*Journal of Ethics and Social Philosophy*
VOL. 14, NO. 2 · DECEMBER 2018
*https://doi.org/10.26556/jesp.v14i2.476*


# EVOLUTION AND MORAL DISAGREEMENT

*Michael Klenk*

THE BURGEONING DEBATE about the metaethical implications of the Darwinist view of morality focuses on which epistemic principle(s) allegedly support debunking arguments against moral objectivism.[1] Moral objectivism is the view that (at least some) moral truths are metaphysically necessary as well as constitutively and causally independent of human attitudes or beliefs.[2] Though objectivists must, of course, explain how objectivist moral beliefs can be justified in the first place, a central question is whether objectivist moral beliefs can be undercut, assuming that they are at least *prima facie* justified.[3]

So, what is that "something" in virtue of which a Darwinist view of morality creates a problem for objectivist moral beliefs? It has been claimed that evolutionary explanations of morality might show that moral beliefs are prone to error or fail to be modally secure, or that the best explanation of moral beliefs does not entail that they are (mostly) true.[4] None of these theses has found widespread support.

1 E.g., Clarke-Doane, "Justification and Explanation in Mathematics and Morality"; Klenk, "Survival of Defeat"; Sinclair, "Belief-Pills and the Possibility of Moral Epistemology"; Vavova, "Evolutionary Debunking of Moral Realism." The "Darwinist view of morality" is shorthand for "an evolutionary explanation of morality"; the view will be specified in section 2, below. Some debunkers take the argument to potentially undercut moral judgments that can be justified, e.g., Sinclair, "Belief-Pills and the Possibility of Moral Epistemology." If you accept this view, you can take the evolutionary challenge to have a wider scope, though I cannot address these further epistemological and metaphysical assumptions about morality in this paper.

2 E.g., Enoch, *Taking Morality Seriously*; Shafer-Landau, *Moral Realism*; Wielenberg, *Robust Ethics*. Objectivism is defended not only by robust realists but also by relaxed realists, such as Scanlon, *Being Realistic about Reasons*, and some naturalists, such as Jackson and Pettit, "Moral Functionalism and Moral Motivation." Arguably, they all face the evolutionary challenge; see Barkhausen, "Reductionist Moral Realism and the Contingency of Moral Evolution"; and Clarke-Doane, "Objectivity and Reliability."

3 See Wielenberg, *Robust Ethics*, and Klenk, review of *Robust Ethics*, for critical discussion.

4 Vavova, "Evolutionary Debunking of Moral Realism" and "Moral Disagreement and Moral

In light of this controversy, a new thesis is quickly gaining currency. A number of philosophers have argued that a Darwinist view of morality is metaethically significant because it shows that moral beliefs are counterfactually subject to disagreement.[5] This line of thought could also explain why the often-noted contingency of our moral beliefs is epistemically problematic.[6] Belief-contingency is problematic, on this view, *because* it reveals a problem with disagreement.[7] So, a Darwinist view of morality could yet play a metaethical role if it piggybacks on the epistemic significance of disagreement. For example, Mogensen writes that any metaethical implications that follow from a Darwinist view of morality "will be due to the epistemic significance of moral disagreement."[8] The disagreement in question is *hypothetical* or *counterfactual* disagreement: had our evolutionary history been different, our moral beliefs would conflict with our actual moral beliefs.[9] The consequence of this counterfactual moral disagreement is that the justification of all affected moral beliefs (objectively construed) is undercut, or so these philosophers argue. Let this be the disagreement view:

---

Skepticism"; Clarke-Doane, "Justification and Explanation in Mathematics and Morality"; Joyce, *The Evolution of Morality*; Lutz, "What Makes Evolution a Defeater?" See Klenk, "Survival of Defeat," ch. 2, for a discussion of the prospects of these arguments.

5 Bogardus, "Only All Naturalists Should Worry about Only One Evolutionary Debunking Argument"; Joyce, "Arguments from Moral Disagreement to Moral Skepticism"; Mogensen, "Contingency Anxiety and the Epistemology of Disagreement" and "Disagreements in Moral Intuition as Defeaters"; Sinclair, "Belief-Pills and the Possibility of Moral Epistemology"; Tersman, "Debunking and Disagreement"; White, "You Just Believe That Because…." Further suggestions about the relevance of disagreement for debunking, though not explicit endorsements, are provided by Street, "Objectivity and Truth," 314f; Clarke-Doane, "Justification and Explanation in Mathematics and Morality," 100; Ballantyne; "The Problem of Historical Variability," 246–54; and Horn, "Moral Realism, Fundamental Moral Disagreement, and Moral Reliability."

6 E.g., Handfield, "Genealogical Explanations of Chance and Morals"; Leiter, *The Future of Philosophy*; Lillehammer, "Methods of Ethics and the Descent of Man"; Sher, "But I Could Be Wrong."

7 Mogensen, "Contingency Anxiety and the Epistemology of Disagreement."; White, "You Just Believe That Because …."

8 Mogensen "Contingency Anxiety and the Epistemology of Disagreement," 591.

9 According to the ordinary understanding of disagreement, disagreement requires actual disputants and actual disputes. For example, one does not disagree about household chores if one's partner is merely lazy. Thus, on that understanding, whatever is implied by the evolutionary hypothesis seems far removed from disagreement. The relevant idea, however, is that some imaginary disagreements could easily be actual, in which case learning about them seems epistemically significant. I further specify the relevant type of disagreement in section 2.

> *Disagreement View*: Evolutionary explanations of morality imply that there is justification-defeating counterfactual disagreement about all moral beliefs (as conceived of by moral objectivists).

The disagreement view rests on two important claims about the epistemology of disagreement. First, the proper response to genuine peer disagreement is to be concessive: the disputants epistemically ought to reduce confidence in the disputed belief. Second, the same holds for *hypothetical* peer disagreement. The concessive view is controversial, and so is the view that hypothetical disagreement is epistemically significant.[10] For this paper, however, I will assume that the concessive view is true. Another route to attacking the disagreement view would be to deny that hypothetical disagreement is epistemically significant. There are good reasons, however, not to place too much weight on the actual/possible distinction in arguments about disagreement.[11]

Instead, I aim to show that the disagreement view is false by focusing on the requirements of epistemic peerhood, a rather underexplored issue in recent epistemology and uncharted territory in relation to evolutionary debunking arguments in metaethics.[12] My strategy is to raise a dilemma for proponents of the disagreement view. They have to claim that evolution creates counterfactual moral disagreement in nearby or non-nearby scenarios. However, in non-actual nearby scenarios, there will not be disagreement about all moral beliefs. In remote scenarios, there will be disagreement, but not with peers. So, evolutionary explanations of morality do not reveal *epistemically significant disagreement* about all moral beliefs, and the disagreement view is false, or so I argue.

Moreover, if it is true that debunking arguments are epistemically significant if and only if they reveal epistemically significant disagreement, then the argument presented in this paper implies that evolutionary explanations of morality are epistemically insignificant.[13] Independently of that claim, this paper speaks to what we can and cannot learn about counterfactual moral disagreement from

10 On the concessive view see, e.g., Enoch, "Not Just a Truthometer." On epistemic significance, see Kelly, "The Epistemic Significance of Disagreement"; Tersman, "Moral Disagreement."

11 My main worry is that drawing the actual/possible distinction will depend on counterfactual analyses to explain when possible but absent disagreements are significant, and counterfactual analyses have a bad track record in philosophy.

12 King ("Disagreement") and Gelfert ("Who Is an Epistemic Peer?") are notable exceptions, though their arguments do not speak directly to the disagreement view.

