to moral intuitions. These findings suggest, for instance, that moral intuitions vary with cultural background, gender or framing. If it could be shown that all or certain kinds of moral intuitions are unreliable, this would diminish the plausibility of the claim that they provide good reasons to believe in the truth of their content, especially given the growing evidence that professional philosophers are just as biased and susceptible to distorting factors as lay people. In this paper I critically discuss three recent defenses of RE against this unreliable intuitions objection (UIO). Two of these defenses argue that the UIO misses its target, because RE does not depend on intuitions, or because the relevant considered judgments in RE are not intuitions. The third defense grants that RE depends on intuitions and that some Rawlsian considered judgments are intuitions. But it argues that the notion of considered judgments can be revised, and that wide reflective equilibrium (WRE) has the means to adopt the criticism because it can use information about unreliable intuitions and exclude those from further consideration. I argue that the first two defenses do not succeed. The third defense is the most promising one. However, I argue that this defense also fails. In this defense, RE is not a distinctive moral epistemic theory: different versions of RE would be conceivable, all of which are deeply problematic. I conclude with a discussion of the broader philosophical implications of the UIO, specifically whether or not accepting the UIO would warrant skepticism regarding moral knowledge.

Keywords Reflective equilibrium · Moral intuitions · Experimental philosophy · Skepticism
In contemporary philosophy, moral reasoning largely proceeds through the systematization of intuitions about moral problems or cases on all levels of generality. The idea is that moral problems or cases elicit moral intuitions, and that these intuitions provide defeasible reasons to believe in their content. Having the intuition that it is morally wrong to push the large man off the bridge in order to save the five workers on the tracks is thus a defeasible reason to believe in the truth of the judgment that it is morally wrong to push him. Ethicists often seek to justify certain moral principles by reference to such intuitions. They search for principles that can explain and ground the content of these intuitions (Kahane 2013; McMahan 2013). Probably the best known method that works in this way is reflective equilibrium, or RE [probably best explained in Rawls (1974: 8), see also his 1951 and his 2005; for recent reviews of the RE debate see Cath 2016 and Tersman 2018].

To engage in the method of RE is to start from intuitions about certain moral problems or cases and to systematize those with moral principles. In that process of systematization, intuitions and principles are reconsidered and revised until the set of intuitions and principles is coherent. If successful, this method leads to a state of RE in which the coherent set of beliefs is considered to be justified. But it might turn out that the state of RE cannot be reached. Then the set of beliefs is not considered to be justified. In contrast to this “narrow” RE, where only intuitions about particular moral problems and cases and more general principles have to be brought into coherence, “wide” RE also requires coherence with a range of relevant background theories in order for a set of moral beliefs to be justified. Acknowledging its widespread use, Arras titled his handbook chapter on RE: “The way we reason now” (2007). Scanlon even holds that RE is “the only defensible method: apparent alternatives to it are illusory” (2002: 149).

However, RE has also been criticized for various reasons ever since the publication of Rawls’ A Theory of Justice, mainly for its apparent emphasis on pre-theoretical intuitions. Brandt, for instance, cautioned that the method of RE “may be no more than a reshuffling of moral prejudices” (1979: 22). And Singer asked why we do “not rather make the opposite assumption, that all the particular moral judgments we intuitively make are likely to derive from discarded religious systems, from warped views of sex and bodily functions, or from customs necessary for the survival of the group in social and economic circumstances that now lie in the distant past?” (1974: 516).

Recent empirical research into moral decision-making questions the reliability of moral intuitions. It has been argued that this research discredits moral intuitionism in general (Sinnott-Armstrong 2006, 2008) and RE in particular (McPherson 2015) insofar as they ascribe epistemic value to moral intuitions. RE has been a target of the first paper in experimental philosophy, in which it is provocatively described as the paradigmatic instance of “intuition driven romanticism” in philosophy (Weinberg et al. 2001), and Singer has used findings from experimental philosophy and cognitive sciences to reinforce his earlier criticism of RE (2005). These findings suggest, for instance, that moral intuitions vary with cultural background (Ahlenius

---

1 I will discuss this claim later in the paper.
and Tännsjö 2012), gender (Buckwalter and Stich 2013) or framing (Nadelhofffer and Feltz 2008). This research thus reaffirms some of the earlier criticisms and poses new challenges to RE.

For the purposes of this paper I do not assume that the empirical evidence available today shows conclusively that moral intuitions are, in fact, unreliable. Rather, I want to investigate what it would mean for RE if the evidence for the unreliability of moral intuitions were conclusive. Now, if it could be shown that moral intuitions are unreliable, this would diminish the plausibility of the claim that they provide good reasons to believe in the truth of their content, especially given the growing evidence that professional philosophers are just as biased and susceptible to distorting factors as lay people (Tobia et al. 2013; Schwitzgebel and Cushman 2012; Horvath and Wiegmann 2016). Let us call this the *unreliable intuitions objection against RE (UIO)*:

\[ UIO: \text{If it could be shown that moral intuitions are unreliable, this would discredit RE.} \]

Let me stress right from the start that the UIO is a hypothetical objection. That is, as it stands, it does not discredit RE. In this sense it is different from the traditional and more recent criticism of RE mentioned before. Other than these critics, I do not assume that something actually discredits RE. Rather, my strategy is to investigate recent defenses of RE that claim, in one form or another, that no matter what the empirical evidence regarding the reliability of intuitions turns out to be, such evidence could never discredit RE—even if it were the worst possible evidence imaginable. So although the UIO is a hypothetical objection to RE, it targets the idea that RE is somehow immune to empirically-based criticism.

When moral intuitions are said to be “unreliable” this can be understood in at least two ways: first, such that the moral intuitions about test scenarios systematically vary with factors that the test participant herself would regard as morally irrelevant. That is, she actually judges the scenarios on the basis of factors she does not regard as those that should drive the judgment. Her intuitions are then unreliable by her own lights because she does not control the formation of her moral beliefs. Second, “unreliable” can be understood more objectively. Normally, beliefs are taken to be unreliable if they are likely to be in error (Goldman 1986). It is hard to pin down when exactly this is the case. In the moral realm it can mean that the moral intuitions about test scenarios systematically vary with factors that are, objectively, morally irrelevant, irrespective of whether or not the participant regards them as relevant. For instance, empirical findings seem to show that “personalness”—understood as having two prongs: the use of personal force and spatial distance—is a factor that

---

2 In fact, although the evidence suggests far-reaching and sometimes unexpected unreliability of intuitions (for an overview, see Machery (2017), chap. 2), some of the studies do have methodological shortcomings or can be interpreted in ways more favorable for intuitionists (Bengson 2013; Pölzler 2017).

