



The playing field of empirical facts: on the interrelations between moral and empirical beliefs in reflective equilibrium

Manuel Cordes¹

Received: 1 March 2023 / Accepted: 5 March 2024
© The Author(s) 2024

Abstract

What exactly is the role of empirical beliefs in moral reflective equilibrium (RE)? And if they have a part to play, can changes in our empirical beliefs effectuate changes in the moral principles we adopt? Conversely, can empirical beliefs be adjusted in light of certain moral convictions? While it is generally accepted that empirical background theory is of importance to the method of wide reflective equilibrium (WRE), this article focuses on a different aspect, namely the role of empirical beliefs that is intrinsic to the coherence relation of moral beliefs in any (narrow or wide) conception of RE. First, it is shown that in the application of RE, empirical beliefs are crucial to the procedure of matching principles to considered judgments. Changes in our empirical beliefs may therefore upset the relevant coherence relation and motivate changes in the moral principles we adopt. This more detailed account of the interplay between empirical and moral beliefs can help evaluate the overall philosophical appeal of RE. Finally, it is argued that, at least in WRE, the relation between empirical and moral beliefs need not be a one-way street, that is, moral convictions can provide us with some epistemic means of adjudicating between competing empirical descriptions of the world.

Keywords Reflective equilibrium · Coherence · Coherentism · John Rawls · Peter Van Inwagen · Folke Tersman

✉ Manuel Cordes
m.cordes@cs.l.mpg.de

¹ Max Planck Institute for the Study of Crime, Security and Law, Günterstalstraße 73, Freiburg i. Br. 79100, Germany

1 Introduction

What exactly is the role of empirical beliefs in moral reflective equilibrium (RE)? And if they have a part to play, can changes in our empirical beliefs effectuate changes in the moral principles we adopt? Conversely, can empirical beliefs be adjusted in light of certain moral convictions? It is generally accepted that a wide reflective equilibrium (WRE) includes certain empirical background theories (e.g., social theory) in the sets of beliefs among which an ethical inquiry seeks coherence.¹ However, this article focuses on a different aspect, namely the role of empirical beliefs that is intrinsic to the coherence relation of moral beliefs in any (narrow or wide) conception of RE.²

Here is a straightforward way in which empirical beliefs are relevant to RE: if one core aspect of the method consists in testing abstract principles against considered judgments on particular cases, then empirical beliefs inform us about *what is the case*. Starting from this observation, I will analyze in more detail the role of empirical beliefs in the process of matching moral principles to considered judgments (Sect. 2). As will become apparent, empirical beliefs are crucial not only for establishing the match between principles and considered judgments *about particular cases*, but also for establishing the match between principles and judgments at higher levels of generality. Any (narrow or wide) conception of RE must therefore rely on a set of empirical beliefs. Additionally, this account can be linked to a more general coherentist³ framework, such as the one proposed by Tersman (1993). Drawing on the upshot of Sect. 2, I will then shed light on how empirical changes and discoveries may upset RE, thereby motivating changes in the moral principles we adopt (Sect. 3). While this more detailed understanding of the relation between moral and empirical beliefs within RE might be interesting in its own right from a purely descriptive standpoint, it can also, as I will argue in Sect. 4, contribute to evaluating the overall philosophical appeal of RE: first, by demonstrating the empirical adaptability of RE, which in turn

¹ On Daniels's (1979) influential account, WRE aims at producing coherence between an ordered triple of sets of beliefs, namely (a) a set of considered moral judgments, (b) a set of moral principles, and (c) a set of relevant background theories. The latter encompasses both moral and empirical bodies of theory (Arras, 2009, p. 52; Daniels, 1979, p. 260).

² Most philosophers agree with Rawls that for the justification of moral claims, “[w]ide and not narrow reflective equilibrium is plainly the important concept” (Rawls, 2001, p. 31). I do not wish to contest this view by explicitly mentioning narrow reflective equilibrium (NRE), but rather to describe more precisely the relation between moral and empirical beliefs within RE. To this end, it is worth noting that—and how exactly—empirical beliefs *already* play a role in NRE, which is solely characterized by matching proposed moral principles to considered judgments. In other words, I do not regard NRE as a serious theoretical contender to WRE, but primarily as a certain *level of description*. As I will show in this article, empirical beliefs are crucial even to the narrow description of RE.

³ Following Olsson (2023), I understand coherentism as the thesis that a belief or set of beliefs is justified if and only if “the belief coheres with a set of beliefs, the set forms a coherent system or some variation on these themes.” For the purposes of this article, I think I can remain neutral on the question of what exactly the epistemic goal of the justificatory process is. While the classical view holds that justification is a necessary condition for propositional knowledge, more recent proposals argue that the justificatory process, and RE in particular, aims at objectual understanding rather than knowledge (Baumberger & Brun, 2021; Elgin, 2017).

may count as a theoretical advantage of the method, and second, by helping assess (and partly answer) a criticism of the method raised by Kelly and McGrath (2010).

In the final section, I will explore the question of whether and how empirical beliefs may be adjusted in light of certain moral convictions (assuming that coherence considerations can in principle affect the justification of empirical beliefs). Like the preceding sections, then, Sect. 5 is concerned with the interrelations between moral and empirical beliefs within RE, but addresses the opposite direction of impact, so to speak. While Sects. 2–4 state an account (and some further implications) of how the justification of moral beliefs is informed by empirical ones, Sect. 5 focuses on how the justification of empirical beliefs may be informed by moral ones. I think that van Inwagen's "An Essay on Free Will" provides a good example for discussing this latter question. After assessing and ultimately rejecting van Inwagen's deductive inference from a belief in the existence of moral responsibility to a belief in the existence of free will, I will be in a position to make a more general point about the possible adjustment of empirical beliefs in light of moral ones. My central claim in this context is that in applying WRE, when in empirical doubt, we may have a reason to prefer an empirical belief that ensures a higher degree of *ethical* consistency (all other things being equal). Therefore, even if deductive arguments from moral premises to empirical conclusions fail, moral beliefs can still provide us with some epistemic means of adjudicating between competing empirical descriptions of the world. In other words, at least within WRE, the relation between empirical and moral beliefs need not be a one-way street.

2 The matching relation

On an initial sketch, the RE procedure consists in seeking the greatest possible coherence among a set of moral principles and a set of considered judgments, working back and forth between both elements and revising any of them if deemed necessary to produce the required coherence. Let us leave aside for the moment the distinction between narrow and wide RE and focus instead on the coherence relation between moral principles and considered judgments required in any (narrow or wide) conception of RE. How are we to understand this coherence relation?

2.1 Matching principles and considered judgments on particular cases

In his early "Outline of a Decision Procedure for Ethics", Rawls (1951) demands an "explication" of considered judgments on particular cases by moral principles. In his later writings, he speaks simply of a "match" between considered judgments and principles (Rawls, 1974, p. 7; 1999, pp. 17–18, pp. 41–43). The criteria that are needed for this relation to hold, however, are the same on all his accounts: a principle explicates, or matches, a non-principled considered judgment about a particular case if and only if that same judgment could be inferred from the principle applied to that same case.

Now it seems clear that to render such moral judgments—principled or not—about particular cases, we need to know the facts. In defining the relevant class of considered moral judgments, Rawls (1951, p. 182) sets forth the following condition:

It is required that the judgment be one which has been preceded by a careful inquiry into the facts of the question at issue [...]. This requirement is justified on the ground that it is only by chance that a just decision can be made without a knowledge of the relevant facts.

The application of a principle to a particular case also requires an assessment of the facts. Otherwise, one could not decide whether or not the abstract criteria of the principle are met. Rawls (1999, p. 41) hints at this when he writes that principles yield concrete judgments “when conjoined to our beliefs and knowledge of the circumstances”.

In light of the above considerations, the relevant notion of a match can be defined as follows:

(Matching Relation 1) A principle p matches a considered judgment c about a particular case if and only if p accounts for c . A principle p accounts for a particular moral judgment c if and only if c can be inferred from p in conjunction with a set of empirical facts about the particular case in question.

Given this definition, how exactly does one proceed to achieve a match between a moral principle and a considered judgment on a particular case? I think the relevant procedure is characterized by two main features:

1. The inquirer compares a principled and a considered non-principled judgment about the same case (call it the test case), seeking equivalence between them by adjusting both the moral principle and the considered judgment, if necessary. When equivalence is achieved between the judgment inferred from the principle and the considered non-principled one, the principle accounts for and thus matches the latter.
2. Both principled and non-principled judgments are crucially informed by an inquiry into the empirical facts of the test case.

The matching relation between moral principles and considered judgments on particular cases can therefore be understood as a three-place relation: Principles and considered judgments are related to each other, while both are related to a common test case. One could also say that the relevant moral beliefs only ever match *in view* of a test case.

This means that if a method of RE—even NRE—makes use of Matching Relation 1, it relies on some set of empirical beliefs. One might say that the empirical makeup of a particular case constitutes the playing field on which moral principles and considered judgments are tested against each other. The outcome of this procedure—as I will argue shortly—may well change if the playing field changes, or if we change our mind on what exactly the playing field looks like.

But first, let us turn to Rawls's later modification, according to which the considered judgments relevant to RE do not concern only particular cases:

“People have considered judgments at all levels of generality, from those about particular situations and institutions up through broad standards and first principles to formal and abstract conditions on moral conceptions.” (1974, p. 8).

In the next two subsections, I will seek to make sense of the matching relation between moral principles and considered judgments at higher levels of generality. Here, too, empirical beliefs prove to be highly relevant.