13 On epistemically significant disagreement, cf. Bogardus, "Only All Naturalists Should Worry about Only One Evolutionary Debunking Argument," 656; White, "You Just Believe That Because . . . ." See Klenk, "Third Factor Explanations and Disagreement in Metaethics" and "Survival of Defeat," for extended discussions of the epistemic insignificance of evolutionary explanations of morality.

evolutionary considerations. These findings should be of interest to both moral objectivists and their critics. Section 1 clarifies the context and the metaethical significance of the disagreement view. Section 2 reconstructs the argument for the disagreement view in greater detail. Section 3 introduces my argument against the disagreement view and sections 4 and 5 defend the two horns of the dilemma of the disagreement view.

## 1. Counterfactual Disagreement and Evolutionary Defeat

Evolutionary explanations of morality maintain that the *capacity* for normative guidance or the *content* of at least some of our most fundamental moral beliefs is a product of human evolutionary history.[14] For example, bravery appears to be evolutionarily useful, and it is evaluated positively across a wide range of societies.[15] So it stands to reason that the positive (moral) evaluation of bravery has an evolutionary origin. Thus:

> *Evolutionary Hypothesis*: A significant proportion of human moral beliefs are the product of human evolutionary history.

For this paper, quite a few critical and controversial issues about the evolutionary hypothesis have to be swept under the rug. That is acceptable, however, because virtually all discussants in the metaethical debate accept two corollaries of the evolutionary hypothesis. First, the evolutionary determinants of our moral beliefs are *contingent*: had human evolutionary history been different, human moral beliefs would have been different. Second, objective moral truths were *causally* irrelevant in the evolutionary genesis of our moral beliefs.[16] The evolutionary hypothesis, with its two corollaries, provides the basis for so-called evolutionary debunking arguments.[17] This paper follows the proponents of the disagreement view and focuses on those variants of evolutionary debunking arguments that aim to conclude that all objective moral beliefs are unjustified.

14 Joyce, *The Evolution of Morality*; Street, "Objectivity and Truth" and "A Darwinian Dilemma for Realist Theories of Value."

15 Curry, Mullins, and Whitehouse, "Is It Good to Cooperate?; Curry, "Morality as Cooperation."

16 Enoch, "The Epistemological Challenge to Metanormative Realism"; FitzPatrick, "Debunking Evolutionary Debunking of Ethical Realism"; Street, "A Darwinian Dilemma for Realist Theories of Value."

17 Though the precise role of empirical information in debunking arguments is controversial; cf. Clarke-Doane, "Justification and Explanation in Mathematics and Morality."

## 2. CLARIFYING THE DISAGREEMENT VIEW

This section reconstructs the argument for the disagreement view. I clarify each premise as we go along. My focus will be on Mogensen's and Bogardus's defenses of the disagreement view. Some steps differ in the details, but they ultimately reach the same conclusion.

First, they argue that the evolutionary hypothesis implies that our counterfactual selves might have had different moral beliefs from us:

> Had our species evolved elsewhere—as easily might have happened—and we later formed moral beliefs using the same method we actually used, our beliefs may easily have been incompatible with our actual beliefs, false by our own lights.[18]

There is reason to suppose that the moral intuitions of human beings reflect our place on the tree of life: had the conditions for the evolution of moral thought been realized in some distantly related species, their moral outlook would most likely incorporate certain fundamental differences in moral intuition, appropriate to their form of life.[19]

Both quotes reflect the idea that the evolutionary hypothesis implies the contingency of at least some of our moral beliefs. To make that idea more precise, let $M_{actual}$ be the set of moral propositions whose members are the objects of our moral beliefs, where "our" refers to US, the set of all human beings who live or lived in the actual world. Let $M_{counterfactual}$ be the set of moral propositions believed by THEM, where THEM stands for the set of all human beings who live in some fixed counterfactual evolutionary scenario. According to the evolutionary hypothesis, $M_{counterfactual}$ could diverge significantly from $M_{actual}$.

Second, proponents of the disagreement view claim that the divergence of $M_{actual}$ and $M_{counterfactual}$ amounts to hypothetical *disagreement* with our counterfactual selves. Compare Darwin's famous thought experiment:

> If men were reared under precisely the same conditions as hive-bees, there can hardly be a doubt that our unmarried females would, like the worker bees, think it a sacred duty to kill their brothers, and mothers would strive to kill their fertile daughters, and no one would think of interfering.[20]

18 Bogardus, "Only All Naturalists Should Worry about Only One Evolutionary Debunking Argument," 656.

19 Mogensen, "Contingency Anxiety and the Epistemology of Disagreement," 607.

20 Darwin, *The Descent of Man, and Selection in Relation to Sex*, 70.

Following Darwin's conjecture, the contents of $M_{counterfactual}$ might radically conflict with the members of $M_{actual}$.[21] Then we would find the moral beliefs of THEM "false by our own lights."[22] This constitutes hypothetical moral disagreement.

Third, proponents of the disagreement view claim that the hypothetical moral disagreement with our counterfactual selves is epistemically significant. We have the same type and quality of evidence for our moral beliefs, which, Bogardus claims, shows that there is an epistemically significant "evidential *symmetry*" between us and our counterfactual selves.[23] Mogensen, in contrast, takes the evolutionary hypothesis to show that we and our counterfactual selves have different evidence, which shows that there is an evidential *asymmetry* between us and our counterfactual selves. Such asymmetry is epistemically significant nonetheless because the moral disagreement implied by the evolutionary hypothesis bottoms out in pure conflicts of intuition.[24] Indeed, the claim is that hypothetical disagreement with our counterfactual selves is, all else being equal, as epistemically significant as actual disagreement. Bogardus qualifies this by saying that the hypothetical disagreement is "near enough to cause [epistemic] trouble" such that had we run a different evolutionary course, we would have *easily* ended up disagreeing with our counterfactual selves.[25] To paraphrase, hypothetical disagreement is relevant if it could be *easily present* (Bogardus) or if it is *arbitrarily absent* (Mogensen).

Fourth, the correct response to epistemically significant disagreement is to withhold judgment about the disputed belief. This claim is reminiscent of a *concessive* view about disagreement.[26] Concessive views imply that intractable disagreement about $p$ among interlocutors of comparable epistemic standing undercuts their justification for endorsing or rejecting $p$, provided that there is no *independent* evidence for or against $p$.[27]

In conclusion, the evolutionary hypothesis implies that there is justifica-

21 Mogensen, "Contingency Anxiety and the Epistemology of Disagreement," 607.

22 Bogardus, "Only All Naturalists Should Worry about Only One Evolutionary Debunking Argument," 656.

23 Bogardus, "Only All Naturalists Should Worry about Only One Evolutionary Debunking Argument," 638, emphasis added.

24 Mogensen, "Disagreements in Moral Intuition as Defeaters," 286ff.

25 Bogardus, "Only All Naturalists Should Worry about Only One Evolutionary Debunking Argument," 657.

26 Bogardus, "Only All Naturalists Should Worry about Only One Evolutionary Debunking Argument," 656; Mogensen, "Contingency Anxiety and the Epistemology of Disagreement," 607.

27 Elga, "Reflection and Disagreement"; Feldman, "Epistemological Puzzles about Disagreement"; Sidgwick, *The Methods of Ethics*. As noted in the introduction, I will assume for the

tion-defeating disagreement about $M_{actual}$. Given a concessive view about disagreement, and in the absence of independent evidence in favor of $M_{actual}$, we should give up our belief in $M_{actual}$.[28]

Before turning to my argument against the disagreement view, two clarifications are in order. First, Bogardus and Mogensen do not specify the nature of disagreement; I suggest understanding disagreement about $p$ as follows:[29]

> *Disagreement*: There is disagreement about $p$ iff there exists a $p$ such that
> (a) $S1$ believes that $p$ and $S2$ believes that $\sim p$, or $S1$ believes that $p$ and $S2$ suspends judgment on whether $p$, or $S1$ believes that $\sim p$ and $S2$ suspends judgment on whether $p$.
> (b) $S1$ and $S2$ have the same understanding of $p$.