3 I leave open the question which moral intuitions, or which proportion of them, is understood to be unreliable. I will discuss this problem later in the article. For now it suffices to make the quite realistic assumption that one just doesn’t know which intuitions are reliable and which are not.
actually drives responses to trolley cases (Greene 2014: 709 f.). Significantly more people find it morally acceptable to sacrifice the large man on the bridge if they do not have to push him, but can simply hit a switch next to the man that opens a trap door so he will fall onto the tracks. An even larger number of people find this acceptable if the switch is not next to the large man but further away. The general rule is this: the further away one is and the less one has to employ personal force, the more likely one is to find it permissible to give up on one person in order to save five. If this empirical claim is correct, and if one accepts the normative premise that the presence or absence of “personalness” is irrelevant to the moral acceptability of actions in such trolley scenarios, then one can—using the first understanding of “unreliable”—conclude that “judgments in response to trolley problems are strongly influenced by at least one morally irrelevant factor…, and are therefore at least somewhat unreliable.” (Greene 2010: 16; see also Paulo 2018).

The second understanding of “unreliable” is more demanding because it is harder to demonstrate that some factor is or is not—objectively—morally relevant. In this paper I rely solely on the less demanding first understanding of “unreliable.” Similarly to the “personalness” example, I assume that most philosophers would agree that feelings of disgust—completely incidental to the moral scenario and triggered by a filthy desk, bad odor (Schnall et al. 2008, 2017) or induced through hypnosis (Wheatley and Haidt 2005)—or the mere order of presentation of two or more test scenarios (Liao et al. 2012) should not be driving moral judgment.4

This being said, recall the hypothetical UIO: If it could be shown that moral intuitions are unreliable, this would discredit RE. Different attempts have been made to defend RE against the UIO. In what follows, I discuss three recent defenses in turn. All of them say that, no matter how damning the evidence against the reliability of moral intuition turns out to be, such evidence does not discredit RE as a moral epistemic method or standard. The critical discussions of the three defenses provide the opportunity to say more about what intuitions are and how RE can be understood. Two of the defenses argue that the UIO misses its target, because RE does not depend on intuitions, or because the relevant “considered judgments” in RE are not intuitions. I argue that these two defenses do not succeed and that the UIO holds. The third defense grants that RE depends on intuitions and that some Rawlsian considered judgments are intuitions. But it denies that this implies that RE is discredited (as the UIO has it): it argues that the notion of considered judgments can be revised, and that wide reflective equilibrium (WRE) has the means to adopt the criticism

4 Demaree-Cotton argues that the influence of such factors—and of framing effects in particular—can be seen as yielding objective unreliability because the judgments in relevantly similar cases would vary with the presence/absence of these factors and thus result in inconsistent beliefs (2016: 3–4). But note that moral consistency reasoning can also be understood as not being objective but as being judged by the lights of the individual agent (Campbell and Kumar 2012). Moreover, even though most of the examples of potentially disturbing factors might seem to be likely candidates for objective unreliability, this is not necessarily the case. As I show in another paper (Paulo 2020a), objective notions of unreliability are very hard to define. That most philosophers would agree on something certainly doesn’t imply that it is objectively the case. Also, the objective understanding invites criticism from moral relativists which I would like to avoid in this paper.
because it can use information about unreliable intuitions and exclude those from further consideration. In this way the third defense escapes the verdict of the UIO. However, I argue that this defense also fails. In this defense, RE is not a distinctive moral epistemic theory: different versions of RE would be conceivable, all of which are deeply problematic. I conclude with a discussion of the broader philosophical implications of the UIO, specifically whether or not accepting the UIO would warrant skepticism regarding moral knowledge.

1 Defense 1: RE Does Not Depend on the Use of Moral Intuitions

In a recent paper, Georg Brun offers a detailed argument for the claim that “even if one thinks that the recent attacks (from experimental philosophy) have discredited intuitions completely, this in itself is not a serious problem for the method of reflective equilibrium” (Brun 2014: 238). The argument is that the attacks miss their target because RE does not depend on the use of moral intuitions. To many, this will be surprising. After all, as pointed out above, RE is widely assumed to be an “intuitionistic” method or at least one systematizing initial moral intuitions into a coherent set of intuitions about particular cases on the one hand and more general moral principles on the other (and, in WRE, background theories) in order to reach a state of RE. Daniels, who first formulated an elaborate account of WRE, calls the method “modern” intuitionism (1979: 268). Against these common understandings Brun reconstructs the method of RE not as a relation between intuitions about particular cases and general principles but, instead, as one between “commitments” and “systematic elements.” Drawing on Goodman (1983), who was Rawls’ inspiration for important elements of the method of RE (Rawls 2005: 18; Føllesdal 2005), Brun takes “commitments” to “include a propositional attitude involving a certain degree of commitment, which need not be definite or unwavering,” whereas “systematic elements” are “those propositions which are part of some given moral theory” (Brun 2014: 239 f.; Scheffler 1954). Both can have the same content and appear at all levels of generality, but one need not actually be committed to systematic elements. So the idea is to start the process of searching for a state of RE with “antecedent commitments” and systematic elements and to refine and readjust both until one reaches a coherent set of “current commitments” and systematic elements, i.e. with a consistent set in which the commitments can be derived from the systematic elements. Moreover, the current commitments must have some epistemic standing independent of the coherence relations in RE; they must also “respect” the antecedent commitments in the sense that one has to have reasons to depart from an antecedent commitment (Brun 2014: 241).

So what is the role of moral intuitions, if any, in RE so construed? Brun argues that antecedent commitments might well be intuitions, but that they are not necessarily intuitions. He discusses four characteristics often ascribed to intuitions: first, they have propositional content; second, they have a certain epistemic credibility or reliability; third, they have a certain strength (i.e., they are not easily given up); and, fourth, that they are non-inferential. Focusing on non-inferentiality and strength Brun argues that understanding RE in terms of intuitions would be both too narrow
and too wide. It would be too narrow because not all commitments are intuitions, simply because not all commitments are non-inferential: they can also be based on prior theorizing, perception, memory or testimony. Also, once a state of RE is reached, all current commitments can be derived from the systematic elements, and it is possible that some will be believed because they are so derived. Understanding RE in terms of intuitions would also be too wide, Brun argues, because not all intuitions are commitments: Some intuitions are “recalcitrant” or “incorrigible,” i.e., they are hard or even impossible to abandon despite evidence that they are wrong. “If a subject believes that p is false but nonetheless does not manage to give up the intuition that p, he is not committed to p and consequently must ignore his intuition that p in striving for reflective equilibrium” (Brun 2014: 245).