2.2 Matching principles and summary judgments

One form of generalized moral statement that people sometimes make refers not to one particular case but to a range of cases. Consider the following assertion:

- a) *In many cases*, when an intellectually mature and mentally healthy person commits a wrong, they are responsible for their wrongdoing.

Such judgments (call them “summary judgments”) summarize a set of concrete moral judgments on particular cases. Given that any one of these particular judgments in the summarized range is informed by empirical beliefs regarding the facts of the case (as seen above), it plainly follows that summary judgments are informed by the sum of these empirical beliefs. Now perhaps some of the more general considered judgments relevant to RE are of such summary form. How, then, can we check whether they match a set of moral principles? Given that a summary judgment refers to a set of judgments on particular cases, we can apply a proposed principle to these very cases and see whether the resulting set of principled judgments licenses the same summary. Extending the example in a), one could apply some principle of moral responsibility to all cases of wrongdoing by an intellectually mature and mentally healthy person and see whether the resulting set of principled judgments can be summarized in the same way, namely by saying that in many cases an intellectually mature and mentally healthy person is responsible for their wrongdoing. Thus, the relevant notion of a match can be defined as follows:

(Matching Relation 2) A principle p matches a considered summary judgment s if and only if p accounts for s . A principle p accounts for a summary judgment s if and only if s can be inferred from p in conjunction with a set of empirical facts about the range of particular cases in question.

With this definition in hand, we can again identify two main features of the procedure of matching principles to summary judgments.

1. The inquirer compares a principled and a non-principled summary judgment about the same range of test cases, seeking equivalence between them by adjusting, if necessary, both the moral principle and the considered judgments in the

summarized range. When equivalence is achieved between the principled and non-principled summary judgments, the principle accounts for and thus matches the latter.

2. Both principled and non-principled summary judgments are crucially informed by an inquiry into the empirical facts of each test case in the summarized range.

The matching relation between moral principles and summary judgments can therefore be understood as a three-place relation: Principles and considered judgments are related to each other, while both are related to a common range of test cases. One could also say that the relevant moral beliefs only ever match *in view* of a range of test cases.

2.3 Matching principles to each other

Next, consider the assertion:

- b) *Whenever* an intellectually mature and mentally healthy person commits a wrong, they are responsible for their wrongdoing.

Note that b) differs crucially from the summary judgments mentioned in the previous subsection in that it asserts a *moral principle*. Instead of referring to an accidental set of concrete judgments about particular cases (as summary judgments do), a moral principle describes a (repeatable) type of situation (as in “an intellectually mature and mentally healthy person commits a wrong”) and yields a certain judgment for *every possible case* that instantiates this type, no matter how the remaining particulars of the case play out. The distinction between summary judgments and moral principles thus mirrors the distinction between accidental and non-accidental (i.e., necessary) generalizations familiar from the natural sphere (see Fine, 2002). Moral principles are law-like, such that they are able to sustain counterfactuals (Fine, 2002) and license deductive arguments (Rawls, 1999, p. 41).⁴ This is not true of summary judgments, since they are merely summaries of accidental sets of concrete-particular judgments.

As Rawls (1974, p. 8) plausibly suggests in the passage quoted above, some of the considered judgments relevant to RE concern principles themselves. This raises the question of how to distinguish these principle-asserting moral judgments from the “principles” against which they are to be tested in the RE process. The core difference seems to be that the latter are systematic in the sense of being part of a moral theory (Brun, 2014, p. 240). Still, the question remains as to how the corresponding matching relation is to be understood: How can a proposed moral principle (as part of a systematic body of moral theory) *match* a principle-asserting (but pre-theoretical) considered judgment? A natural answer to this question is the following: the match

⁴ I think that this characterization of moral principles is consistent with a *ceteris paribus* (or *pro tanto*) reading of moral principles. Even though *ceteris paribus* laws are non-universal (which is why the formulation “*whenever*” would be inappropriate in this case), they establish an abstract relation between a situation type and a moral consequence (even if that consequence consists only in a *pro tanto* or *prima facie* moral claim). Moreover, *ceteris paribus* laws can support counterfactuals and play a role in deductive reasoning (see Reutlinger et al., 2021).

we are looking for consists in a candidate theory containing or entailing a principle that is *identical* to some principle-asserting considered judgment to which one is already committed. The identity of two principles is in turn plausibly determined by their propositional content. However, I do not think it is necessary for a match between pre-theoretical and systematic moral principles that they have the same propositional content. Rather, it is sufficient to define the relevant notion of a match in the following, less demanding way:

(Matching Relation 3) Two moral principles p and q match if and only if all instances of q can be accounted for by p . An instance of a principle q is defined as any particular case where q yields a particular moral judgment. A moral principle p accounts for an instance of another principle q if and only if the particular moral judgment yielded by q can also be inferred from p in conjunction with a set of empirical facts about the particular case in question.

My central argument for this less demanding definition draws on the epistemic function of the matching relation. As Tersman (1993) argues, one crucial purpose of introducing moral principles and theories is to establish explanatory relations. The Rawlsian matching relation considered in this section can be conceived as an important instance of such an explanatory relation. This is well illustrated by the fact that Rawls originally spoke of an “explication” instead of a match (see Rawls, 1951). Now, if a candidate theory contains a principle to which one is already pre-theoretically committed, then it is surely appropriate to say that the theory *explains* the pre-theoretical principle. But does the same explanatory relation hold if the conditions of Matching Relation 3 are satisfied, even if the two principles in question are not synonymous or otherwise conceptually related? I think it does. Let me illustrate this with the example of two physical principles:

- c) In principle, heavy objects fall to the ground when dropped.
- d) In principle, objects with mass exert gravitational force on each other.

Certainly, most people are pre-theoretically committed to c). It is also very plausible to say that the scientific principle d) *explains* the pre-theoretical principle c). At the same time, c) and d) do not stand in an analytical or conceptual relation (like synonymy or entailment). Neither is “heaviness” synonymous with or entailed by “mass”, nor is “falling to the ground” synonymous with or entailed by “exerting gravitational force”. It is just that every concrete exemplification of principle c) is also an exemplification of—and thus accounted for by—principle d), so that the extension of the term “heavy objects fall...” in c) is a subset of the extension of the term “objects with mass exert...” in d). Meaningful scientific explanations of pre-theoretical beliefs, it seems, may very well appeal to sophisticated concepts that carve out “deep”, non-obvious features of reality and are therefore radically different from any concepts to which we are pre-theoretically committed. I see no reason why this should be any less true in the case of moral inquiry.

One additional, though related consideration in favor of my less demanding definition is that, given the pre-theoretical character of principle-asserting considered judg-

ments, it would be an implausibly strong demand on a match if these pre-theoretical principles must stand in an analytical or conceptual relation (like synonymy or entailment) to the more sophisticated, systematic principles. Especially if one endorses some form of meaning holism, it seems hardly possible that a pre-theoretical principle is synonymous to or entailed by a systematic one, because the meaning of the latter will be determined to some extent by the theory it is part of.⁵

If we accept the less demanding definition of Matching Relation 3, how exactly can we proceed to establish a match? According to my characterization above, moral principles describe repeatable types of situations and yield a certain judgment for every possible case that instantiates this type. Let us assume that a principle *p* yields a concrete judgment *c* (e.g., a judgment about moral responsibility) in every particular case that meets the description *P*, and that another principle *q* yields the same concrete judgment *c* in every particular case that meets the description *Q*. Suppose further that *q* has only actual-world instances (so that a physically impossible case will never be an instance of *q*). Now, if we knew that every particular case that instantiates *Q* necessarily instantiates *P* (so that the extension of *Q* is a subset of the extension of *P*), we would be entitled to the conclusion that in every particular case in which *q* yields a particular judgment, this same judgment can also be inferred from *p*. This, in turn, ensures the fulfillment of Matching Relation 3. In light of this reasoning, we can formulate the following corollary:

(Corollary) If there is a principle *p* that yields a concrete judgment in every particular case that meets the description *P* and another principle *q* that yields the same concrete judgment in every particular case that meets the description *Q* (and if *q* has only actual-world instances), then *p* and *q* match if and only if the extension of *Q* is a subset of the extension of *P*.⁶

⁵ Note, additionally, that in Matching Relations 1–3 it was assumed that a systematic principle and a pre-theoretical commitment can yield an identical judgment in a particular case. Given that the compared propositions are theory-guided in one case and not theory-guided in the other, the same considerations as the ones just mentioned might speak against the requirement that the judgments in question have identical propositional content. One way of solving this problem could be to require mere topic continuity instead of exact conceptual identity. If, for example, a considered judgment ascribes moral responsibility to an agent in a particular case, a matching systematic principle ascribing moral responsibility to the same agent in the same case need not, on this view, understand the notion of “moral responsibility” in exactly the same way. Rather, it is sufficient that the topic of “moral responsibility” be preserved. As a consequence of this view, the above definitions of “account for” would have to be modified to say that a pre-theoretical judgment *c* is accounted for by a systematic principle *p* if and only if *c*, or *some judgment c’ topic continuous with respect to c*, could be inferred from *p*. Such a view certainly requires more fleshing out, especially with regard to how exactly topic continuity is to be understood. For some proposals see Brun (2022, 2016), Baumberger and Brun (2021), and McPherson and Plunkett (2021). For the sake of simplicity, I will continue to speak of an identity between theoretical and pre-theoretical judgments. It should be borne in mind, however, that I am open to a less strict definition of identity that requires only topic continuity.