Condition (a) is standard.[30] Condition (b) is sensible to preclude problems with merely apparent disagreement that turns out to be a sort of confusion of tongues.[31] Hence, when I write that there is disagreement about whether or not $p$, I mean that the disputants are referring to the same thing and that they are not merely talking past each other.

Second, the argument for the disagreement view is not explicitly presented as a case of peer disagreement. Nonetheless, it appeals to cases in which our counterfactual selves appear to be our epistemic peers in the minimal sense that their moral beliefs matter for the evaluation of our epistemic standing regarding morality. Bogardus emphasizes, as we have seen above, the evidential symmetry between us and our counterfactual selves. This affords the interpretation that he accepts what might be called a narrow conception of epistemic peerhood, which can be understood as follows:[32]

---

sake of argument that some such independence principle is valid and will not discuss it in what follows.

28 Of course, objectivism as a metaphysical thesis would still stand. Nonetheless, virtually every objectivist is in fact committed to the possibility of moral knowledge, and so the conclusion of the disagreement view would be a truly devastating result for their view; see Enoch, *Taking Morality Seriously*, 166; Shafer-Landau, *Moral Realism*.

29 I leave out complications about differences in credence regarding the disputed proposition between interlocutors. As far as I can see, nothing substantial depends on it in this paper.

30 Kölbel, "Faultless Disagreement," 54.

31 Tersman, *Moral Disagreement*, 22ff.

32 See Cohen, "A Defense of the (Almost) Equal Weight View," 98; Feldman and Warfield, *Disagreement*, 2; Kelly, "The Epistemic Significance of Disagreement," 174–75; King, "Disagreement," 252ff; Matheson, *The Epistemic Significance of Disagreement*; Wedgwood, "The Moral Evil Demons," 226.

> *Peerhood—Narrow Conception*: *S*1 and *S*2 are epistemic peers in regard to *p* iff *S*1 and *S*2 are equals regarding their evidential possession and their evidential processing with respect to *p*.[33]

Mogensen, in contrast, does not think that we and our counterfactual selves share equal moral evidence. Instead, he believes that we should treat the moral intuitions of our counterfactual selves as equally likely to be an excellent guide to the truth.[34] This sits very well with what might be called a broad conception of epistemic peerhood, which can be understood as follows:[35]

> *Peerhood—Broad Conception*: *S*1 and *S*2 are epistemic peers in regard to *p* iff *S*1 and *S*2 are equally likely to be right about *p*.

Neither specification of peerhood is entirely satisfactory *as a specification of peerhood*. For example, even on a narrow conception, a full specification of peerhood would doubtlessly require further conditions, such as "similar openness to experience."[36] In the present context, however, my concern is not so much with an entirely accurate specification of the concept of peerhood, but instead with the fixation of our ideas about which interlocutors the proponents of the disagreement view consider to be epistemically relevant. As such, less strict criteria for epistemic peerhood benefit the proponents of the disagreement view, since it would be easier for them to show that there is peer disagreement *on either such conception*.[37]

With these clarifications in place, it is evident that the truth of the disagreement view depends on whether or not the evolutionary hypothesis implies either narrow or broad peer disagreement (or both).

## 3. The Argument against the Disagreement View

So much for the argument in favor of the disagreement view; it is time to introduce my argument against it, which challenges the claim that the hypothetical

33 We can distinguish between acknowledged peer disagreement and non-acknowledged peer disagreement; see Kelly, "The Epistemic Significance of Disagreement," 168; King, "Disagreement," 261. In line with an internalist account of defeat, awareness of the disagreement is required to have an effect on justification. In the definition of peerhood, however, we can leave out this criterion.

34 Mogensen, "Disagreements in Moral Intuition as Defeaters," 294ff. Cf. Wedgwood, "The Moral Evil Demons," 241f.

35 E.g., Elga, "Reflection and Disagreement"; Vavova, "Moral Disagreement and Moral Skepticism," 308.

36 Cf. Gelfert, "Who Is an Epistemic Peer?"; King, "Disagreement."

37 See Gelfert, "Who Is an Epistemic Peer?" for problems with these accounts.

moral disagreement implied by the evolutionary hypothesis is epistemically significant (thus, proponents in the disagreement view go wrong in the third step of their argument):

P1. Hypothetical disagreement about p is epistemically significant only if the disputants are epistemic peers in regard to p, according to either a narrow or a broad conception of peerhood.[38]
P2. In non-actual nearby scenarios, the evolutionary hypothesis does not imply hypothetical disagreement about objectivist moral beliefs.
P3. In non-nearby scenarios, the evolutionary hypothesis does not imply hypothetical disagreement with epistemic peers, in the narrow or broad sense, about objectivist moral beliefs.
P4. So, the evolutionary hypothesis does not imply epistemically significant disagreement in either nearby or remote scenarios.
C. So, the evolutionary hypothesis does not imply epistemically significant disagreement.

The argument is deductively valid. P1 specifies two necessary conditions for the epistemic significance of disagreement. P2 and P3 state that neither non-actual nearby nor non-actual non-nearby scenarios exhibit disagreement about the objectivist moral beliefs (P2) or with our peers (P3); that means that there is no epistemically significant disagreement in either case. "Nearness" is a notoriously vague notion. I do not expect to offer a fully satisfactory account of it in this paper. For present purposes, nearby scenarios are those in which our counterfactual selves resemble the members of human societies on the ethnographic record (incidentally, this also implies closeness in time).[39] Non-nearby scenarios are those that depart in more or less extreme ways from the known ethnographic record. P4 is an intermediary conclusion that I will not discuss further. The argument's main conclusion implies that the disagreement view fails.

## 4. FIRST HORN OF THE DILEMMA: NO RELEVANT DISAGREEMENT IN NEARBY SCENARIOS

I will first precisify P1 by showing why peerhood is an important criterion for assessing the epistemic significance of hypothetical disagreement. Then I will consider whether the evolutionary hypothesis implies hypothetical disagreement about the truth of the relevant moral beliefs to vindicate P2.

38 To wit, relevant for the disputant's epistemic justification for or against endorsing the belief that *p*.

39 See Curry, "Morality as Cooperation."

*4.1. Relevant Interlocutors*

P1 states, as a necessary restriction on the epistemic significance of hypothetical disagreement in general, that the hypothetical disagreement must be between epistemic peers.[40] The narrow conception of peerhood disjunctively connects with the broad conception, such that two thinkers are peers if they are equals regarding evidential possession or equally likely to get it right (or both). The need to limit the epistemic relevance of possible disagreement through a peerhood requirement is suggested by the potentially devastating effects of combining an uncurbed epistemic relevance of possible disagreement with a concessive view, as suggested in the following example.

Suppose that experts *E*1 and *E*2 are, before their encounter, defeasibly justified in believing *p* and ~*p*, respectively. According to a simplistic version of the concessive view, *E*1 and *E*2 lose their justification for maintaining either belief once they learn of their disagreement. To maintain their belief, they have to appeal to independent evidence for or against p, or find independent evidence that suggests that their interlocutor is not their epistemic peer, to settle the question of whether or not *p* is true. Brushing aside thorny issues about the relevant sense of independence here, suppose that *E*1 and *E*2 do find independent evidence, *q*, about whether or not *p*. Usually, that would settle the disagreement. However, with the suggestion about the relevance of possible disagreement on the table, *E*1 and *E*2 cannot yet stop thinking about whether or not *p*, because it might be that *E*3's belief that ~*q* could either be *easily present* or *arbitrarily absent.* In the absence of a reason to think that *E*3's disagreement is too modally distant, *E*1, *E*2, and *E*3 would, being diligent adherents of the concessive view, have to consider whether there is independent evidence about whether or not *q* or about E3's epistemic status (while *E*1 and *E*2 remain agnostic about whether or not *p*), *ad infinitum*. So, on the face of it, a concessive view about disagreement paired with a view about the epistemic significance of possible disagreement leads to a vicious regress that leaves us unjustified in holding any belief at all.