To sum up, Brun argues that RE does not depend on the use of moral intuitions. Its notion of commitments is different from intuitions as understood in philosophy. RE thus escapes the verdict of the UIO. So, according to Brun, no matter how damning the empirical findings about the reliability of intuitions turns out to be, this will have no effect on the credibility of RE.

1.1 Response

Brun’s is an interesting contribution to the RE literature in that it combines it with the philosophical literature on intuitions. However, it is worth noting that his is a very particular understanding of RE, modeled after Goodman, not so much after Rawls [for a strictly Rawlsian interpretation of RE see Hübner (2017)]; and it is certainly not what most commentators take RE to be. As it stands, his argument is limited to this particular Goodman-version of RE. Those who stick to the common understanding of RE as a more or less “intuitionistic” method still have to answer the UIO.

Even if one accepts Brun’s Goodman-version of RE, his argument presupposes that the UIO is restricted to intuitions as understood in philosophy. In formulating the UIO I deliberately did not specify the term “moral intuitions.” What philosophers take intuitions to be might overlap with but need not exactly be what has been investigated in experiments [for an overview of different uses of “intuition” in philosophy and psychology see Kauppinen (2015); for more details see the chapters in Osbeck and Held (2014)]. It seems at least plausible to assume that the philosophical understanding of “intuition” is relevantly different (relevant, that is, for the UIO) from the cognitive scientific understanding. Take, for example, Greene’s argument regarding the role of the Doctrine of Double Effect (DDE). The DDE is often used to justify responses to trolley scenarios, and it is a paradigmatic example of the justification of moral principles by reference to case-specific intuitions. But Greene holds that the DDE is a result of mere intuition-chasing, i.e., a form of ex-post rationalization of unreliable intuitions:

5 This situation can at least arise when one understands intuitions as intellectual seemings or inclinations to believe that p. If they were understood as beliefs this would amount to inconsistent beliefs about p.
It’s been on the books for a long time, which gives it a dusty air of authority. But how did it get on the books? It seems that the DDE was codified because it was observed that certain intuitive patterns in moral judgment could be summarized by a set of principles now known as the DDE. We may infer this (inductively, not deductively) from the finding that lay moralists the world over make judgments consistent with the DDE while having no clue that they are doing so. (2014: 721)

Sure, to this one can respond that even if the DDE were merely an ex-post rationalization of intuitions, it can still be justified if we have an adequate normative justification for it (Lott 2016: 252). But when ethicists come up with further reasons supporting the DDE, these reasons can be subjected to empirical inquiry again (Paulo 2018). Greene’s point is that large parts of moral reasoning are mere ex-post rationalization of unreliable intuitions (see also Haidt 2001). This bold claim might or might not be true. What is important for my critique of Brun’s defense of RE against UIO is that, in Greene’s empirical research, the boundaries between pre-theoretical and theoretical, pre-reflective and reflective, inferential and non-inferential etc. are blurred—and the same holds, I would claim, for other experimental investigations of moral reasoning (e.g., Björklund et al. 2000; for critical discussion Hindriks 2015). In other words, what philosophers take intuitions to be does overlap with but is not exactly what has been investigated—and found to be unreliable—in empirical research.

So there seems to be some initial plausibility to the assumption that some “commitments” are just as unreliable as moral intuitions. What Brun would have to show is not that “commitments” in RE are relevantly different from “intuitions” as understood in philosophy. What would have to be shown is that “commitments” are relevantly different from “intuitions” as used and investigated in the cognitive sciences. Only a successful argument to that effect would immunize RE from empirical criticism regarding the unreliability of intuitions.

2 Defense 2: Considered Judgments are Not Intuitions

Others argue that RE escapes the verdict of the UIO because it draws on considered judgments, which are not intuitions in the sense of the cognitive sciences.

According to the common understanding—and in contrast to Brun’s understanding of the method—RE starts from intuitive judgments about particular cases. Greenspan has recently argued that the influential work of moral psychologist Haidt and others in moral decision making does not directly affect the method of RE because RE relies on considered judgments only (Greenspan 2015). For Rawls, Greenspan explains, considered judgments “are pre-theoretical judgments not

---

6 In the same paper Greenspan also argues that at least some instances of what Haidt and others criticize as mere ex post rationalization of moral intuitions can be reconstructed as credible inferences to the best explanation in a Peircean understanding. I will not discuss this argument here because it does not directly concern unreliable intuitions in RE.
derived from general principles—but not necessarily pre-reflective. His considered intuitive judgments may rest on factual inquiry, comparison of consequences, and mid-level rules... The only kind of reasoning he excludes is systematic derivation from general moral principles" (2015: 108). Already in his *Decision Procedure for Ethics*, Rawls demanded that the cases used to test moral principles should be reasonably simple and “likely to arise in ordinary life … in order that the judgment in question be made in the effort to settle problems with which men are familiar and whereupon they have had an opportunity to reflect” (1951: 182).

Greenspan then contrasts this notion of considered judgments with Haidt’s understanding of moral intuition as a “sudden appearance in consciousness of a moral judgment, including an affective valence (good-bad, like-dislike), without any awareness of having gone through steps of searching, weighing evidence, or inferring a moral conclusion” (2001: 818). This notion of pre-reflective intuitions is obviously different from Rawlsian considered judgments. If this is so then RE escapes the UIO because it draws on considered judgments and not on pre-reflective intuitions in Haidt’s sense. So, Greenspan’s argumentative strategy is similar to Brun’s. If successful, her argument would show that, no matter how damning the empirical findings about the reliability of intuitions turns out to be, this will have no effect on the credibility of RE.