⁶ What if a principle has non-actual instances, such that it claims validity not only in the actual but also in some other possible worlds? One example of this would be a principle that claims metaphysical necessity, i.e., validity in all possible worlds. I think it highly unlikely that people are pre-theoretically committed to such claims. However, if they are, matching a pre-theoretical principle *q* with non-actual instances to some other principle *p* would require that the situation-type descriptions *P* and *Q* have the same intension. This means that the relevant matching procedure depends on the content of the requisite considered

We can now think of various considerations sustaining the assertion that the extension of Q is a subset of the extension of P. It seems worthwhile to look at two kinds of such considerations, namely conceptual and empirical considerations. Compare the following three principles:

- e) Whenever an intellectually mature and mentally healthy person commits a wrong, they are responsible for their wrongdoing.
- f) Whenever a person has the cognitive capacity x and the volitional capacity y at the time of their wrongdoing, they are responsible for it.
- g) Whenever a person has a fair opportunity to avoid wrongdoing, they are responsible for it.

It is conceivable that being an intellectually mature and mentally healthy person *just means* having some cognitive capacity x and some volitional capacity y . Then, naturally, following Frege's slogan that sense determines reference, the two principles also have the same extension. This in turn means that the extension of the situation type described by e) is a subset of the extension of the situation type described by f) (and vice versa), so that the two principles match according to my corollary. This is in line with the above reasoning that identical propositional content is indeed sufficient (albeit not necessary) for a match.

It is a difficult question whether the conceptual relation just described can be established independently of some larger body of background theory. However, at least in WRE, inquirers are explicitly asked to draw on theoretical considerations. Therefore, in applying WRE, we might appeal to some conceptual work in moral and/or psychological theory to reach the conclusion that principles e) and f) are indeed matching. Similarly, moral theory might provide us with the insight that possessing the cognitive capacity x and the volitional capacity y at the time of wrongdoing *is exactly what it takes* to have a fair opportunity to avoid wrongdoing⁷, thereby establishing a conceptual match between principles f) and g). Hence, an analysis of the concepts invoked by moral principles, and maybe also the theoretical explication or re-engineering of such concepts, can (and perhaps *must*) inform the matching procedure in RE.⁸ It is worth noting that even when matching proposed moral principles to considered judgments *on particular cases* (as examined in Sect. 2.1 above), we cannot do without an understanding of the concepts invoked by the principle, otherwise we

judgments: Are they to apply only in the actual world or in all possible worlds? Even if there are some principle-asserting considered judgments that are meant to hold in other (or all) possible worlds, there are certainly still other, more local judgments that are meant to hold in the actual world only.

⁷ For a similar view, see Brink (2021). To be exact, we could not establish a conceptual match between f) and g) on Brink's account, because, according to him, a set of cognitive and volitional capacities is not sufficient for having a fair opportunity to avoid wrongdoing. Additionally, he requires the fulfillment of certain external or situational conditions providing for what he calls "situational control".

⁸ Brun (2022) underlines that in the RE process, theory development involves not only articulation but re-engineering of contested concepts. However, in the spirit of Carnap's (1950) requirement of similarity, it seems important that the claimed conceptual identity—which in turn provides for the matching relation—does not depart too far from the common usage of the concept in question, otherwise it would be implausible to speak of a match between (pre-theoretical) considered judgments and theory-guided moral principles.

would be unable to decide whether the principle applies to the case in question. Also, if it is true that the meaning of a concept cannot be determined without some larger body of background theory, we have another reason to think that NRE—insofar as it relies on any of the aforementioned matching procedures, which in turn rely on a conceptual understanding of moral principles—is not so narrow at all.

Moreover, the assertion that the extension of Q is a subset of the extension of P can be sustained by laws of nature. Let us assume for the sake of argument that we cannot establish a conceptual match between e) and f). Still, it is conceivable for there to be a law of nature ensuring that whenever a person is intellectually mature and mentally healthy, they also *happen to have* the cognitive capacity x and the volitional capacity y. Then, by natural necessity, each particular case that instantiates the situation type described by e) also instantiates the situation type described by f), so that the extension of the situation type described by e) is a subset of the extension of the situation type described by f).

More realistically, there might be a probabilistic empirical relation such that whenever a person is intellectually mature and mentally healthy, they are also *likely to have* the cognitive capacity x and the volitional capacity y. Naturally, this consideration would not provide a perfect match between e) and f), as there are still instances of e) that are not accounted for by f). But due to the probabilistic empirical relation, these instances are relatively rare. Call this situation a *decent match*. Since it is only decent and not perfect, i.e., there are remaining inconsistencies between the two principles in question, either the pre-theoretical or the systematic (or both) will have to be revised as the RE process continues. Still, finding a decent match shows inquirers that they are “almost there”. Reforming either of the principles involved to achieve a perfect match will have little moral consequence, quantitatively speaking, because there are already relatively few instances in which the two principles in question yield diverging judgments. This observation may prove useful in making pragmatic decisions in the course of the RE process. In particular, if RE is also understood as a dialogical method—as in the example of law reform considered in Sect. 3.2 below—it might be prudent to prioritize major inconsistencies over already decent matches when deciding where to invest time and resources in the pursuit of a state of complete equilibrium. Note, however, that the notion of a decent match is defined in purely quantitative terms. Among the few cases in which two decently matching principles yield diverging judgments might be cases that we deem particularly important. Such a situation may motivate the pragmatic decision to give priority to resolving the remaining inconsistencies, even if the principles in question are already decently matching.

In subsections 2.1–2.3, I have sought to demonstrate in which precise way the procedure of matching moral principles to considered judgments—as a core element of any (narrow or wide) method of RE—relies on empirical beliefs. This account can be linked to a more general coherentist framework, such as the one proposed by Tersman (1993). Tersman’s theory seems particularly amenable to comparison with my account, as he also discusses in some detail the relation between moral and nonmoral beliefs in RE (1993, ch. 3).

2.4 A more general framework according to Tersman (1993)

Tersman's coherentist approach starts from the assumption that the relevant notion of coherence describes a relation between a particular belief and an agent's belief system. Now, according to Tersman, a necessary (though not sufficient) condition for a belief p to cohere with one's system S is that there be a justificatory argument for p in S based on premises that are themselves supported by justificatory arguments (1993, p. 40).

In this framework, the Rawlsian matches examined in the previous subsections serve as a particular type of justificatory argument for moral beliefs (see Tersman, 1993, p. 44, p. 50): On the one hand, principles may support less general judgments by explaining and entailing them. On the other hand, less general judgments may support principles as an effect of being explained by them. These relations—which are equivalent to the Rawlsian matches examined in the previous subsections—are thus important relations of evidential support within his RE framework (Tersman, 1993, p. 50).

How, according to this picture, are nonmoral beliefs relevant to the evidential support of moral beliefs? Tersman's comments on this are quite general and rather vague, perhaps due to the generality of his overall approach, but they are entirely consistent with the findings of the previous subsections. Indeed, Tersman claims that “[m]oral beliefs often explain, entail, or conflict with other moral beliefs only if some context of nonmoral assumptions is presupposed” (1993, p. 54). Given that this relation of mutual explanation, entailment, and conflict is an essential part of justifying moral beliefs, it follows that “some of the beliefs that are positively relevant to our moral beliefs are likely to be nonmoral” (1993, p. 55).⁹ As far as the evidential relation provided by Rawlsian matches is concerned, the findings of the previous subsections allow a more precise formulation of Tersman's general point:

- (1) If Matching Relation 1 serves as a justificatory argument for a moral belief, this argument *necessarily* contains empirical premises (as without them the match could not be established).
- (2) If Matching Relation 2 serves as a justificatory argument for a moral belief, this argument *necessarily* contains empirical premises (as without them the match could not be established).
- (3) If Matching Relation 3 serves as a justificatory argument for a moral belief, this argument *possibly* contains empirical premises (as some of these matches can only be established by reference to laws of nature).

Tersman (1993, pp. 55–56) continues to argue that insofar as moral beliefs gain evidential support from arguments containing nonmoral premises, the justification of our moral beliefs also requires that we are justified in holding these nonmoral prem-

⁹ In Tersman (1993), the relation of “positive relevance” is a core notion of evidential support. Essentially (and somewhat simplistically), a belief p is positively relevant to another belief q if and only if p serves as a premise of a valid justificatory argument for q . See Tersman (1993, p. 38) for further necessary specifications.

ises. This highly plausible demand, in conjunction with my assertions (1)–(3) above, underscores the importance of investigating empirical facts when applying the RE method. I will now show how empirical developments may upset RE and motivate changes in the principles we adopt, precisely by affecting the matching procedure I have just described.

3 Changes in the empirical world, changes in the equilibrium

According to the above account of the matching relation, empirical beliefs are crucial to the process of establishing matches between moral principles and considered judgments. Changes in our empirical beliefs, possibly due to actual changes in the empirical world or to new empirical discoveries, are therefore likely to affect the matching relation. In particular, once we have arrived at a match between a set of moral principles and a set of considered judgments—and thus have reached a state of (partial) RE—new empirical developments may disrupt this state and necessitate further changes in our moral principles or our considered judgments, or both. Let me explain exactly how this happens.

First, consider Matching Relation 1 above and recall that both principled and non-principled judgments are crucially informed by the empirical facts of a particular case. Let us call the sum of empirical beliefs about a given case M . We can now specify the above definition of Matching Relation 1 as follows: a match consists in the fact that a considered non-principled judgment c (which is informed by M) is also inferable from a principle p in conjunction with some subset of M . Suppose that in a particular case p matches c with respect to M . But what happens if our empirical beliefs about this case change, resulting in a new set of beliefs M^* ? Possibly M^* leads us to modify our considered judgment about the particular case in question (call this modified judgment c^*). The judgment c^* , however, may not be inferable from p in conjunction with some subset of M^* . Moreover, p conjoined with some subset of M^* may yield a modified principled judgment c^{**} . If the considered non-principled judgment, on the other hand, changes to c^* (or even remains unchanged), p can no longer account for it. A new set of empirical beliefs M^* about a particular case can therefore disrupt Matching Relation 1 in multiple ways. Restoring the match will usually require a change of the principle or the considered judgment, or both.¹⁰

Analogously, for Matching Relation 2, we can define a set of empirical beliefs N that comprises all empirical beliefs about a given range of test cases. According to the same reasoning already demonstrated for Matching Relation 1, changes in N resulting in a new set of empirical beliefs N^* may well disrupt Matching Relation 2 between a considered summary judgment and a moral principle.