So, lest general skepticism be embraced, the epistemic relevance of hypothetical disagreement must somehow be curbed. Peerhood among the interlocutors is a natural suggestion as a criterion for the epistemic relevance of a given disagreement. More pertinently, in the case of *possible* disagreement, there are countless hypothetical interlocutors, $E_n$, which might be relevant to the dispute existing between any two disputants $E_1$ and $E_2$. Limiting the set of relevant (hypothetical) interlocutors to those who are in equal evidential possession *or* antecedently

40 Naturally, proponents of the disagreement view should consider only disagreements that are plausibly implied by the evolutionary hypothesis.

equally likely to get things right would at least partly impede the potential regress that is made possible by invoking hypothetical disagreements. Hence, for hypothetical disagreement to be epistemically significant, it must be among peers.

*4.2. Peer Agreement about the Relevant Moral Beliefs*

In this section, I will narrow down the range of relevant beliefs that objectivists have to defend, which will, based on plausible evolutionary considerations, vindicate P2. According to proponents of the disagreement view, the evolutionary hypothesis must imply hypothetical disagreement about the contents of moral beliefs. However, objectivists need not defend *all* members of $M_{actual}$ against the evolutionary challenge, and hardly any objectivist aims to do so.[41] This is because $M_{actual}$ certainly does not contain only true and justified moral beliefs. It contains moral beliefs that reflect biases, conceptual errors, and other infelicities. It also contains highly specific beliefs that refer to idiosyncratic sociocultural factors, such as food taboos, which are often moralized.[42] Objectivists do not claim that all of these beliefs are justified.

Rather, defending the justification of some moral beliefs is enough to guard objectivism against the evolutionary challenge. In particular, objectivists defend the justification (and truth) of the following moral beliefs:

1. Survival and reproductive success … is at least somewhat good.[43]
2. Pleasure is usually good, and pain is usually bad.[44]
3. We have rights because we are reflective beings.[45]

These platitudes have a similar structure: some plausibly evolutionarily relevant natural property or event (e.g., being an instance of survival, being painful, being capable of self-reflection, etc.) is related to a moral property such as being good. The normative concept alluded to is always a thin moral concept: GOODNESS, BADNESS, or RIGHT (in the sense of "having a right," rather than being correct).

Let the set of moral platitudes be $M_{basic}$. $M_{basic}$ is a proper subset of $M_{actual}$. I do not attempt to outline the contents of $M_{actual}$. It suffices to distinguish $M_{basic}$ from $M_{actual}$. The members of $M_{basic}$ are the moral platitudes that are universally accepted (by actual humans). Moral platitudes have two components. First, they are picked out by *thin* moral concepts. Thin moral concepts are evaluative concepts without descriptive content: RIGHT, GOOD, and OUGHT are standard

41 Cf. Shafer-Landau, *Moral Realism*, 17.

42 Cf. Henrich and Henrich, "The Evolution of Cultural Adaptations," 3717.

43 Enoch, "The Epistemological Challenge to Metanormative Realism," 430.

44 Skarsaune, "Darwin and Moral Realism," 232.

45 Wielenberg, "On the Evolutionary Debunking of Morality," 447.

examples. Second, moral platitudes latch onto the nonmoral facts that are evolutionarily relevant. The members of $M_{basic}$ are thus the beliefs that combine thin moral concepts with evolutionarily relevant causal factors, such as pain, procreation, and survival.[46] Judging by the ethnographic record, every society accepts $M_{basic}$. For example, Curry, Mullins, and Whitehouse found that every (studied) society has moral rules about problem-centered domains of resource allocation, coordination to mutual advantage, exchange, and conflict resolution.[47] They also provide further evidence that people across societies evaluate positively behavior that represents optimal solutions to domain problems (e.g., all appreciate food sharing and bravery). These findings are supported by an overarching hypothesis that suggests that some morals differ across societies, but that there is a fundament of basic moral beliefs, indeed those that seem to belong to $M_{basic}$, that are held constant and shared across societies.[48]

This characterization of the relevant domain in terms of $M_{basic}$ suggests that in *nonactual nearby scenarios*, we have good reason to suppose that the basic moral beliefs of our counterfactual selves will be like those of individuals in our society or other societies on the ethnographic record. Thus, given the ubiquity of beliefs in the platitudes of $M_{actual}$, it seems very likely that $M_{basic}$ is a proper subset of $M_{counterfactual}$ too. That is, all the basic moral beliefs, whose contents are in $M_{basic}$, will also be endorsed by our counterfactual selves in nonactual nearby scenarios.[49] Thus, turning back the wheel of life only a tiny bit will show that there is agreement rather than disagreement about $M_{basic}$.

We have arrived at an important intermediary conclusion: if we confine ourselves to nearby possible scenarios, then the evolutionary hypothesis implies agreement about some moral beliefs that can be explained evolutionarily, rather than disagreement. While the evolutionary hypothesis might suggest disagreement about *some* moral beliefs in nearby possible scenarios, these disagreements are merely disagreements about the application of thick moral concepts, rather than disagreements about the members of $M_{basic}$.[50] As such, moral objectivists need not worry about these kinds of disagreement; P2 is thereby vindicated.

46 I take no stance on whether or not the members of $M_{basic}$ stand in deductive or inferential relations to each other.

47 Curry, Mullins, and Whitehouse, "Morality as Cooperation."

48 Cf. Morris, *Foragers, Farmers, and Fossil Fuels*.

49 They need not be explicitly endorsed, as explicit representation is not necessarily required for believing something; cf. Harman, *Change in View*, 13ff. Thanks to an anonymous referee for prompting me to clarify this point.

50 Barkhausen, "Reductionist Moral Realism and the Contingency of Moral Evolution"; Morris, *Foragers, Farmers, and Fossil Fuels*.

However, proponents of the direct approach will probably be unimpressed by the lack of disagreement in nearby scenarios. They might argue that considering only nearby scenarios betrays a lack of imagination. Recall Darwin's thought experiment about the bees, which is supposed to illustrate that "we" could have ended up being very different organisms after all. In that case, it seems that the intersection between $M_{actual}$ and $M_{counterfactual}$ will get smaller and smaller as we replay the tape of life until we arrive at counterfactual selves that do not agree about any member of $M_{actual}$ and thus, by extension, any member of $M_{basic}$. In other words, P3 might still be false: when our counterfactual selves are like Darwin's human bees, $M_{counterfactual}$ will diverge radically from $M_{basic}$. So, the proponents of the disagreement view might claim that the evolutionary hypothesis will reveal epistemically significant disagreement in non-nearby scenarios. Let us follow them there.

## 5. SECOND HORN OF THE DILEMMA: NO DISAGREEMENT WITH MORAL PEERS

In *non-nearby* scenarios, the evolutionary hypothesis would be very likely to yield some disagreement about $M_{basic}$ (that is, $M_{basic}$ would not be a subset of $M_{counterfactual}$). However, to vindicate P3, I aim to show that any disagreement we may find in non-nearby scenarios is not *peer* disagreement.

### *5.1. Disagreement between Peers on a Narrow Conception*

Let us consider *narrow* peer disagreement first. Recall that a narrow conception of peerhood says that two persons are peers if and only if they are in equal evidential possession and their processing of the evidence regarding moral issues is also equal. The crucial question is thus the following: Does the evolutionary hypothesis imply that our counterfactual selves in non-actual, non-nearby scenarios have the same evidence as we do and yet disagree with us about $M_{basic}$?