### 2.1 Response

This certainly is a fair position, and one that deserves serious discussion. Starting with a concession, I take it to be clear that Haidt’s understanding of “intuition” is indeed significantly different from Rawlsian considered judgments. However, it is worth looking beyond those definitions and to ask what the UIO actually refers to. I have already pointed out above that in Greene’s empirical research, the boundaries between pre-theoretical and theoretical, pre-reflective and reflective, inferential and non-inferential etc. are blurred, which indicates that his research into factors that distort moral judgment, relevant for UIO, includes but goes well beyond what Haidt defines as “intuitions.” And the same is true of some of Haidt’s own research. Consider his dumbfounding experiments. In these, Haidt and his collaborators designed test scenarios in which participants had to evaluate offensive but harmless actions. The participants were later asked to justify or revise their evaluations. It is worth quoting Haidt’s description of the experiments in length. As he explains, the scenarios included:

- eating one’s dead pet dog, cleaning one’s toilet with the national flag, or eating a chicken carcass one has just used for masturbation. The stories were carefully constructed so that no plausible harm could be found, and most participants directly stated that nobody was hurt by the actions in question, yet participants still usually said the actions were wrong, and universally wrong. They frequently made statements such as, “It’s just wrong to have sex with a chicken.” Furthermore, their affective reactions to the stories (statements that it would bother them to witness the action) were better predictors of their moral judgments than were their claims about harmful consequences… [A]
ffective reactions were good predictors of judgment, whereas perceptions of harmfulness were not... [P]articipants were often “morally dumbfounded”...; that is, they would stutter, laugh, and express surprise at their inability to find supporting reasons, yet they would not change their initial judgments of condemnation. (2001: 817)

Far from directly investigating Rawlsian considered judgments, these studies do concern pre-reflective intuitions as defined by Haidt, but they go beyond the limitations of this definition. In asking the participants why they condemned these presumably harmless actions and if they would want to revise their verdict, the experimenters triggered “steps of searching, weighing evidence, or inferring a moral conclusion,” which are explicitly excluded by Haidt’s definition of “intuition.” And still, the findings are relevant for the UIO, not least because the initial intuitions seem to survive at least some forms of scrutiny, thus probably evolving from mere intuitions to something not too different from considered judgments. It will often be hard to tell whether something is a pure pre-reflective intuition falling under Haidt’s definition or whether the subject has several reasons to believe the relevant content—“overdetermination” (Audi 2008: 477; Brun 2014: 244)—some of which might not be captured by the definition of “intuition.”

Moreover, even if the relevant research would only address intuitions as defined by Haidt, it is not clear that considered judgments are not subject to the UIO. In a recent study Schwitzgebel and Cushman (2015) specifically tested for the difference familiarity, certainty and reflection make for the reliability of judgments. As pointed out above, proponents of RE take these conditions to make moral judgments more reliable and to be conducive to considered judgments in the Rawlsian sense. In the study, Schwitzgebel and Cushman specifically limited the test subjects to philosophers with relevant expertise and to philosophers who reported having stable opinions on the matter [i.e., certainty and familiarity with the specific types of moral dilemma in question (inter alia, trolley cases)]. Moreover, they chose an experimental setting that encouraged reflection on the moral judgment. To their own astonishment they found that neither expertise, nor familiarity with the moral problem used in the scenarios, nor certainty, nor reflection reduced the size of order or framing effects on judgments of specific cases: “We confess that we find our main result surprising: that is, our across-the-board failure to find evidence for philosophical expertise and reflection in moderating biased moral judgment... However, if there is a level of philosophical expertise that reduces the influence of factors such as order and frame upon one’s moral judgments, we have yet to find empirical evidence of it” (Schwitzgebel and Cushman 2015: 136).

So even if one concedes that Haidt’s understanding of intuitions as pre-reflective is different from Rawlsian considered judgments in RE (which I granted), and even if one holds that the relevant research only addresses intuitions as defined by Haidt (which I questioned), this in itself does not defend RE against the UIO: first, because Haidt’s research about unreliable intuitions goes beyond intuitions in the sense of pre-reflective snap judgments and arguably includes at least some instances of considered judgments. Second, because Schwitzgebel and Cushman’s study—together with a host of other studies suggesting similar conclusions—questions the
assumption that considered judgments are more reliable than pre-reflective intuitions. Philosophical expertise, familiarity with the problem and time to reflect on one’s response do not seem to make moral judgments more reliable.

These experimental findings are certainly not yet a sufficient basis for a conclusive answer to the question whether considered judgments are reliable, or at least more reliable than pre-reflective snap judgments. But they do provide initial evidence against the common philosophical assumption—from the armchair—that they are. A good defense strategy for proponents of RE would be to show—empirically—that considered judgments are, in fact (and in contradiction of recent findings), more reliable than intuitions in the sense of pre-reflective snap judgments. Moreover, as I have shown by two examples, Greenspan’s argument does not immunize RE from empirical criticism. Even if Haidt’s experiments would merely investigate intuitions as he defined them, this would not mean that experimental research that includes Rawlsian considered judgments could not be conducted in the future.

### 3 Defense 3: Information About Unreliable Intuitions is Valuable in WRE

Another strategy to defend RE against the UIO is to argue that RE can cope with the UIO even if some Rawlsian considered judgments are (unreliable) intuitions. This line of defense can be argued for in two different, though related, ways:

**Suggestion 1:** One might argue that the very notion of considered judgments should be revised. For instance, Tersman claims that although Rawls might not have thought about the kind of distorting factors Greene and others found, due to its “dynamic nature” the method of RE is always open for further development and revision in light of new evidence (Tersman 2008: 398). So in contrast to Greenspan’s defense of RE, Tersman does not stick to the canonical Rawlsian explication of considered judgments but suggests revising the notion in a Rawlsian spirit. One need not, Tersman holds, take Rawls’ statements about considered judgments literally when, in fact, Rawlsian spirit indicates that considered judgments are those that are most reliable given the state-of-the-art knowledge about moral decision-making. That is, proponents of RE can reformulate the understanding of considered judgments in light of new empirical findings. This first suggestion is a defense of RE against the UIO because unreliable intuitions would not figure in RE and thus not discredit it as a theory in moral epistemology.