Put less formally, changes in our empirical beliefs present us with new test cases for Matching Relations 1 and 2. Even if some set of moral principles has hitherto matched our considered judgments, it is by no means clear that the match can be

¹⁰ Perhaps the match could also be restored by revising the set of empirical beliefs M^* . I will consider the possibility of revising our empirical beliefs in light of certain moral convictions in Sect. 5 below. I believe that this possibility does exist, but is restricted to a specific set of circumstances.

upheld in the face of these new test cases. It follows that a state of (partial) RE, insofar as it relies on the above Matching Relations 1 and 2, is stable only relative to a set of empirical beliefs. If these beliefs change, the achieved equilibrium is likely to be upset.

The mechanism described seems particularly important bearing in mind that considered judgments are not constrained by any theory or principle. Consequently, *all available empirical information* about a particular test case can potentially be relevant to a considered judgment. Moral principles, on the other hand, select only a limited number of empirical case features as relevant. For example, if one endorses an act-utilitarian principle, only information about the consequences of acts is relevant to any concrete judgment that follows from the principle. In short, principled and non-principled judgments *differ in flexibility*. It is therefore very likely, in my view, that moral principles and considered judgments often respond differently to new empirical developments, which in turn leads to a disruption of previously accepted matching relations. Hence, the difference in flexibility between principled and non-principled moral beliefs is an additional source of instability for RE in the face of new empirical developments.

Finally, consider Matching Relation 3 and recall that it can be established by reference to laws of nature. Changing our beliefs about certain laws of nature can thus directly affect our knowledge of the matching relation, both by disrupting previously accepted matches and by revealing new matches that we previously did not know about. Similarly, changing beliefs about probabilistic empirical relations directly affect our knowledge about decent matches. This in turn may influence some of our pragmatic decisions in the RE process.

In conclusion, RE can account well for the fact that when our empirical beliefs change, our moral beliefs very often change with them. Let me illustrate this with two examples.

3.1 Van Inwagen's dilemma

Consider the following position proposed by van Inwagen (1983). First, according to van Inwagen, free will is incompatible with the truth of determinism, a thesis that depends crucially on his famous transfer principle (β).¹¹ Second, free will exists. This latter thesis rests on two premises, namely the undeniable reality of moral responsibility and the fact that moral responsibility entails free will.

After having defended both theses, van Inwagen (1983, pp. 219–221) addresses a potential threat to his position: what if science were to provide convincing reasons—as he believes is possible in principle—that determinism (as a physical theory about the world) is true, or at least that human beings should be regarded “for all practical purposes” as deterministic systems? Such a case would confront van Inwagen with

¹¹ The principle (β) reads as follows: $Np, N(p \rightarrow q) \vdash Nq$. “Nx” abbreviates the proposition that “x and no one has, or ever had, any choice about whether x” (van Inwagen, 1983, pp. 93–94). The transfer principle (β) is at the heart of van Inwagen’s Consequence Argument for incompatibilism, which states that, on the assumption that determinism is true, since one has no choice about the remote past and about what the laws of nature are, one also has no choice about the consequences of these things, including one’s own present acts (1983, p. 56).

the dilemma that his two theses above cannot be true together. In this hypothetical situation, van Inwagen argues, he would probably abandon his transfer principle (β)—and thus his incompatibilism—because he is even more deeply convinced of the existence of moral responsibility than of his incompatibilist position.

In this line of reasoning, van Inwagen does not invoke RE. In fact, since he views the existence of moral responsibility as a kind of foundational truth (see van Inwagen, 1984, p. 209) from which the correct solution to the dilemma is deduced, his argument fits better with strong or moderate forms of foundationalism (following the terminology of Bonjour, 1985). RE, on the other hand, is commonly associated with coherentist or only weakly foundationalist epistemologies.

However, the truth of moral responsibility judgments as claimed by van Inwagen is a highly contentious matter. Instead, one could argue—presumably less controversially—that the proposition

We are often morally responsible for what we have done.

is a summary considered judgment that many people are inclined to make. In fact, van Inwagen's case for the truth of our everyday judgments of moral responsibility rests mainly on the pervasiveness and importance of judgments of moral responsibility in our getting along in the world (1983, pp. 206–209). While it is by no means clear how these observations can speak to the *truth* of responsibility judgments, they do support the claim that most people would, upon due consideration, say that human agents are often responsible for what they do.

Now, if the moral responsibility of human agents is not an epistemically basic truth but a summary considered judgment, then van Inwagen's (hypothetical) reaction to scientific evidence for the truth of determinism can be well reconstructed as an application of RE. Assuming determinism to be false, an incompatibilist principle *matches* a summary considered judgment affirming the moral responsibility of human agents. However, this match would be compromised if we discovered scientific evidence for the truth of determinism: while the summary considered judgment would plausibly remain unchanged, an incompatibilist principle would now deny the moral responsibility of human agents in all (physically) possible test cases. Such a discrepancy will usually be met by revising either the considered judgment or the principle.¹² Given the pervasiveness and importance of judgments of moral responsibility in our getting along in the world, we would probably have good reason to abandon the incompatibilist principle and stick to our summary considered judgment affirming the moral responsibility of human agents.

My intention is not to present a conclusive argument that van Inwagen's reasoning is *more plausible* when reconstructed as an application of RE.¹³ For now, I merely wish to show that this reconstruction, if accepted, provides a nice example of a (partial) RE being upset by an empirical discovery.

¹² Again, the discrepancy could be met also by revising the empirical claim. See Sect. 5 for a detailed discussion of this possibility (which I take to be restricted to a specific set of circumstances).

¹³ In Sect. 5, however, I will offer some reflections on van Inwagen's argument for libertarianism and suggest that it could perhaps be improved if he were to adopt WRE.

3.2 Double jeopardy reforms

Legal reforms are often prompted by the emergence of new cases that pose problems for established legal norms or principles. In this respect, legal reforms can be well reconstructed as the application of (dialogical) RE. A good example is the 2003 double jeopardy reform in the United Kingdom. Historically, the principle of double jeopardy was introduced to provide absolute protection against the risk of re-prosecution for the same offense after a definitive court acquittal. Around the turn of the century, however, there was growing concern in British politics that this absolute protection could lead to injustice, particularly in a number of cases at the time in which acquitted individuals were reported to have subsequently confessed their guilt (see Coffin, 2010, p. 773). These included the case of William Dunlop, who repeatedly admitted (e.g. to his prison nurse, in letters to friends, and in a custody proceeding) to the murder of Julie Hogg, an act of which he had previously been acquitted by the Crown Court in 1991.

Moreover, concerns about the absolute force of the double jeopardy principle were fueled by scientific advances in DNA testing:

[...] blood samples taken at a murder scene in the early 1980s might not have produced sufficient identification evidence at that time. The prime suspect may have been prosecuted on the basis of other evidence. If the prosecution failed to satisfy the jury that the defendant was guilty beyond reasonable doubt, the defendant would have been acquitted and left the court a free man. A decade later, advances in DNA testing could enable the original blood samples to be analysed and show with near certainty that the acquitted person had been at the crime scene.”¹⁴

These concerns led to the passage of the Criminal Justice Act of 2003, which codifies an exception to the principle of double jeopardy for serious offenses “if there is new and compelling evidence against the acquitted person”.¹⁵ A very similar attempt at reform in Germany in 2021 was prompted by just such a “cold case” from the early 1980s—namely the murder of Frederike von Möhlmann—in which a subsequent re-examination of the evidence using modern DNA testing heavily incriminated a previously acquitted suspect.¹⁶

What happened in these examples can plausibly be interpreted as a (partial) RE being upset by empirical developments. The emergence of new test cases involving subsequent confessions and modern DNA testing disrupted the match between the principle of double jeopardy and people’s considered judgments. In these new cases, the principle of double jeopardy yielded an unchanged judgment, namely that the acquitted are absolutely protected from re-prosecution. However, in the same cases, it seemed wrong or unjust to many people, upon due consideration, to grant the sus-

¹⁴ The Law Commission, *Double Jeopardy and Prosecution Appeals*, 2001, Cm 5048, § 1.5.

¹⁵ Criminal Justice Act, 2003, c. 44, § 78 (U.K.).

¹⁶ See Hörnle (2022) and Slognsat (2021) for a detailed discussion of the 2021 double jeopardy reform in Germany. In 2023, the reform was overturned by the German Federal Constitutional Court.

pect protection from re-prosecution. Thus, the considered judgments on the newly emerged cases were not accounted for by the principle of double jeopardy in its previously accepted absolute form. The discrepancy was resolved, as a consequence, by reforming the principle. Note further that the constellation in the present example is exactly opposite to the constellation in the previous example of van Inwagen's dilemma. In the latter, the principled judgment changed in response to new empirical developments, while the non-principled considered judgment remained the same. In the present example of double jeopardy reform, by contrast, the non-principled considered judgment changed while the principled judgment remained the same.

Up to this point, I have given a detailed account of the relevance of empirical beliefs to the formation of moral convictions in the RE process. While such an account may be interesting in its own right, it is worth asking whether it has any implications for the general appeal of RE as an approach to moral justification. This question will be pursued in the following section.