In order to assess the crucial question, we first have to take a brief detour back to the evolutionary hypothesis (which said that a significant proportion of human moral beliefs are the product of human evolutionary history). Proponents of the disagreement view must adopt a stringent interpretation of the evolutionary hypothesis, according to which there is a close connection between the ancestral environment, evolutionary forces, and the contents of moral intuitions and moral beliefs.[51] On this view, evolutionary processes influence the raw material based on which we form our moral beliefs to such a degree that if you change the evolutionary environment, you change the raw material and thereby the moral

51 Environment is to be widely understood to encompass sociocultural factors.

beliefs that our counterfactual selves will hold.[52] Of course, the details of this story may differ depending on the correct view about the bases of moral beliefs, but the result will be the same for proponents of the disagreement view: a sufficiently different evolutionary trajectory changes *whatever* determines our moral beliefs, and so our counterfactual selves end up with different moral beliefs.[53]

Proponents of the disagreement view need the stringent interpretation of the evolutionary hypothesis to fend off an objectivist objection: objectivists might claim that there is a subset of $M_{basic}$ that is not subject to evolutionary contingency. That is, if moral beliefs are not determined by moral intuitions (and/or mental states with nonmoral content) that, in turn, are determined by evolutionary forces, then moral beliefs might reliably track moral facts after all.[54] For example, objectivists could argue that our beliefs about, say, the (*pro tanto*) badness of pain are not subject to evolutionary contingency and thus our counterfactual selves will *not* adopt diverging beliefs about the badness of pain.[55] The stringent interpretation will, in contrast, entail that our counterfactual selves will have different moral beliefs about the badness of pain *because* our counterfactual selves will have different moral intuitions and different moral intuitions *because* their sensory input is different on different evolutionary tracks.

Importantly, the stringent interpretation of the evolutionary hypothesis implies an inclusive notion of moral evidence (to wit, evidence for the moral truth). According to the inclusive notion of moral evidence, mental states with moral content (such as moral intuitions) and with nonmoral content (such as sensory input) can count as evidence for moral truth.[56] As we have seen, the stringent interpretation of the evolutionary hypothesis implies that both factors

52 Cf. Mogensen, "Contingency Anxiety and the Epistemology of Disagreement," 593.

53 For example, my point applies even when moral intuitions are themselves doxastic states (versus perceptual states) and even when moral beliefs are also based on mental states with nonmoral content (versus exclusively based on moral intuitions). For discussions of these views, see Climenhaga, "Intuitions Are Used as Evidence in Philosophy"; Chudnoff, "What Intuitions Are Like"; Bengson, "Grasping the Third Realm"; and Huemer, *Ethical Intuitionism*. Importantly, even if one thinks, like Bengson and Huemer, that moral beliefs are not based on intuitions as I understand them here, but instead on "direct perception," the stringent interpretation of the evolutionary hypothesis appears to be a threat to the form of moral objectivism they defend, too, as desired by proponents of the disagreement view.

54 FitzPatrick, "Debunking Evolutionary Debunking of Ethical Realism."

55 Objectivists might even use these beliefs, given their untarnished epistemic credentials in this scenario, to set up so-called third-factor accounts in order to vindicate other moral beliefs that are subject to evolutionary contingency; cf. Enoch, "The Epistemological Challenge to Metanormative Realism."

56 Strictly speaking, it is the fact that one has an intuition that may count as evidence, not the intuition, or seeming, itself.

will change on alternative evolutionary trajectories. Both factors are also usually considered to be relevant to the question of which moral beliefs it is rational for one to hold.[57] It is a controversial issue whether both factors or only moral intuitions count as evidence for moral beliefs, but the issue need not be settled here: as long as either factor counts as moral evidence, we will find different evidence on different evolutionary paths (given the stringent interpretation of the evolutionary hypothesis).[58] This means that, on a view of moral evidence charitable to proponents of the disagreement view, moral intuitions count as moral evidence.

I can now compare the input/output relations of our moral faculty with the input/output relations of our counterfactual selves.[59] The options are exhausted by four cases, where *Input* refers to the forces that shaped moral intuitions in THEM and US ($Input_{US}$ and $Input_{THEM}$, respectively) and *Output* refers to the set of moral beliefs (again with the subscript indicating whether they are OUR beliefs or THEIR beliefs):

1. $Input_{US} = Input_{THEM}$; $Output_{US} = Output_{THEM}$
2. $Input_{US} \neq Input_{THEM}$; $Output_{US} = Output_{THEM}$
3. $Input_{US} = Input_{THEM}$; $Output_{US} \neq Output_{THEM}$
4. $Input_{US} \neq Input_{THEM}$; $Output_{US} \neq Output_{THEM}$

Cases 1 and 2 signify agreement (since both outputs are identical) and are thus not relevant here. Cases 3 and 4 signify a divergence of $M_{actual}$ and $M_{counterfactual}$

57 Wedgwood, "The Moral Evil Demons," 226.

58 This understanding of "evidence" is unorthodox insofar as it does not signify an epistemic support relation: not every determinant of a moral belief is also an epistemically *good* reason for that belief (for some subject); see Huemer, "The Problem of Defeasible Justification," 376. But such an inclusive notion is required for dialectical reasons. Adopting a restrictive interpretation would be uncharitable for proponents of the disagreement view. For example, the view of Williamson (*Knowledge and Its Limits*) would be restrictive in the present context because, assuming that knowledge requires truth, the evolutionary hypothesis could not, per se, imply peer disagreement about objective moral facts, narrowly construed. Either us or them would have evidence, but not both, and thus there would not be peer disagreement between THEM and US. Of course, this would be one way to argue against the disagreement view, though one that I do not pursue here mainly because an adequate discussion of a theory of evidence is beyond the scope of this paper. Thus, the more inclusive the notion of evidence, the more likely it is that the evolutionary hypothesis implies that there is disagreement with peers that share the same evidence (thus fulfilling the criteria for an epistemically relevant disagreement).

59 I do not place too much weight on the term "information" here. I wish mainly to exploit the thought of an input/output process whose relation between input and output is systematic, since this is what proponents of the evolutionary hypothesis claim; see Harms, *Information and Meaning in Evolutionary Processes*, for a relevant discussion.

and could be relevant for assessing the disagreement view. However, case 3 is not implied by the evolutionary hypothesis and case 4 is not relevant disagreement.

Consider case 3 first. It does indeed appear to be a relevant peer disagreement. Our counterfactual selves disagree about some beliefs in $M_{basic}$, and given that these beliefs are based on the same input, the narrow conception of peerhood tells us that we have a peer disagreement. However, either case 3 is *not* implied by the evolutionary hypothesis, or our counterfactual selves mean different things when they use moral terms. Consider the first option. Case 3 illustrates that subjects that base their moral beliefs on the very same input will generate differing moral beliefs. In other words, the output is not correlated with the input—a sign of a random process. However, the evolutionary hypothesis does not imply that our moral beliefs are the products of a random process. Indeed, as suggested above, the interpretation of the evolutionary hypothesis that proponents of the disagreement view require implies that moral beliefs are based on moral intuitions to such an extent that changing the moral intuitions would change the organism's moral beliefs.