**Suggestion 2:** In the same article, Tersman already hints at a second possible argument, namely to adopt WRE and let background theories do the work (2008). As mentioned earlier, in contrast to “narrow” RE, where only intuitions about particular moral problems and cases and more general principles have to be brought into coherence, WRE also requires coherence with a range of relevant background theories which might, with respect to the UIO, include theories about moral decision-making, distorting factors and biases. Such background theories, in turn, affect the credibility of considered judgments. So just as we have theoretical knowledge
about the reliability of our observations, in WRE we should strive for theoretical knowledge about the reliability of our moral intuitions. Once we have that knowledge we should accept the reliable intuitions as “evidence” and discard the others (Tersman 2018: 5). According to this defense, and in contrast to Tersman’s first suggestion, one should keep the notion of considered judgments as outlined by Rawls and thus let some moral intuitions enter the process of refinements and revisions in the method of RE. But this second suggestion demands that in that process, unreliable intuitions be abandoned or revised in light of relevant empirical findings that figure as background theories in WRE. So the difference between the two suggestions is that, in the first, unreliable intuitions do not even enter the method of RE, whereas, in the second, they might enter the method, but get filtered out in the search for a state of RE. Moreover, the first suggestion works in “narrow” RE as well as in WRE; the second only works in WRE.

The second suggestion is a defense of RE against the UIO because the pursuit of the method of RE (the back and forth between intuitions, principles and background theories, refining and potentially revising all elements) does not guarantee to reach a state of RE. It is possible that one follows the method perfectly well without ever arriving at a set of justified moral beliefs. So the idea of RE as a standard or ideal for justified moral beliefs is not affected by the UIO, because the UIO merely implies that we don’t have justified moral beliefs. This situation is perhaps more familiar in the sciences, where we often know that we would need better technical means—better microscopes or telescopes, say—in order to reliably form beliefs about very small or far-away objects. One might say that in these situations we have a proper epistemic theory but still lack the means to do what it demands in order to arrive at justified beliefs about the subject matter.

3.1 Response

Both suggestions in this line of defense of RE against the UIO aim at accommodating the criticism but reject the implication that this discredits RE. The first suggestion does so by revising the notion of considered judgments, the second by demanding coherence with the findings about unreliable intuitions in the use of RE. I respond to both suggestions in turn.

Ad Suggestion 1: What the first suggestion could amount to is nicely summarized by Huemer. Though moral theorizing must take intuitions into account, we must not uncritically accept the content of moral intuitions, even if we feel very confident in them. “Rather,” he argues,

we should screen intuitions according to the following criteria: (1) Seek a substantial body of ethical intuitions that fit together well, rather than placing great weight on any single intuition. (2) Eschew intuitions that are not widely

---

7 For instance, our observations are unreliable when made under the influence of certain substances, when one is distracted or focused on other tasks. But they can also be unreliable when made in ideal circumstances, just consider the well-known Müller-Lyer illusion.
shared, that are specific to one’s own culture, or that entail positive evaluations of the practices of one’s own society and negative evaluations of the practices of other societies. (3) Distrust intuitions that favor specific forms of behavior that would tend to promote reproductive fitness, particularly if these intuitions fail to cohere with other intuitions. (4) Distrust intuitions that differentially favor oneself, that is, that specially benefit or positively evaluate oneself, as opposed to others. (5) Distrust intuitions that line up with strong emotions one has about the things evaluated. (2008: 381 f.)

This list might not entail suggestions to counter all distorting factors uncovered so far. But since RE has a revisionary nature, it could potentially adopt all of these and many more, thus prolonging and refining the list. One might also design a list with open demands such as “distrust those intuitions which are discredited by decision making research.” At some point however, the resulting RE would hardly resemble the RE Rawls proposed. Consider, for example, the possibility that empirical research shows that 90%, 95% or 100% of relevant moral intuitions are unreliable. In such a situation, proponents of RE would have two options, both of which are problematic.

The first option is to acknowledge that the method of RE does not get off the ground, because the revised and refined list specifying which intuitions count as considered judgments would produce an empty set. That is, in the described situation, the method of RE could not even be used because empirical findings suggest that all potential considered judgments are unreliable. There would be no reliable “data” to work with. This implies that, in the context of the UIO (i.e., if it could be shown that moral intuitions are unreliable), RE is of no use. It simply would not be applicable.

The second option is to acknowledge that moral intuitions should not enter the method of RE whatsoever, and thus abandon reliance on considered judgments as a core element of RE. This option would turn RE into a significantly different moral epistemic theory, namely one that does not rely on moral intuitions as a kind of “evidence.” This reliance, however, was one of the key elements Rawls took from Goodman’s idea of how to understand the problem of induction, namely by searching for equilibrium between widely accepted practices of inductive reasoning (similar to considered moral judgments in RE) and rules of inductive reasoning (similar to moral principles in RE). This link between Goodman’s main idea and Rawls’ parallel understanding of RE would be cut. This option would pose significant theoretical costs: proponents of RE would have to explain the ground shifting deviance from RE as Rawls conceived of it. After all, even Singer, a fierce critic of RE, would accept RE thus conceived: “[i]f the interpretation is truly wide enough to countenance the rejection of all our ordinary moral beliefs, then I have no objection to it. The price for avoiding the inbuilt conservatism of the narrow interpretation, however, is that reflective equilibrium ceases to be a distinctive method of doing normative ethics” (2005: 347). And it is the very idea that intuitions as considered judgments are a central element of RE that must never be eliminated which Scanlon, one of the method’s main advocates, brings up to defend RE against the charge that it is empty (Scanlon 2002: 151).
Moreover, RE advocates would need to explain how RE would, despite such a fundamental change (abandoning reliance on considered judgments), still provide moral justification. For Rawls, and for many later advocates of RE, the justification of RE as a standard for moral justification largely rests on the epistemic value of considered judgments. The elimination of this element raises serious doubts about the justification of RE.

One can thus conclude that both options to defend RE against the UIO by revising the notion of considered judgments (suggestion 1) come with high theoretical costs.

Ad Suggestion 2: Recall Tersman’s second suggestion to leave the notion of considered judgments unaltered. This would let some unreliable moral intuitions enter the method of RE, where they would be abandoned or revised in light of relevant empirical findings that figure as background theories in WRE. In contrast to the first, this second suggestion thus keeps considered judgments as an important element of RE as a moral epistemic theory modeled after Goodman’s thoughts about the problem of induction. The second suggestion also has some use for RE in the context of the UIO (i.e., if it could be shown that moral intuitions are unreliable). In this context, the method of RE is applicable because there are some intuitions that count as considered judgments.