4 Some implications for the RE debate

My account of the role of empirical beliefs within RE can contribute to the assessment of the latter's overall philosophical appeal in at least two ways. First, my account demonstrates the empirical adaptability of RE, which in turn may count as a theoretical advantage of the method. Second, my account can help assess (and partly answer) a criticism of the method raised by Kelly and McGrath (2010). I will address each of these two points in turn.

4.1 The advantage of empirical adaptability

As seen above, any of the Matching Relations 1–3 may be upset by empirical developments, and in restoring the match, no privileged group of beliefs is immune to revision. This means that even abstract moral principles can be adjusted and refined in response to empirical developments. Call this characteristic *empirical adaptability*. Recognizing this empirical adaptability can, I believe, contribute to a broader discussion about the adequacy of RE as an approach to moral justification. Specifically, I wish to argue that empirical adaptability may be seen as an advantage of RE.

It is only against the background of a certain metajustification that some feature of a moral epistemological approach can plausibly be considered an advantage. In other words, the advantages and disadvantages of a theory of moral justification depend on what criteria such a theory must fulfill to count as true or convincing. The classic answer to this question is that justification should be truth-conducive (Bonjour, 1985, p. 9; Kappel, 2006). Now, unless we make some contentious metaethical assumptions, it is by no means clear that tracking moral truth requires an empirically adaptable moral epistemology. Thus, I do not think that my account of the role of empirical beliefs within RE can by itself (that is, without additional metaethical assumptions) contribute to assessing the truth-conduciveness of RE. However, there may be general problems with the criterion of truth-conduciveness. For example, Tersman (1993, p. 99–102) argues that with any given theory of epistemic justifica-

tion, it is difficult to find a neutral, non-question-begging way of determining whether it is properly connected to truth.

In any case, there are alternative approaches to the problem of metajustification. One is the claim that a moral epistemological theory should account for our actual argumentative practice (Tersman, 1993). Now, the criteria and considerations employed in actual moral practice seem to be highly heterogeneous across different empirical contexts, and it is unclear whether this state of affairs can be accounted for by a “universal grammar” of unchanging basic principles (Birnbacher, 1999).¹⁷ And even if there is a fixed set of basic principles at some point, this need not be a necessary state of affairs, i.e., it may be contingent on a particular time and space and liable to change with empirical developments. Consider again the above example of double jeopardy reforms, where an important principle of criminal procedure was revised in response to a range of newly arisen cases and advances in DNA testing. This dialectical process exemplifies the empirical adaptability of actual argumentative practice. The fact that RE can neatly account for this should count as an advantage of this theory, assuming Tersman’s (1993) metajustification.

Alternatively, one could take up a pragmatic conception of moral justification by asking how it should be understood in light of some core functions of moral discourse. One such function might be to provide convincing solutions to social conflicts. Given the great variety of conflicts that can arise empirically (and that one cannot even imagine yet), there may be pragmatic reasons for adopting an epistemology that is characterized by a high degree of adaptability to different empirical contexts.

Of course, if empirical adaptability is to provide a reason to favor RE over alternative approaches to moral justification, one must show that RE does *better* than other theories in terms of empirical adaptability. I do not wish to present a comprehensive discussion to that effect here, as this is beyond the scope of this article. However, I suspect that while RE is certainly not the only moral epistemology that offers a high degree of empirical adaptability, it does do better in this respect than some of its theoretical competitors. Take, for example, a foundationalist approach to applied ethics as described by Birnbacher (1999, pp. 321–322). According to this approach, applied ethics essentially consists of “a deduction of concrete consequences from a set of basic ethical principles together with a number of empirical premises” (ibid., p. 322). In other words: when it comes to concrete empirical contexts, the epistemology described encourages a purely deductive mode of reasoning that proceeds from context-insensitive basic principles. I think that such a method offers a far lower degree of empirical adaptability than does RE, which, as I have argued above, allows for a constant and successive adjustment of moral principles in response to empirical developments and scientific discoveries.

I turn now to the second way in which my account of the role of empirical beliefs within RE may contribute to the debate on the theoretical adequacy of RE. Specifi-

¹⁷ Interestingly, according to Birnbacher (1999), this question can be investigated empirically, as one might ask whether there are decisive empirical differences between different contexts that—according to some fixed set of basic principles—warrant diverging “surface grammars” of particular moral judgments or practice rules.

cally, the account presented here can help assess and to some extent rebut a purported counterexample to the method presented by Kelly and McGrath (2010).

4.2 The transcendent metaphysics counterexample against coherentist methods of moral justification

There is a familiar concern that considerations of coherence as sanctioned by RE are too weak to provide a plausible justification for moral claims.¹⁸ Imagine, for example, that somebody started from a set of confidently held astrological beliefs and, in applying the method of RE, managed to bring them into perfect coherence. Certainly, we would not consider these astrological beliefs justified despite the internal coherence among them. In discussing this analogy, Scanlon (2002, p. 146) points to an important difference between astrology and ethics: astrology is committed to causal claims about physics and psychology that are clearly false. Ethics, on the other hand, is not in an analogous situation since it has no such external commitments.

Kelly and McGrath (2010, pp. 343–344) countered Scanlon’s argument with a more challenging comparison than astrology, namely theology understood as “transcendent metaphysics”—the latter being characterized by the fact that its claims make no difference whatsoever to the empirical world, and thus cannot even in principle be contradicted by empirical claims. Now, if a religious believer produced a perfectly coherent theological system of divine entities that qualified as transcendent metaphysics, we would still not consider these religious beliefs justified even though they would be internally coherent and not contradicted by (or even in tension with) claims from other domains like physics or psychology. Analogously, as Kelly and McGrath (2010, p. 344) conclude, “[...] it is not enough for our moral judgments to be justified, that they cohere well with one another and are not contradicted by well-confirmed views from outside of morality.”

Yet there is a potential difference between the coherence relation of moral beliefs within RE and an internally coherent piece of transcendent metaphysics. As stated above, in applying RE, the coherence between moral beliefs is likely to be upset by empirical changes and discoveries. On the other hand, it does not seem clear how a piece of transcendent metaphysics, understood as “its own self-contained subject matter” (Kelly & McGrath, 2010, p. 343), could be as sensitive to changing empirical beliefs. While the astrology comparison goes wrong because moral beliefs are *not directly contradicted* by empirical claims, the transcendent metaphysics comparison goes wrong because within RE moral beliefs *are not completely independent* of empirical claims either.

But perhaps a counterexample can be found that better fits the coherence relation of moral beliefs within RE. We could imagine, for instance, a kind of nature reli-

¹⁸ For the sake of simplicity, I understand “justification” in this context as the kind of justification that figures in the traditional analysis of propositional knowledge. It should be noted, however, that RE has also been proposed as a method aimed at objectual understanding rather than propositional knowledge (Baumberger & Brun, 2021; Elgin, 2017). Yet I take the discussion of this subsection to be a variant of the *ex-nihilo* objection, which seems to be a problem for RE-based accounts of objectual understanding as well (see Baumberger & Brun, 2021, pp. 7933–7935; Elgin, 2017, p. 75). Thus, I think the considerations of this subsection are also applicable to RE-based theories of objectual understanding.

gion in which judgments about transcendent metaphysical occurrences—such as the activities of nature spirits or ghosts—are always linked to specific empirical events or “cases.” Empirical changes or additional knowledge about empirical events can thus influence the believer’s assessment of what is going on in the metaphysical world of nature spirits. Again, rendering these nature-religion judgments into a perfectly coherent form (against the backdrop of the relevant empirical events or “cases”) would probably not provide them with sufficient justification.

But what exactly do these examples suggest? Perhaps they show that systems of belief can be internally coherent and at the same time radically detached from reality. Kelly and McGrath formulate this general point as follows:

“After all, the coherence of a system of beliefs is, presumably, something that supervenes on the relations that obtain between those beliefs, as opposed to any relations that obtain between those beliefs and anything outside the system. But this makes salient the possibility that a system of beliefs could be arbitrarily coherent while being radically detached from the very subject matter that it purports to accurately represent. [...] [T]he problem is that, at least in principle, an individual might maintain a perfectly coherent set of beliefs while being completely unresponsive to relevant and easily perceptible changes in his or her environment” (2010, p. 333).

This “No Contact with Reality” objection (as Kelly and McGrath call it) seems closely related to the familiar objection that mere coherence does not provide a plausible indication of truth (and is therefore inadequate in light of the metajustification of truth-conduciveness mentioned above). In the case of moral epistemology, this would amount to the worry that even if one has reached a state of RE, one may be radically detached from moral reality. As suggested earlier, I do not think that my account of the role of empirical beliefs within RE can—at least without some contentious meta-ethical assumptions—contribute to the debate about the truth-conduciveness of RE. Therefore, I will not address further this dimension of detachment, which pertains to the possible lack of connection to an independent reality.

However, there is another side to the worry of radical detachment that might explain at least part of the intuitive response to the two examples above. Transcendent metaphysics, conceived as “its own self-contained subject matter” (Kelly & McGrath, 2010, p. 343), is also characterized by a lack of connection to beliefs about other subject matters. Indeed, it is a common idea among coherentists that in the curved and ultimately circular structure of coherentist justification, large circles are better than small ones (Harman, 1986, p. 33; Tersman, 1993, p. 42). Accordingly, the more a belief is entrenched and integrated into a person’s overall system of beliefs, the better it coheres with the system and the higher its degree of justification (Tersman, 1993, p. 43). Whether or not a belief is well entrenched in turn depends on whether it obtains evidential support by a variety of considerations from different domains of the belief system (*ibid.*).