To illustrate the first interpretation of case 3, suppose that our counterfactual selves live in a world exactly like ours in all nonmoral aspects. Given that they form their moral beliefs in the same way as we do, by relying on their intuition, there is no indication that their intuitions are any different in a world that is just the same as our world. The point is that disagreement is only a problem insofar as differences in output cannot be traced to differences in input.[60] Therefore, case 3 does not follow from the evolutionary hypothesis and thus it does not help the proponents of the direct approach. Of course, the assumption that moral intuitions shaped by evolutionary forces determine the content of moral beliefs is a stark idealization. If evolutionary forces fully determine the bases of moral beliefs, *pace* the evolutionary hypothesis, then organisms subject to the same evolutionary history might have different moral intuitions and correspondingly different moral beliefs. Note, however, that this line of reasoning is no help for proponents of the disagreement view. Pursuing the same thought about the disconnect between evolutionary influences and moral beliefs, objectivists can argue that truth-conducive methods such as reasoning or understanding can lead to true beliefs based on intuitions that are not influenced by evolutionary forces.[61]

Alternatively, when the outputs of our counterfactual selves differ from ours, even though they are based on the very same inputs, we have reason to suspect

60 Wright, *Truth and Objectivity*, 91ff.

61 FitzPatrick, "Debunking Evolutionary Debunking of Ethical Realism"; Huemer, "A Liberal Realist Answer to Debunking Skeptics."

that they are using moral terms differently from us. After all, based on the stringent interpretation of the evolutionary hypothesis and the fact that they had the same input, we should expect our counterfactual selves to agree with us about moral matters. The best explanation of their ostensible disagreement would be that we are not really disagreeing, but merely talking past each other.[62] Consequently, we would not have a genuine disagreement at all (irrespective of whether it is between peers) on this interpretation of case 3.

Case 4 also shows a disagreement, and the evolutionary hypothesis plausibly implies it. Suppose, for example, that our counterfactual selves live in a world where their overall fitness is increased by sacrificing their children. They might indeed be rather like Darwin's bees. Due to various evolutionary processes, they might have different intuitions about how to treat their children from those we have, and consequently, they will form moral beliefs whose contents conflict with some of the members of $M_{basic}$. Thus, we certainly have a relevant disagreement in case 4.

However, case 4 does not exhibit peer disagreement, narrowly construed, because US and THEM do *not* have the same evidence. Our counterfactual selves had different sensory inputs, and thus they have different moral intuitions: when they consider, say, whether to sacrifice their children, they might feel a warm glow of anticipation and a resounding positive feeling toward the thought—quite unlike our moral intuitions about infanticide. Therefore, US and our counterfactual selves will not be in equal evidential possession: case 4 is not peer disagreement, narrowly construed.

This line of argument, which says that difference in intuitions prevents people from counting as peers, might imply that paradigmatic cases of disagreement do not count as peer disagreement either.[63] For example, two expert mathematicians who disagree about the truth of Goldbach's conjecture, based on diverging mathematical intuitions, would not be having a peer disagreement. However, insofar as this is a problem, it is only a problem for the narrow conception of peerhood or for the view that intuitions constitute the sole evidence relevant for beliefs about Goldbach's conjecture (or moral beliefs, in the analogous case).

62 This resembles a point made by Davidson (*Inquiries into Truth and Interpretation*) about radical interpretation. Davidson argues, roughly, that in cases of radical disagreement about a subject matter, the "principle of charity" demands that we should regard the other party as talking about a different subject altogether. Since we are concerned with merely conceivable disagreement, I suppose we can conceive that there is *no* talking at cross purposes going on and so we need not be charitable. Still, in agreement with Davidson, I believe that we should not take seriously the disagreement in this case.

63 I am grateful to an anonymous referee for raising this objection and for suggesting a possible solution.

However, the validity of the narrow conception is not at issue here (and section 6 aims to show that my argument holds if we adopt the alternative broad conception of peerhood). Moreover, case 4 would still not be a case of peer disagreement if we counted mental states with nonmoral content as moral evidence, too. Mental states with nonmoral content would *also* be different for our counterfactual selves, since this difference explains why they have different moral beliefs, and so we would end up with different evidence again. For example, if our counterfactual selves had different factual beliefs about the effects of, say, fratricide, their moral beliefs about fratricide would plausibly differ from ours, but then again, we would have different evidence and, again, case 4 would not be a case of narrowly construed peer disagreement.

Therefore, on a narrow conception of peerhood that is congenial to the proponents of the disagreement view, the evolutionary hypothesis does not imply relevant peer disagreement in non-nearby scenarios, which partly vindicates P3. Granted, however, the narrow conception of peerhood is not, though congenial to Bogardus's view, the most felicitous conception of peerhood for proponents of the disagreement view. Their argument could yet be saved if there were peer disagreement on the broad conception of peerhood. In the next section, we stay in non-nearby scenarios but consider whether any of our counterfactual selves are peers on a broad conception of peerhood.

### *5.2. Total Disagreement between Peers on a Broad Conception*

On a broad conception of peerhood, you will recall, our counterfactual selves count as our peers insofar as they are equally likely to be right about moral matters. Since relevant disagreement is about $M_{basic}$, we can distinguish between two relevant cases for analysis: total disagreement and partial disagreement about $M_{basic}$. Neither case, however, creates a problem for moral objectivism.

Consider total disagreement about $M_{basic}$ first. Total disagreement means that our counterfactual selves disagree about *all* beliefs in $M_{basic}$, which is a tremendously extreme situation.[64] Our counterfactual selves would not even agree that pain is *pro tanto* bad, that survival is good, that people are fundamentally equal,

64 This does not exclude the fact that our counterfactual selves agree about moral logical truths, such as "survival is either good or not good." But tautologies are either part of $M_{basic}$ or they are not. If they are not, then the agreement is irrelevant for my argument. If they are, then we and our counterfactual selves still cannot validly infer points of agreement that lie within $M_{basic}$ that are not tautologies, which suffices for my argument. My view is that platitudes based on thin moral concepts, but not tautologies, are part of $M_{basic}$. The reason is that the latter, but not the former, are evaluative and action guiding and thus useful to have in an evolutionary sense. Thanks to an anonymous referee for prompting me to clarify this point.

or that we should not cause unnecessary harm.[65] This is noteworthy because *at least some* agreement about some subject matter provides one with *prima facie* reasons to accept someone as a peer. For example, consider whether our counterfactual selves would be our peers if they were like the Neanderthals. Suppose we know nothing about our phylogenetic relatedness, their social habits, or their ventures into early forms of art. Despite ignorance on these matters, a good reason (not necessarily a sufficient reason) to maintain "default trust" in the moral intuitions of Neanderthals would be *some* agreement about $M_{basic}$.[66] Indeed, Neanderthals plausibly endorsed some of the members of $M_{basic}$: they believed that it is good to take care of one's family and community, they cherished survival, and they generally avoided pain.[67] Such agreement about $M_{basic}$ would constitute common ground to accept Neanderthals as our moral peers, despite the 30,000 years that separate us from them. Would we have reason to withdraw the default trust we bestowed upon Neanderthals *in the absence of* agreement about $M_{basic}$? To aid our imaginations, let us conjure up some evolutionary path on which our counterfactual selves do not endorse any member of $M_{basic}$ and let us call that species *Homo sapiens peregrinus*, the strange man.[68]

My claim is that total disagreement about $M_{basic}$ with our Peregrinus-like counterfactual selves gives us reason to reject their default peerhood status (which they do possess, according to proponents of the disagreement view) and so we would not have peer disagreement on a broad conception of peerhood either. I will support my claim by looking at the reasons for granting and withdrawing default trust, which proponents of the disagreement view themselves accept.

65 When extreme cases of disagreement, in which disputants share little common ground, are considered in the literature on peer disagreement, even defenders of concessive views concede that disagreement might then lose its epistemic significance; cf. Elga, "Reflection and Disagreement," 495f; Kornblith, "Belief in the Face of Controversy," 50. Most discussions, however, focus on cases in which there is at least *some* agreement, as in disagreements with psychopaths, who disagree about many but not all moral beliefs; cf. Ballantyne, "The Problem of Historical Variability," 254; Sinnott-Armstrong, "Moral Disagreements with Psychopaths," 53. In such cases, Tersman's worry that "the mere fact that a person disagrees with us … cannot itself count as a shortcoming" seems apt (*Moral Disagreement*, 34ff). However, the extremity of rejecting $M_{basic}$ altogether implies that reasons to take even "moral monsters" as our peers do not straightforwardly apply. While psychopaths, for example, recognize the method we use to arrive at moral beliefs and what constitutes good moral reasoning, they are simply unperturbed by it. Total disagreement with our counterfactual peers, in contrast, does not even allow agreement about what good methods of moral reasoning are.