However, since the broader belief system does not contain resources for a defense of the reliability of most or all relevant moral intuitions, the second suggestion of RE implies that the moral beliefs under consideration are not justified. One might think that this is merely a problem for ethicists but none whatsoever for RE. After all, the method of RE does generate a result, namely that the beliefs are not justified. Tersman could thus argue that the idea of RE as a standard or ideal for justified moral beliefs is not affected by the UIO. That is, if it could be shown that moral intuitions are unreliable, RE implies that we don’t have justified moral beliefs, which is the kind of result one should expect from a theory in moral epistemology. In other words, Tersman could argue that the UIO does not hold because RE would not be discredited by findings about the reliability of moral intuitions.

At first glance, this sounds reasonable. However, in the context of the UIO this suggestion presupposes an implausible all-or-nothing view of justification that does not allow for degrees of justification. Consider this situation: out of five sets of moral beliefs currently under consideration, four are completely random and ill-formulated moral principles coupled with absurd background theories. One of the five sets, however, consists of well-founded background theories that cohere with a significant number of well-formulated moral principles. The problem is only that in this latter set of beliefs, too, there are no reliable moral intuitions (as the UIO has it). I take it that the one set of beliefs should be regarded not as fully but as better justified than the other four sets. But, given the insistence on the reliance on intuitions in RE, I do not see how proponents of RE could come to the same result. They would have to regard all five sets of beliefs as equally unjustified.

Tersman mentions the possibility of sets of beliefs being justified to some degree when a subject “has not managed to bring all her considered judgments and moral beliefs into equilibrium” (2018: 2). This I find very plausible. But note that this does not cover the situation I described for it assumes that some (though maybe not all)
considered judgments are reliable. In the context of the UIO, all intuitions (and considered judgments) are unreliable. The subject knows from her (cognitive scientific, say) background theories that she has no reliable considered judgments whatsoever to draw on. Given RE’s reliance on intuitions as some kind of evidence (which is kept in this second suggestion in contrast to the first), I do not see how advocates of RE could explain why the one set of beliefs is more justified than the other four without facing the same problems discussed for the first suggestion’s second option: They would have to explain how RE, despite having no considered judgments to rely on, would still provide moral justification. The inability to assign degrees of justification in the context of the UIO is thus a problem for the method of RE rather than an instance of its well-functioning.

Another more general objection speaks against the understanding of RE as the search for coherence between moral principles and all relevant state-of-the-art background theories as this second suggestion has it. This understanding comes dangerously close to an immunization of RE, rather than a defense of it. For if coherence and revisability in the light of all relevant background theories were the core of RE, how could it ever be shown to be defective? It seems that RE would—almost by definition—always be the right method to use, exactly because it asks for coherence with all relevant state-of-the-art background theories [for a similar objection McPherson (2015: 661)]. Immunization is not only a theoretical vice, it also comes with the practical problem of indeterminacy: the risk of immunization occurs because the characteristics of coherence and revisability are so indeterminate. It is hard to tell what they demand and what they exclude so that RE thus understood risks being of little help in figuring out what to believe, which is what a theory in moral epistemology should do.

4 Implications for the Use of RE

So far, I have critically discussed three strategies to defend RE against the hypothetical UIO. All of them tried to defend the view that, no matter how damning the evidence against the reliability of moral intuition turns out to be, such evidence does not discredit RE as a moral epistemic method or standard. I have argued that none of these three strategies is successful.

One might take the thus established UIO to support a skeptical position concerning moral knowledge. Since the UIO only mentions RE, the skeptical conclusion would only be warranted if objections similar to the UIO would also affect alternative moral epistemic theories. This obviously presupposes two things. First, that there actually is damning evidence regarding the unreliability of intuitions, and, second, that there are such alternative theories. For the purposes of this paper, I want to bracket the first question about the actual evidence, which I discuss in other papers (Paulo 2020a, b). Here I want to focus on the hypothetical question what condemning evidence would mean for RE. So let’s focus on the second point. Whether or not there really are alternatives to RE is not a trivial question. As mentioned earlier, some proponents of RE regard it as “the only defensible method: apparent alternatives to it are illusory” (Scanlon 2002:}
The Unreliable Intuitions Objection Against Reflective... 347

149; see also DePaul 2006: 618). The idea behind that claim is that WRE already includes all other methods sometimes regarded as alternatives to RE because, as Scanlon argues, the search for WRE

allows for what might have been seen as alternative methods of justification to be incorporated within it. … [RE] thus allows all possible sources of justificatory force to be considered. But the method is not vacuous because it is incompatible with some views about these sources. It is incompatible, first, with the idea that any particular class of judgments or principles can be singled out in advance of this process as justified on some other basis and, second, with the idea that any class of considered judgments should be left out of this process (for example that “intuitions” about what is just or unjust in particular cases should not be given any weight in justifying general principles but must be derived from them.) (2002: 151)

Note that even if all of this was true, it would not warrant the rather strong conclusion that RE does not have any alternatives. First, alternative methods which are excluded by these two conditions could well be correct. Scanlon offers no reason why it would be “illusory” to conceive of the possibility that they are correct. Second, if WRE really included all the alternative methods, it would very likely also include some alternative methods that are wrong. In other words, maybe RE is too inclusive to be true. So it is at least possible that a moral epistemic theory completely different from, or more restrictive than, RE is better suited to yield justified moral beliefs. It is thus not “illusory” to consider alternatives to RE.

Maybe what Scanlon has in mind is that philosophers should continue using RE—no matter how much evidence against the reliability of moral intuitions is found—as long as there is no better alternative. Consider Jonathan Floyd’s recent argument for such a position: “The ultimate defence of Rawls’ method,” Floyd argues, “is that unless we can construct an alternative, together with a convincing argument regarding its superiority, we should just ‘keep calm and carry on’” (Floyd 2017: 378). To make his point Floyd draws on Freeman and Pogge. However, Freeman argues for a slightly different position, namely that

[w]e have to begin somewhere in moral thinking, and reliance upon considered moral convictions is unavoidable in so far as we do. Unless the critic is prepared to reject moral thinking as illusory, at least under current conditions, this criticism carries no force, at least not until we are provided with a convincing explanation why our moral convictions are, en masse, unreliable (2006: 35).