This view can account for the intuition that it is unreasonable to hold a religious belief system that qualifies as transcendent metaphysics, even if it is internally coherent. Since we are dealing with a “self-contained subject matter”, any argument in

support of a transcendent metaphysical belief will presumably rest exclusively on premises that themselves qualify as transcendent metaphysical. In light of the above account of Rawlsian Matching Relations 1–3, we can now see a clear difference between transcendent metaphysics and a system of moral beliefs in RE. Recall, specifically, the necessity claims (1) and (2) in Sect. 2.4 above, according to which any justificatory argument built on Matching Relations 1 and 2 *necessarily* contains empirical premises. This means that insofar as Matching Relations 1 and 2 play an evidential role in the achieved state of (partial) RE, the resulting system of moral beliefs has a certain degree of *necessary entrenchment* in a person's empirical beliefs.¹⁹ I conclude that almost any system of moral beliefs held in RE—or, more precisely, any system built at least in part on Matching Relations 1, 2, or the empirical variant of Matching Relation 3—is significantly *more justified* than any coherent system of transcendent metaphysics conceived as its own self-contained subject matter. It is another question whether this minimal, necessary degree of entrenchment is *sufficient* for justification. The above example of a nature-religious belief system (which is modeled to contain the three-place Rawlsian matching relations defended in this article) suggests that it is not.²⁰

The response to the transcendent metaphysics counterexample defended in this subsection is similar to, but also somewhat different from, that of Tersman (2018). Tersman argues that in WRE, in contrast to a coherent system of transcendent metaphysics, empirical beliefs provide support for our moral beliefs, e.g., by vindicating the credibility of our moral judgments with the help of an evolutionary account of their origins. In fact, Tersman believes the transcendent metaphysics example to show that, in the method of RE, some such empirical support is necessary to plausibly justify our moral beliefs. I fully subscribe to the latter claim, which is also consistent with the entrenchment criterion explained above, but think that this necessary empirical support can be had more cheaply than with a contentious evolutionary vindication argument. My account shows that some degree of empirical support is *necessarily inherent* in Matching Relations 1–3, and therefore necessarily inherent in any state of RE that relies on these central relations of evidential support.

Thus, to respond to Kelly and McGrath's counterexample, it suffices to point out the necessary empirical entrenchment of moral beliefs within RE. Recall, however, that this necessary entrenchment is not *sufficient* for justification. The evolutionary vindication argument cited by Tersman (2018) might provide a fuller, perhaps sufficient, justification of moral beliefs—and thus refute the nature-religion counterexample mentioned above. Although I think this is possible, I see no reason to believe that the credibility of our moral judgments must necessarily be vindicated by an *empiri-*

¹⁹ Note that this view presupposes (plausibly, I think) that entrenchment comes in degrees.

²⁰ Note, however, that the argument presented here seems to imply that the nature-religious belief system just mentioned is *more justified* than a theology that qualifies as transcendent metaphysics. Is this consequence reasonable? I think it is at least not obviously mistaken. Kelly and McGrath (2010, p. 343) note that most actually held religious belief systems are not plausibly interpreted as transcendent metaphysics because they typically involve some sort of intervention in or interaction with the empirical world (even if only through its creation) on the part of divine entities. I suggest that this may well be so because it is indeed *less reasonable* to believe in a completely self-referential, detached transcendent metaphysical theology than in a theology that has at least something to do with the physical world we experience.

cal argument. The necessary empirical entrenchment of moral beliefs within RE can by itself alleviate the worry that our moral beliefs become radically detached from the empirical world. Now, setting aside this worry, it is not clear to me that what is missing for sufficient justification is necessarily further *empirical* support. Providing sufficient justification for our moral beliefs is probably a complicated matter, requiring a variety of arguments from many different domains. However, neither Kelly and McGrath's transcendent metaphysics counterexample nor my nature-religion counterexample shows that empirical (perhaps evolutionary) arguments for the credibility of our moral judgments must necessarily be a part of it.

5 The opposite direction: adjusting empirical beliefs in light of moral beliefs

So far, I have considered the role of empirical beliefs within Rawlsian matching relations and how moral beliefs may be revised in light of empirical beliefs. I turn now to the opposite direction of impact. Thus, my question in this subsection is whether and how empirical beliefs may be adjusted in light of certain moral convictions (assuming that coherence considerations can in principle affect the justification of empirical beliefs²¹). I think that van Inwagen's "An Essay on Free Will" again provides a good example for this discussion. I will therefore examine in some detail van Inwagen's deductive inference from beliefs about moral responsibility to the existence of free will, and show why it ultimately fails (at least if one denies a certain implausible variant of moral foundationalism). After that, I will be in a position to make a more general point about the adjustment of empirical beliefs in light of moral beliefs within the WRE framework.

5.1 Van Inwagen's deductive argument

A straightforward way in which moral beliefs might influence empirical beliefs is through a deductive argument from moral premises to empirical conclusions. Such an argument is used by van Inwagen (1983, pp. 206–213) to defend his view that free will exists. His argument can be stated in the following way:

(Premise 1) We have good reason to believe that the existence of moral responsibility entails the existence of free will (in the incompatibilist sense that requires the empirical falsity of determinism²²).

(Premise 2) We have good reason to believe that moral responsibility exists.

(Conclusion) We have good reason to believe that free will exists.

²¹ If, on the other hand, one advocates some form of moderate or strong foundationalism according to which all empirical knowledge is ultimately derived from certain *empirical* basic beliefs, the question of this subsection should appear misguided at the outset.

²² In what follows, the term "free will" is understood exclusively in this specific sense.

To assess the epistemic merits of this argument, it is useful to ask why the first premise is supposed to hold true. One possible reason lies in moral and/or conceptual considerations about the correct attribution of moral responsibility, in particular, that a necessary condition for the correct attribution of moral responsibility to an agent is that they possess free will. But if so, one surely cannot justify a belief in the instantiation of moral responsibility *without presupposing* a belief in the instantiation of free will, unless one wishes to adopt an unattractive form of intuitionism in which justifying a moral responsibility ascription does not require any reference to the necessary conditions for that ascription.²³ Now, since the conclusion of the argument must already be presupposed to justify the second premise, the argument is epistemically circular (though valid) and leads to no new insight.

To avoid circularity, one could deny that the necessary relation between moral responsibility and the possession of free will is provided by moral and/or conceptual considerations about the correct attribution of moral responsibility. But what then provides it? Are there any plausible candidates for a physical or metaphysical necessary relation between the two notions in question? The only plausible candidate I can think of is the relation of (partial) metaphysical identity, meaning that being morally responsible (partly) *consists in* possessing free will. But then the problem of epistemic circularity reappears. For surely one cannot justify belief in the reality of a thing *without presupposing* the reality of its (partly identical) components, unless one wishes again to adopt an unattractive form of intuitionism.

So far I have assumed that the “good reason” for believing the second premise consists in a justificatory argument from other beliefs (which already contain the conclusion of the argument). One way to avoid epistemic circularity, then, is to claim that the justification of the second premise does *not* depend, inferentially or otherwise, on other beliefs, and that this independent justification is sufficient to confer justification on further beliefs, such as the belief that free will exists.²⁴ This would amount to a (moderately or strongly) foundationalist epistemology in which the belief in the reality of moral responsibility takes on the role of an “unmoved mover” not only of the moral but also, by way of deductive argument, of the empirical epistemic realm.²⁵ I see two main challenges in defending this view. First, one would need to give an account of exactly how the belief in the reality of moral responsibility can be justified if not by reference to other beliefs. In particular, one would have to explain how this basic moral belief can in some sense be even *more* basic than certain empirical convictions, given that the latter are supposed to be inferable from the former. Second,

²³ Indeed, van Inwagen seems to employ exactly such an intuitionist epistemology when he establishes the truth of the second premise: “That we are convinced that we know something does not, of course, prove that we do know it or even that it is true. But it *is* true that we are morally responsible, isn’t it? And we *do* know it to be true, don’t we?” (1983, p. 209, *emphases his*). I think, however, that we should not accept such an (implausible-seeming) epistemology without argument.

²⁴ Alternatively, one might claim that there need be no good reason for accepting the second premise. Yet I see no way how epistemic justification can rest on arbitrarily chosen beliefs that one has no reason to accept at all. To avoid epistemic circularity, therefore, it is much more plausible to concede that there must be a reason to accept the second premise, but to deny that this reason must consist in some further *belief*.

²⁵ The characterization of basic beliefs as “unmoved movers of the epistemic realm” goes back to Chisholm (1977, p. 25).

a proponent of this view would need to provide general reasons for its acceptance, i.e., reasons that are independent of the ad hoc purpose of getting van Inwagen's free-will argument off the ground. Although van Inwagen's intuitionist-sounding justification of the reality of moral responsibility may perhaps qualify as foundationalist in the sense just described, it is introduced only casually and does not meet these two explanatory challenges that I think must be overcome to make it a plausible epistemological option. In any case, my main concern in this paper is with RE-based epistemologies. Therefore, I will not further address moderately or strongly foundationalist interpretations of van Inwagen's argument. As will become clear shortly, I think it is already a valuable insight that van Inwagen's deductive argument does not work in the framework of RE.²⁶

Van Inwagen attempts to add plausibility to his free will argument by presenting a similarly structured argument, the acceptance of which, in his view, commits us to accept his free will argument as well. His analogous argument reads as follows:

(Premise 1) We have good reason to believe that if most of the propositions we think we know to be true *are* true, then there exists no Cartesian Universal Deceiver.

(Premise 2) We have good reason to believe that most of the propositions we think we know to be true *are* true.

(Conclusion) We have good reason to believe that there exists no Cartesian Universal Deceiver.