66 Cf. Mogensen, "Disagreements in Moral Intuition as Defeaters," 283.

67 Cf. Wynn and Coolidge, *How to Think Like a Neandertal*, 19–21.

68 It is doubtful that our Peregrinus selves would be plausible products of an evolutionary process. I will return to this biological objection against the disagreement view in section 5.3.

According to proponents of the disagreement view, our counterfactual selves enjoy default trust because of a fundamental symmetry between THEM and US. Insofar as we can (defeasibly) trust our own moral beliefs without independent vindication, we must extend this trust to others.[69] Let us accept this controversial claim for the sake of argument.[70] Second, proponents of the disagreement view claim that default trust is defeasible if considerations *independent* of the disputed proposition let us assign a higher likelihood of us being right about the disputed proposition compared with the disputing interlocutor.[71] For example, suppose that you judge that torture is morally impermissible, but you learn that within a week you will judge that torture is morally permissible. In light of this disagreement with your future self, is your current belief about the impermissibility of torture defeated?[72] No, because you can justifiably reject your future self's peer status based on the privileged access you have to your reasons for thinking that torture is impermissible.[73] Hence, default trust is defeasible if we have reason to think that we are more likely to get things right compared to our peers. For example, given a disagreement, we might know we were, say, not intoxicated when thinking about the disputed issue, but lack such knowledge about our peers. Such cases, where we have *no* or *little* information about our peers, but a lot of (introspective) information about ourselves, are highly relevant for defeating default trust.

It follows from this view of defeating default trust that if we *only* know that our counterfactual selves totally disagree about the truth of the contents of $M_{basic}$, then our counterfactual selves will lose their status as peers. The crucial question is therefore whether we know *anything* about our Peregrinus-like counterfactual selves, apart from the fact that we disagree, that gives us reason to maintain our trust (our trust, that is, that our counterfactual selves are as likely as us to get moral matters right). In other words, given that (a) we totally disagree about $M_{basic}$ and (b) we know $X$ about our counterfactual selves, are we equally likely to get moral matters right?[74] If nothing can replace $X$, in combination with the fact that we totally disagree about $M_{basic}$, suggests that they and we are equally

69 Cf. Foley, *Intellectual Trust in Oneself and Others*; Mogensen, "Disagreements in Moral Intuition as Defeaters."

70 See Enoch, "Not Just a Truthometer," 962ff, for critical discussion.

71 Mogensen, "Disagreements in Moral Intuition as Defeaters," 294ff.

72 Wedgwood, "The Moral Evil Demons," 241.

73 Mogensen, "Disagreements in Moral Intuition as Defeaters," 294–95.

74 Again, this is in light of the assumptions of the evolutionary challenge: our moral beliefs are *prima facie* justified, so the methods we are currently using to arrive at $M_{basic}$ are by and large accurate.

likely to get moral matters right, then our Peregrinus-like counterfactual selves do not count as our peers.

My aim is the modest one of showing that two plausible factors for maintaining trust in our counterfactual selves, two factors to "fill in" *X*, turn out to be mistaken. If that is right, then we have strong reason to judge that cases of total disagreement about $M_{basic}$ are not cases of peer disagreement, which further vindicates P3.[75] We could look, first, to Peregrinus's *cognitive capacities including nonmoral beliefs*, or, second, to their *belief-forming methods*. Neither aspect, however, is convincing. Consider cognitive capacities first.

> $X_1$ = *Cognitive capacities including nonmoral beliefs.*[76] Our peers are those who, in general, reason as well as us. They are as good as we are in obtaining scientific knowledge. They compose logical proofs, understand physics, and perform as well in standardized intelligence tests as average humans.

Cognitive development might be a necessary condition for peerhood in moral matters, but it is not a sufficient condition. Cognitive development of a certain level might function as a kind of "enabler" for making correct moral judgments. For example, if Peregrinus lacked a theory of mind, similar to very young children, he would be prone to making egoistic decisions and would lack the ability to recognize that other beings have their own plans and wishes.[77] It might thus be thought that cognitive abilities alone provide reason for expecting Peregrinus to be a good moral reasoner. However, the opposite is true. Just because a specimen of Peregrinus can realize that you would be hurt by something he does, this does not imply that he will respect that consideration.

Moreover, people can be experts in one area but still be (systematically) wrong in another area, and it is generally the case that assessments of peerhood seem domain specific.[78] Otherwise, it would make good sense to ask expert chess players to sit on ethics committees and top-notch nuclear physicists to weigh in on Europe's border policies just *because* their cognitive abilities are taken as evidence of their ethical expertise. More generally, an ability to get moral matters right does not seem to be directly inferable from competence in nonmoral mat-

75 I say that my goal is modest because I do not provide a conclusive case; there might be other reasons for maintaining trust in our counterfactual selves that I cannot (for lack of space) address here. However, I cannot think of more, nor have proponents of the disagreement view provided such reasons yet.

76 Capacities are understood here as having the ability to function on a certain level.

77 Kohlberg and Hersh, "Moral Development."

78 Goldberg, "Disagreement, Defeat, and Assertion," 169; Weatherson, "Disagreements, Philosophical and Otherwise," 56.

ters, otherwise we could argue, for example, that our ability to ascertain mathematical truth indicates that we are good at grasping moral truth. These proposals do not look promising, and thus the mere fact that we have shared cognitive capacities with Peregrinus (in the absence of any agreement about $M_{basic}$) gives us no reason to think that they are as likely as us to get moral matters right.[79]

Consider the following alternative instead:

> $X_2$ = *Similarity of belief-forming methods.* Our peers are those who use similar methods of belief formation. Peregrinus is as good as we are at obtaining knowledge about nonnormative matters. Peregrinus also relies on his intuition in forming moral beliefs, and so do we.

The proposal might seem promising, but it just pushes back the crucial question.[80] We wanted to know whether we have reason to believe that Peregrinus is as likely as us to get moral matters right (given total disagreement about $M_{basic}$), but now we need to know whether we are using the same method of belief formation. However, how can we know what method Peregrinus is using, when the outputs of whatever method he is using are completely different from the outputs of our method?

The problem is one of individuating methods, and there is a metaphysical and an epistemic dimension to it. Metaphysically, the question concerns which method is being used; epistemically, the question concerns which method we have sufficient reason to believe is being used. Consider the metaphysical question first. We can suppose that methods are partly individuated by input, given

79 Cf. Klenk, "Can Moral Realists Deflect Defeat Due to Evolutionary Explanations of Morality?" for an objection to the view that we can use our reliability in a nonmoral domain to vindicate our moral reliability.

80 There is reason to doubt that similarity of belief-forming methods is a good criterion to use to judge whether others are likely to get things right in the first place, though I will not, for dialectical reasons, build on this argument here. The reason is that the proposal presupposes a sound approach for individuating methods (effectively an answer to the generality problem) and no current approach has found widespread support; cf. Bishop, "Why the Generality Problem Is Everybody's Problem," 285. Any description of a method, M, that we are using (e.g., perception, statistical inference) might also fit the description of the methods used by a class of interlocutors that we do not normally regard as peers, such as hallucinating people (in the perception case) or depressed economists (when it comes to making accurate predictions about the economy). According to the present proposal, we would have to regard them as peers *or* find a better principled way to describe the method. But given the problems with individuating methods, this seems unpromising, so we would have to, counterintuitively, accept that they are peers. This worry should be kept in mind as an additional problem for the proposal that similarity of belief-forming methods is a good criterion for peerhood, though I aim to show that the proposal does not help the disagreement view even if it is *prima facie* acceptable.

that individuating methods externally is a widely accepted approach to method individuation among epistemologists.[81] On this view, the causes of one's beliefs (partly) determine the method one is using. For example, a loving mother who assesses the piano skills of her child is not using an objective method, even though she might think she is, when her belief that her child is a prodigy is caused by her love, not by an assessment of her child's skills. However, if inputs partly individuate methods, then given that Peregrinus will have had, according to the evolutionary hypothesis, different input, they will have used different methods. The sensory input and/or the moral intuitions that cause their moral beliefs will be different from ours, and so they will have used a different method. Thus, even though similarity of belief-forming method might be relevant for assessing our counterfactual selves' accuracy, we get the result that Peregrinus is not using the same method. Hence, we have reason to recant our counterfactual selves' default status as peers.