To Freeman’s point one might object that there are alternatives to RE that do not rely on or start from moral intuitions or “considered moral convictions.” Habermas’s idea of an ideal discourse (1983), Dworkin’s Herculean task to find the one right answer (1986, 2011: pt. 1 and 2) and reliance on transcendental arguments (de Maagt 2017) come to mind as examples for such alternatives. What is more, Freeman seems to be concerned with how to get moral discourse started, which is quite different from assigning epistemic value to intuitions, as RE does.
Moreover, Freeman’s position is different from Floyd’s in that Freeman’s would allow for a situation where we do not have a theory of justification, namely if RE were discredited and if there were no alternative theories. Floyd’s position would not allow for such a situation because if the two conditions would obtain, one were to rely on RE.

When Floyd refers to Pogge he takes him to hold “that there is simply no alternative way of doing political philosophy” (Floyd 2017: 378). However, in the passage Floyd refers to, Pogge makes a different point: He argues that there are such different ways to do political philosophy, but that their forms of reasoning can be reformulated in terms of the method of RE (Pogge 2007: 176). So, provided this observation is correct, RE might be considered a good theory to describe the functioning of various theories in moral epistemology. It might also be considered a good theory alternative theories can be reduced to. But even if that were true, it remains open why this has any bearing on the normative question which moral epistemic theory to rely on. So Floyd’s authority argument—relying on Freeman and Pogge—does not succeed.

But is there independent reason for Floyd’s position? To be sure, as Floyd of course knows, there are alternatives to RE. Every textbook and encyclopedia entry on moral epistemology lists the standard options. To name just some alternatives that have been defended in recent years: Audi (2004, 2013), Huemer (2005) and Katarzyna de Lazari-Radek and Peter Singer (2016) have forcefully defended versions of (rational) intuitionism; and Campbell and Kumar have developed a model of moral reasoning with cases that draws on empirical research into moral decision-making and argued that it is superior to RE (Campbell and Kumar 2012; Campbell 2014; Paulo 2019).

So there is no shortage in alternatives. Floyd’s idea that we should continue using RE as long as no alternative has been shown to be superior seems to be too demanding, for it is not to be expected anytime soon that moral epistemologists miraculously agree that an argument regarding the superiority of RE or of one of its competitors is convincing. Even if there is progress, convergence and agreement in some respects, the question of a moral epistemic theory’s across-the-board superiority is likely to remain contested for the foreseeable future. But this problem to demonstrate a theory’s superiority does not in itself speak for RE. There is no apparent reason to regard RE as the default option in moral epistemology in situations of peer disagreement.

It is on the advocates of every particular moral epistemic theory to demonstrate why philosophers should use this theory as a standard of justification instead of its alternatives. Moreover, it remains possible that in moral matters no satisfactory epistemic theory yet exists. This being said, the burden of proof does not only shift to the advocates of RE when empirical evidence conclusively shows that moral intuitions are unreliable (as the UIO has it). Instead, they have to motivate the use of RE (maybe in comparison to its alternatives). So RE advocates should not keep too calm. They should carry on searching for reasons supporting RE. In doing so they should bear in mind that in a discussion of Brandt’s, Singer’s and others’ early criticism of RE mentioned above, Daniels, one of the method’s main proponents, argued
that the burden of proof had already shifted to the advocates of RE. His argument is worth being quoted at length:

[In WRE] we may have a way of distinguishing initially credible from merely initially believed types of moral judgments. The “no credibility” criticism gains initial plausibility because we are able to assign initial credibility to nonmoral observation reports, but not to moral judgments. The credibility assignment, however, draws implicitly on a broadly accepted body of theory which explains why those judgments are credible. Properly understood, the credibility story about nonmoral observation reports is itself only the product of a nonmoral wide reflective equilibrium of relatively recent vintage. In contrast, we lack that level of theory development in the moral case. What follows from this difference is that the “no credibility” argument succeeds in assigning a burden of proof. Some answer to the question about the reliability of moral judgments must be forthcoming. But the argument is hardly a demonstration that no plausible story is possible. (Daniels 1979: 271 f.)

Assuming that Daniels rightfully conceded that the early “no credibility” criticism had already succeeded in assigning the burden of proof to the proponents of RE, this seems all the more true today, 40 years later, with the early criticism still largely unanswered and ever more empirical insights into the unreliability of moral intuitions. It is not too late for proponents of RE to develop a “plausible story,” as Daniels put it, for RE’s epistemic reliability. But so far, they have not told this story.

5 Concluding Remarks

Recent empirical research into moral decision-making poses new challenges to RE because it questions the reliability of moral intuitions. These findings suggest, for instance, that moral intuitions vary with cultural background, gender or framing. If it could be shown that all or certain kinds of moral intuitions are unreliable, this would diminish the plausibility of the claim that they provide good reasons to believe in the truth of their content. In this paper I critically discussed three recent strategies to defend RE against this hypothetical unreliable intuitions objection (UIO). Two of these defenses argue that the UIO misses its target, because RE does not depend on intuitions (but on “commitments”), or because the relevant considered judgments in RE are not intuitions. The third defense grants that RE depends on intuitions and that some Rawlsian considered judgments are intuitions. But it argues that the notion of considered judgments can be revised, and that WRE has the means to adopt the criticism because it can use information about unreliable intuitions and exclude those from further consideration.

I argued that none of the three defense strategies against the UIO is successful. Therefore, the UIO holds: If it could be shown that moral intuitions are unreliable, this would discredit RE. In other words, RE is not generally immune against empirical criticism regarding the unreliability of intuitions.

Because one might take the thus established UIO to support a skeptical position concerning moral knowledge, I also discussed whether or not accepting the UIO...
would warrant skepticism. If the UIO holds and RE would be discredited through empirical findings about the unreliability of moral intuitions, this would support skepticism, or so the argument goes. Of course, this argument only holds if RE were the only available moral epistemic theory. I critically discussed, and ultimately rejected, claims to the effect that there are, in fact, no alternatives to RE. Since the UIO merely targets RE and not its alternatives, it does not warrant skepticism (even if damning evidence about moral intuitions were found).

Finally, I engaged with the claim that, as long as there is no better alternative, we should just “keep calm and carry on” with RE. I followed Daniels’ argument that even the early criticism of RE was so forceful that the burden of proof to show which moral epistemic theory is preferable is on the proponents of RE. This seems all the more true today, four decades later, with new empirical findings casting even more doubt on the reliability of moral intuitions.