I suggest that, for reasons similar to those given against van Inwagen's free will argument, we should be hesitant to simply accept his argument against the existence of a Universal Deceiver (though there is certainly much more to be said about this kind of anti-skeptical argument). This time, however, let us grant both premises and ask whether they license the conclusion. Naturally, if our best method of inquiry gives us reason to believe that some proposition about the world is true, then *by the standards of this very method* we are not deceived about this proposition. This inference is valid but raises, once again, the problem of epistemic circularity. For, in any reasonable method, one certainly cannot establish the correctness of some proposition without at the same time excluding the possibility that one is deceived about it. Also here then, since the conclusion of the argument must already be presupposed to justify the second premise, the argument is epistemically circular (though valid) and leads to no new insight.

To avoid circularity, one could take the inference to establish that if our best method of inquiry gives us reason to believe in the correctness of some proposition about the world, then *by some independent standard* we are not deceived about this proposition. Such an argument, however, seems to be simply incorrect (Cohen, 1984;

²⁶ One might object that the epistemic circularity of van Inwagen's argument is not a problem, at least for coherentist interpretations of RE, since, according to coherentism, *all* justification is ultimately circular. Recall, however, that coherentism aims at large circles rather than small ones. Thus, an argument in which the justification of one of the premises depends quite directly on the justification of the conclusion (which I think is true in the case of van Inwagen's argument) is of little epistemic value even in a coherentist framework.

Kelly & McGrath, 2010). Assuming there is a truth about the presence or absence of a Cartesian Universal Deceiver that is independent of our best methods of inquiry, how can these methods guarantee that we will find it and thus license the inference in question?

Since van Inwagen's anti-skeptical argument meets with similar doubts as his free will argument, the former does not help van Inwagen's cause to prove the epistemic merits of the latter. Even if one wishes to assent to the anti-skeptical argument, one can still raise reasonable doubts about the analogy: Why, after all, should we assume that the relation between justification and truth is anything like the relation between moral responsibility and free will?²⁷

5.2 An alternative account

If van Inwagen's deductive argument for the existence of free will does not work (at least in the framework of RE), does this mean that he is forced to take an agnostic position on the matter? Compare the position of van Inwagen's imaginary skeptic, who says:

“Therefore, your arguments represent just one more attempt by a philosopher to settle by intellectual intuition and pure reason a question [namely the question about the truth of determinism] that should be left to empirical science. And if it should prove that this question can't be settled by empirical science, owing perhaps to ‘the obscurity of the matter and the shortness of human life’, then we should simply elect to have no opinion about the right answer to it” (1983, p. 210).

I think that if van Inwagen were to adopt the method of WRE, he would not have to take an agnostic position. More generally, I wish to defend the claim that in applying WRE, when in empirical doubt, we may have reason to prefer an empirical belief that ensures a higher degree of *ethical* consistency (all other things being equal).

To see why this is so, it is useful to examine how the central notion of coherence should be explicated. I take the above Matching Relations 1–3 as specific instances of a coherence relation, but there are certainly other relations in WRE that also make for coherence, like the evidential relation between moral beliefs and certain philosophical background theories. Thus, the question arises as to what the *general* notion of coherence in WRE amounts to. In discussing this question, Tersman (1993, p. 36) distinguishes between two main approaches. Namely, coherence can be understood (a) as a property of sets of beliefs and (b) as a relation between a particular belief and a set of beliefs. Among proponents of the first approach, it is undisputed that a necessary (though not sufficient) condition for the coherence of a set is its logical consistency. This minimal requirement is enough for my argument, so I will not bother with attempts to spell out the sufficient conditions for coherence in terms of (a). As for approach (b), I will simply take up Tersman's theory (though I do not think my argument depends on the particular details of his view). According to Tersman

²⁷ See Maier (2014, pp. 261–265) for a detailed argument against van Inwagen's analogy.

(1993, pp. 36–44), the coherence relation between a belief p and a set of beliefs S is essentially provided by justificatory arguments contained in S which support p . On this view, it is not so easy to determine the exact degree of coherence as it depends on many different considerations like the number and strength of the arguments for and against p contained in S , and the number and strength of the arguments for and against the premises of each of the arguments for p , and so forth. However, the general idea should be reasonably clear.

Now, in discussing both options, Tersman argues that in WRE one should opt *exclusively* for approach (b) because (a) leads to an implausibly strong requirement. This is because, on approach (a), it seems that unless all beliefs are justified, no one is (1993, p. 36). A related concern is that according to (a), remote inconsistencies in a person's belief system that are not at all relevant to a particular belief p implausibly prevent p from being justified. However, I contend that, at least if (a) pertains only to a system's logical consistency, one can endorse a combination of (a) and (b) while not falling prey to Tersman's objection. This works by claiming that *both* (a) and (b) are relevant to the notion of coherence, but that (b) takes epistemic priority over (a). In other words: if there is a high degree of coherence according to (b) between a particular belief p and a system S , then p is highly justified even if there are remote and irrelevant inconsistencies in S . Crucially, however, if the degree of coherence between p and S is somewhat unspecified according to (b)—which may be the case when there are no decisive justificatory arguments for either p or not- p , or when there is a stalemate between justificatory arguments for both p and not- p —then there is still a reasonable, albeit “secondary” or “weak” epistemic interest to improve the overall consistency in S . In this case, then, if p provides a higher degree of consistency in S than not- p , then p enjoys a higher degree of justification than not- p . Consequently, the view of coherence defended here can both accommodate Tersman's objection and preserve the plausible assumption that the overall consistency of a person's system of beliefs is of *some* epistemic value.

Let me now apply this view of coherence to the interplay between moral and empirical beliefs within WRE and, by way of example, to van Inwagen's argument for the falsity of determinism. Assume for the sake of argument that we follow van Inwagen in thinking that there are good reasons to accept the propositions that (1) people are often morally responsible for what they do and (2) moral responsibility is incompatible with determinism. Suppose further that determinism and indeterminism score equally well in terms of our best empirical theories about the world and that both options provide a similar degree of overall consistency among our *nonmoral* (especially empirical) beliefs. Now, the question is whether one must take an agnostic position on whether determinism is true or false, or whether there are any epistemic grounds for preferring either alternative. In WRE, this depends on whether any of the two alternatives coheres better with our total system of beliefs. The above criterion (b) leaves this unspecified because *ex hypothesi* there is no decisive *empirical* argument for either alternative, nor, as shown above, is there a valid justificatory argument from *moral* premises to the truth or falsity of determinism. Hence, on my view, the secondary criterion (a) comes into play, according to which an epistemic agent has reason to prefer a belief that guarantees a higher degree of consistency within her total system of beliefs. But this means that, under the above assumptions (and other

things being equal), there is reason to prefer a belief in the *falsity* of determinism, for, if determinism is true, (1) and (2) are not consistent, whereas if determinism is false, (1) and (2) are consistent.²⁸ Note, however, that as soon as our best empirical theories about the world provide us with plausible arguments to prefer one of the two alternatives to the other, criterion (b) would take epistemic priority over (a), so that there would be reason to prefer the belief which is evidentially supported by empirical arguments, even if it provides a lower degree of consistency among our ethical beliefs. Again, and somewhat less technically, the crucial point here is that justificatory arguments take epistemic priority over considerations of general consistency (including the consistency of our empirical beliefs with our moral beliefs). This means that the latter considerations only come into play when there are no decisive justificatory arguments for one of several competing empirical beliefs, or when there is a stalemate between justificatory arguments for several competing empirical beliefs.

So far, I have discussed the adjustment of empirical beliefs in light of certain moral convictions only in the context of van Inwagen's free will argument. But can the findings of this subsection be extended to the general relation between empirical and moral beliefs within RE? I think they can. Recall Rawls's view that both principled and non-principled moral judgments are crucially informed by empirical facts. Recall further that, as demonstrated in Sect. 2 above, the moral justification provided by Matching Relations 1, 2, and the empirical variant of Matching Relation 3 is necessarily based on empirical premises, which means that moral beliefs thus justified have some degree of necessary empirical entrenchment. Consequently, I suspect that in an RE framework, most (perhaps even all) justificatory arguments proceeding from moral premises to empirical conclusions are—just like van Inwagen's argument discussed above—in severe danger of vicious epistemic circularity. This is precisely because the moral premises of the argument are themselves (partially) justified on empirical grounds. Hence, assuming a RE-based framework, there are general reasons to be somewhat skeptical of justificatory arguments proceeding from moral premises to empirical conclusions.²⁹ This in turn means that the adjustment of empirical beliefs in light of moral beliefs will mostly work on the basis of my "sec-

²⁸ Two remarks are in order here. First, contrary to the above assumption, it may very well be that even in the absence of direct evidential relations, one option fits better (i.e. is more consistent) with the rest of our *empirical* beliefs (such as theories, observations, etc.). Then, naturally, this higher degree of empirical consistency can outweigh a lower degree of ethical consistency. While this consideration shows that determining the degree of overall consistency is probably a complex matter, it does not change my main point, namely that under criterion (a), ethical consistency can *in principle* affect the justification of empirical beliefs. Second, I should add that in the case of van Inwagen's argument, belief in the falsity of determinism might *not* lead to a higher degree of ethical consistency if there are *moral* reasons to deal with the empirical uncertainty in a specific way. It might be, for example, that belief in the falsity of determinism is used to justify a practice of blame and punishment. In this case, there are perhaps good moral reasons not to believe in the falsity of determinism if a certain threshold of empirical evidence is not met.