Acknowledgements Open access funding provided by University of Graz. For helpful comments and discussion, many thanks to my colleagues at the University of Graz, to audiences at Central European University, Budapest, at the University of Hamburg and at the University of Regensburg, as well as to Christoph Bublitz, Julia Hermann, Dietmar Hübner, Shelly Kagan, Peter Königs, Leonhard Menges and, especially, to Folke Tersman.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Ahlenius, Henrik, and Torbjörn Tännsjö. 2012. Chinese and Westerners Respond Differently to the Trolley Dilemmas. Journal of Cognition and Culture 12 (3–4): 195–201.

Arras, John D. 2007. The Way We Reason Now: Reflective Equilibrium in Bioethics. In The Oxford Handbook of Bioethics, ed. Bonnie Steinbock, 46–71. Oxford: Oxford University Press.

Audi, Robert. 2004. The Good in the Right: A Theory of Intuition and Intrinsic Value. Princeton: Princeton University Press.

Audi, Robert. 2008. Intuition, Inference, and Rational Disagreement in Ethics. Ethical Theory and Moral Practice 11 (5): 475–492.

Audi, Robert. 2013. Moral Perception. Princeton: Princeton University Press.

Bengston, John. 2013. Experimental Attacks on Intuitions and Answers. Philosophy and Phenomenological Research 86 (3): 495–532.

Björklund, Fredrik, Jonathan Haidt, and Scott Murphy. 2000. Moral dumbfounding: when intuition finds no reason. Lund Psychological Reports 1 (2). http://lup.lub.lu.se/record/1024827. Accessed 21 Jan 2020.

Brandt, Richard B. 1979. A Theory of the Good and Right. Oxford: Oxford University Press.

Brun, Georg. 2014. Reflective Equilibrium Without Intuitions? Ethical Theory and Moral Practice 17 (2): 237–252. https://doi.org/10.1007/s10677-013-9432-5.

Buckwalter, Wesley, and Stephen Stich. 2013. Gender and Philosophical Intuition. In Experimental Philosophy, vol. 2, ed. Joshua Knobe and Shaun Nichols, 307–347. Oxford: Oxford University Press.
The Unreliable Intuitions Objection Against Reflective…

Campbell, Richmond. 2014. Reflective Equilibrium and Moral Consistency Reasoning. Australasian Journal of Philosophy 92 (3): 433–451.

Campbell, Richmond, and Victor Kumar. 2012. Moral Reasoning on the Ground. Ethics 122 (2): 273–312.

Cath, Yuri. 2016. Reflective Equilibrium. In The Oxford Handbook of Philosophical Methodology, ed. Herman Cappelen, Tamar Szabo Gendler, and Hawthorne John, 213–230. Oxford: Oxford University Press.

Daniels, Norman. 1979. Wide Reflective Equilibrium and Theory Acceptance in Ethics. Journal of Philosophy 76 (5): 256–282.

de Lazari-Radek, Katarzyna, and Peter Singer. 2016. The Point of View of the Universe: Sidgwick and Contemporary Ethics. New York: Oxford University Press.

de Maagt, Sem. 2017. Reflective Equilibrium and Moral Objectivity. Inquiry 60 (5): 443–465. https://doi.org/10.1080/0020174X.2016.1175377.

Demaree-Cotton, Joanna. 2016. Do Framing Effects Make Moral Intuitions Unreliable? Philosophical Psychology 29 (1): 1–22. https://doi.org/10.1080/09515089.2014.989967.

DePaul, Michael R. 2006. Intuitions in Moral Inquiry. In The Oxford Handbook of Ethical Theory, ed. David Copp, 595–623. New York: Oxford University Press.

Dworkin, Ronald. 1986. Law’s Empire. Cambridge: Harvard University Press.

Dworkin, Ronald. 2011. Justice for Hedgehogs. Cambridge: Harvard University Press.

Floyd, Jonathan. 2017. Rawls’ Methodological Blueprint. European Journal of Political Theory 16 (3): 367–381. https://doi.org/10.1177/1474885115605260.

Føllesdal, Dagfinn. 2005. The Emergence of Justification in Ethics. European Review 13 (2): 169–182. https://doi.org/10.1071/S096279870500030X.

Freeman, Samuel. 2006. Rawls. London: Routledge.

Goldman, Alvin I. 1986. Epistemology and Cognition. Cambridge: Harvard University Press.

Greene, Joshua D. 2010. ‘Notes on “The Normative Insignificance of Neuroscience” by Selim Berker’. https://static1.squarespace.com/static/54763f79e4b0c4e55fb000c/t/54cb945ae4b001aedee69e81/1422627930781/notes-on-berker.pdf. Accessed 21 Jan 2020.

Greene, Joshua D. 2014. Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics. Ethics 124 (4): 695–726.

Greenspan, Patricia. 2015. Confabulating the Truth. In Defense of “Defensive” Moral Reasoning. The Journal of Ethics 19 (2): 105–123.

Habermas, Jürgen. 1983. Diskursethik—Notizen zu einem Begründerungsprogramm. Moralbewußtsein und kommunikatives Handeln, 53–125. Frankfurt am Main: Suhrkamp.

Haidt, Jonathan. 2001. The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. Psychological Review 108 (4): 814–834.

Hindriks, Frank. 2015. How Does Reasoning (Fail to) Contribute to Moral Judgment? Dumbfounding and Disengagement. Ethical Theory and Moral Practice 18 (2): 237–250.

Hübner, Michael. 2005. Ethical Intuitionism. Basingstoke: Palgrave Macmillan.

Huemer, Michael. 2008. Revisitory Intuitionism. Social Philosophy and Policy 25 (1): 368–392. https://doi.org/10.1017/S026502550808014X.

Kahane, Guy. 2013. The Armchair and the Trolley: An Argument for Experimental Ethics. Philosophical Studies 162 (2): 421–445.

Kauppinen, Antti. 2015. Moral Intuition in Philosophy and Psychology. In Handbook of Neuroethics, ed. Jens Clausen and Neil Levy, 169–183. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-4707-4_163.

Liao, S.Matthew, Alex Wiegmann, Joshua Alexander, and Gerard Vong. 2012. Putting the Trolley in Order: Experimental Philosophy and the Loop Case. Philosophical Psychology 25 (5): 661–671. https://doi.org/10.1080/09515089.2011.627536.

Lott, Micah. 2016. Moral Implications from Cognitive (Neuro)Science? No Clear Route. Ethics 127 (1): 241–256. https://doi.org/10.1086/687337.

Machery, Edouard. 2017. Philosophy within Its Proper Bounds. Oxford: Oxford University Press.
Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.