²⁹ Note, however, that my view of coherence does not *in principle* rule out such arguments. So, if there is a convincing argument from moral premises to empirical conclusions, then it provides for the kind of coherence defended in this article. I simply doubt that *in concreto* many or even any of these arguments are in fact convincing. I think it reasonable that my view of coherence does not preclude certain kinds of arguments from the outset, since such preclusion seems to depend on metaethical and/or metaphysical commitments that are external to the problem of coherence and should be defended separately.

ondary” coherence criterion (a). This conclusion dovetails nicely with a conjecture by Tersman (1993, p. 73):

Prima facie, it appears to me that straightforward empirical beliefs about e.g. the number of people in a room, the age of the earth, etc., are not in general candidates for being rejected in case of conflict with moral beliefs. Our empirical evidence for such beliefs is likely to be overriding. More plausible candidates are surely philosophical theories, underdetermined as they are by empirical evidence—e.g. theories of free will, the concept of alternative, moral discourse, deontic logic, decision theory, and [...] personal identity.

A view that combines the coherence criteria (a) and (b) and gives epistemic priority to (b) fits quite well into this picture. In the case of “straightforward empirical beliefs”, we will usually be able to specify an evidential relation on the basis of criterion (b), that is, through an empirical justificatory argument. The consistency with moral beliefs, which would be relevant for criterion (a), does not come into play in this case. However, in a case that is “underdetermined by empirical evidence”, criterion (a) may become relevant and thus provide us with an epistemic reason to prefer a belief that is consistent with our moral convictions to one that is not. If this is correct, then it follows that, in the application of WRE, moral convictions can provide us with some epistemic means of adjudicating between competing empirical descriptions of the world.

6 Conclusion and outlook

It is a commonplace that in WRE empirical beliefs are part of the relevant background theories. In this article, I have given a more precise account of the interrelations between moral and empirical beliefs within RE. In the first step, I showed that empirical beliefs are crucial to the process of matching moral principles to considered judgments. Metaphorically speaking, empirical beliefs constitute the playing field on which principles and judgments are matched. Rather than just being part of some background theory, then, empirical beliefs permeate the method at its core. Given this important role of empirical beliefs, it is no surprise that empirical developments are likely to bring about changes in the moral principles and theories we adopt in the pursuit of RE.

Noting these facts makes for a more precise description of RE, which may be interesting in its own right. However, my descriptive account of the role of empirical beliefs within RE also yielded two features that may prove important for evaluating RE as an approach to moral justification: first, the adaptability of the method to different empirical contexts, and second, the necessary empirical entrenchment of moral beliefs within RE. The discussion in this respect remained in some ways incomplete, as these two features will be of different epistemic value depending on broader philosophical commitments. My hope is that empirical adaptability and empirical entrenchment may at some point serve as pieces of the puzzle, so to speak, in a more comprehensive defense of RE.

Finally, assuming that coherence considerations can in principle affect the justification of empirical beliefs, I asked whether and how the latter might be adjusted in light of certain moral convictions. In answering this question, I first considered deductive arguments from moral premises to empirical conclusions and argued that one prominent example of such an argument, van Inwagen's argument for the existence of free will, does not succeed. Moreover, I suspected that in an RE framework, the empirical entrenchment of moral beliefs puts many (if not all) arguments from moral premises to empirical conclusions in danger of epistemic circularity. It remains to be seen, however, whether there are concrete examples in which such an argument is in fact successful.

I subsequently argued that even if deductive arguments from moral premises to empirical conclusions fail, moral beliefs may still provide us with epistemic reasons to prefer one empirical belief over another, at least if one accepts my proposed view of coherence. This view holds that both overall consistency and justificatory arguments matter for coherence, but not to the same extent, because the latter take epistemic priority over the former. My proposal may seem attractive because it does justice to the intuitive epistemic value of our belief system's consistency while at the same time avoiding a crucial objection to overall consistency as a criterion of coherence. Applying my new formulation of coherence, I argued that in a situation of empirical uncertainty characterized by an absence or stalemate of justificatory arguments, we may have reason to prefer an empirical belief that ensures a higher degree of *ethical* consistency (other things being equal). My abstract formulation of coherence, as well as the concrete conclusions drawn from it, can contribute to the broader discussion about coherentist theories of epistemic justification.

Acknowledgements I wish to thank Eran Fish, Federica Coppola and Antonia Gillhaus for valuable discussions and comments. I would also like to thank two anonymous reviewers for helpful comments.

Funding No funding was received to assist with the preparation of this manuscript. Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflicts of interest The author has no relevant financial or non-financial interests to disclose.

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

- Arras, J. D. (2009). The way we reason now: Reflective equilibrium in bioethics. In B. Steinbock (Ed.), *The Oxford handbook of bioethics* (pp. 46–71). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199562411.003.0003>
- Baumberger, C., & Brun, G. (2021). Reflective equilibrium and understanding. *Synthese*, 198, 7923–7947. <https://doi.org/10.1007/s11229-020-02556-9>.
- Birnbacher, D. (1999). Ethics and Social Science: Which Kind of Co-operation? *Ethical Theory and Moral Practice*, 2, 319–336. <https://doi.org/10.1023/A:1009903815157>.
- BonJour, L. (1985). *The structure of empirical knowledge*. Harvard University Press.
- Brink, D. O. (2021). *Fair opportunity and responsibility*. Oxford University Press.
- Brun, G. (2014). Reflective equilibrium without intuitions? *Ethical Theory and Moral Practice*, 17, 237–252. <https://doi.org/10.1007/s10677-013-9432-5>
- Brun, G. (2016). Explication as a method of conceptual re-engineering. *Erkenntnis*, 81, 1211–1241. <https://doi.org/10.1007/s10670-015-9791-5>
- Brun, G. (2022). Re-engineering contested concepts. A reflective-equilibrium approach. *Synthese*, 200, 168. <https://doi.org/10.1007/s11229-022-03556-7>
- Carnap, R. (1950). *Logical foundations of probability*. The University of Chicago Press.
- Chisholm, R. M. (1977). *Theory of knowledge* (2nd ed.). Prentice-Hall.
- Coffin, K. G. (2010). Double take: Evaluating double jeopardy reform. *Notre Dame Law Review*, 85(2), 771–808.
- Cohen, S. (1984). Justification and truth. *Philosophical Studies*, 46, 279–296.
- Daniels, N. (1979). Wide reflective equilibrium and theory acceptance in ethics. *The Journal of Philosophy*, 76(5), 256–282. <https://doi.org/10.2307/2025881>
- Elgin, C. Z. (2017). *True enough*. MIT Press. <https://doi.org/10.7551/mitpress/11118.001.0001>
- Fine, K. (2002). The Varieties of necessity. In T. S. Gendler, & J. Hawthorne (Eds.), *Conceivability and possibility* (pp. 253–281). Oxford University Press.
- Harman, G. (1986). *Change in view: Principles of reasoning*. MIT Press.
- Hörnle, T. (2022). Die subjektiven rechte der angehörigen von mordopfern—und ihre relevanz für die wiederaufnahme nach § 362 Nr. 5 StPO. *Goltdammer & Archiv für Strafrecht*, 169(4), 184–194.
- Kappel, K. (2006). The meta-justification of reflective equilibrium. *Ethical Theory and Moral Practice*, 9, 131–147. <https://doi.org/10.1007/s10677-005-9006-2>
- Kelly, T., & McGrath, S. (2010). Is reflective equilibrium enough? *Philosophical Perspectives*, 24, 325–359.
- McPherson, T., & Plunkett, D. (2021). Topic continuity in conceptual engineering and beyond. *Inquiry*. <https://doi.org/10.1080/0020174X.2021.1980095>.
- Maier, J. (2014). The argument from moral responsibility. *Australasian Journal of Philosophy*, 92(2), 249–267. <https://doi.org/10.1080/00048402.2013.808679>
- Olsson, E. (2023). Coherentist theories of epistemic justification. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy* (Winter 2023 Edition). <https://plato.stanford.edu/archives/win2023/entries/justep-coherence>
- Rawls, J. (1951). Outline of a decision Procedure for Ethics. *Philosophical Review*, 60(2), 177–197.
- Rawls, J. (1974). The independence of moral theory. *Proceedings and Addresses of the American Philosophical Association*, 48, 5–22.
- Rawls, J. (1999). *A theory of justice* (Revised Ed.). Harvard University Press.
- Rawls, J. (2001). *Justice as fairness: A restatement*. Ed. by Erin Kelly. Harvard University Press.
- Reutlinger, A., Schurz, G., Hüttemann, A., & Jaag, S. (2021). Ceteris Paribus Laws. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2021 Edition). <https://plato.stanford.edu/archives/fall2021/entries/ceteris-paribus>
- Scanlon, T. (2002). Rawls on Justification. In S. Freeman (Ed.), *The Cambridge Companion to Rawls* (Cambridge Companions to Philosophy, pp. 139–167). Cambridge University Press. <https://doi.org/10.1017/CCOL0521651670.004>.
- Slognsat, F. (2021). Ne bis in idem—Legitimität und verfassungsrechtliche Zulässigkeit einer Erweiterung der Wiederaufnahmegründe zuungunsten des Beschuldigten durch das Gesetz zur Herstellung materieller Gerechtigkeit. *Zeitschrift für die gesamte Strafrechtswissenschaft*, 133(3), 741–776. <https://doi.org/10.1515/zstw-2021-0028>

- Tersman, F. (1993). *Reflective equilibrium: An essay in moral epistemology (stockholm studies in philosophy 14)*. Almqvist & Wiksell International.
- Tersman, F. (2018). Recent work on reflective equilibrium and method in ethics. *Philosophy Compass*, 13, e12493. <https://doi.org/10.1111/phc3.12493>.
- van Inwagen, P. (1983). *An essay on free will*. Oxford University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.