Next consider the epistemic question of how to individuate methods. In this case, it is sensible to suppose that methods are partly individuated by outputs (which is, incidentally, also congenial to individuating methods externally). For example, your friend's genuine belief that you owe $444 for the lunch bill (for a meal that was almost certainly below $100), gives you an abductive reason for thinking that your friend is not using the same method (e.g., arithmetic) that you are to calculate how the bill is shared in the first place. After all, compared with the hypothesis that your generally reliable friend made a mistake, a better explanation is that she is just trying to be funny. More generally, since methods of belief formation ought to be accurate, we should expect that competently using the method yields comparable outputs—a kind of positional objectivity.[82] In the absence of any reason to assume incompetence or insincerity, widely diverging outputs are thus reason to believe that another method was used to form the judgment. Going back to our disagreement with Peregrinus, and the fact that our outputs differ completely, we plausibly get a positive reason to believe that we are *not* using the same method. At best, we have learned nothing that would warrant maintaining trust in THEM (because, at best, we should withhold judgment about whether they are using the same method). In any case, total disagreement about $M_{basic}$ coupled with the, at best, agnostic attitude about whether they are using the same method of belief formation gives us reason to assign less likelihood to our counterfactual selves being accurate in moral matters compared to us. Again, we would have reason to withdraw their peerhood status.

81 E.g., Pritchard, *Epistemic Luck*, 152ff. Thanks to an anonymous referee for suggesting this point.

82 Cf. Sen, "Positional Objectivity."

Therefore, plausible alternatives to *agreement* about $M_{basic}$ suggest that there are *no* good reasons to take our counterfactual selves as our peers on a broad conception of peerhood if there is total disagreement about $M_{basic}$. I have not considered all possible alternatives. However, there are good reasons to think that none will be successful. The crucial point is that we have to presuppose some standard by which we can ascertain what it means to get moral matters right.[83] A necessary component of such a standard seems to be (partial) agreement about $M_{basic}$. Recall that, before any disagreement, we take ourselves to have reasonable grounds to think that the members of $M_{basic}$ are largely true. Thus, we have reasonable grounds to believe that we are getting moral matters right, insofar as we believe in $M_{basic}$. Our counterfactual selves do not. So, if the evolutionary hypothesis implies total disagreement about $M_{basic}$, then objectivists need not be concerned.

In the next section, I consider and rebut a final option on behalf of the disagreement view to close my case for P3: *partial* disagreement about $M_{basic}$.

### 5.3. *Partial Disagreement between Peers on a Broad Conception*

Still staying in non-nearby scenarios, the much more plausible case is that we rewind the wheel of life, but only to a point where there is still *some* agreement about $M_{basic}$. Let partial disagreement be a case in which our counterfactual selves agree about at least one belief that is a member of $M_{basic}$. We might, therefore, have reason (though perhaps not sufficient reason) to count them as our peers on a broad conception of peerhood. Of course, there are fuzzy boundaries, and I do not suspect that we can say with precision whether agreement about some percentage of the members of $M_{basic}$ is required for peerhood. However, there could be enough agreement to raise the suspicion that "there is no reason to suppose that either party to the dispute is in an evidentially superior position."[84] So, debunking explanations could reveal *local* disagreement with our counterfactual selves. This might concern only some members of $M_{basic}$. However, this line of argument does not vindicate the disagreement view for two reasons.

First, if we consider just one counterfactual scenario, in which we end up like, say, Peregrinus, then the most that proponents of the disagreement view could conclude is that the justification of *some* beliefs is challenged. Such a case would not show, however, that all objectivist moral beliefs in $M_{basic}$ are defeated.

For example, it might be that we cannot determine whether it is morally permissible or impermissible to abort fetuses. However, this finding does not imply that all the other moral beliefs in $M_{basic}$ about which there is agreement are also

83 See Elga, "Reflection and Disagreement," 493ff.

84 Cohen, "A Defense of the (Almost) Equal Weight View," 98.

unreliable. To reach that conclusion, proponents of the direct approach would have to appeal to a principle of the following sort:

> *Token-Type*: If there is peer disagreement about a token of a type of proposition, *K*, then beliefs about that type of proposition are epistemically suspect.

However, the Token-Type principle is undoubtedly false. There may be radical disagreements about matters in physics, but we do not judge all beliefs about physics to be unjustified. Instead, it seems appropriate to judge that the question is beyond our (current) abilities to answer. Objectivists can adopt the same reasoning. There might be peer disagreement about the truth of the contents of some members of $M_{basic}$—and we might want to suspend judgment about those—but that need not compel us to suspend judgment about all beliefs in $M_{basic}$.[85]

Second, proponents of the disagreement view might argue as follows: if we consider manifold disagreements with manifold counterfactual selves, we could get cumulative total disagreement about $M_{basic}$. To illustrate, assume that $M_{basic}$ contains two nonoverlapping proper subsets: *A* and *B*. We agree with Peregrinus about *A* and disagree about *B*. Now imagine that there is another of our counterfactual species, say *Homo sapiens cerritulus*, the mad man. We agree with Cerritulus about *B* but disagree about *A*. As a result, there is peer disagreement about all beliefs in $M_{basic}$, albeit not with the same interlocutor.

However, that response is only initially plausible because it is unlikely that the evolutionary hypothesis implies that such a situation is possible. For one, the contents of the beliefs in $M_{basic}$ are ecologically related in worlds that are similar to ours.[86] If debunking explanations imply relevant disagreement about $M_{basic}$ with, say, Cerritulus, then Cerritulus's world would be very different from ours. Thus, it would be unlikely that Cerritulus would agree about the beliefs contained in set *B*. In other words, disagreement about some beliefs in $M_{basic}$ raises the probability of disagreement about other beliefs in $M_{basic}$, such that it is unlikely that there could be a cumulative disagreement about *all* beliefs in $M_{basic}$. Moreover, given that mere agreement about bits of $M_{basic}$ can be considered a necessary but not a sufficient condition for peerhood, it is not clear, and is certainly not established on the broad conception of peerhood, that imagining many deviant species with whom we have partial agreement establishes that there is relevant peer disagreement.

These considerations suggest that partial disagreement about $M_{basic}$ is plausi-

85 This might imply that there are some moral propositions that are unknowable on an objectivist account of morality (Wright, *Truth and Objectivity*).

86 See Morris, *Foragers, Farmers, and Fossil Fuels*, for an extended argument along these lines.

ble to some extent, but that it does not yield the desired conclusion that all beliefs in $M_{basic}$ are subject to justification-defeating disagreement. This concludes the case for premise P3. As I said, it is not a conclusive case. In particular, there might be other criteria based on which we could take our counterfactual selves to be our peers despite them disagreeing about all members of $M_{basic}$. Moreover, it might be possible to conjure up scenarios that are biologically possible in which there is a triangulated total disagreement about $M_{basic}$. However, proponents of the disagreement view have not made that case. As such, the considerations of this section suggest that, no matter how the tape of life is replayed, we will not find relevant, justification-defeating disagreement about objectivist moral beliefs, so moral objectivism has not been refuted by the disagreement view.

## 6. Conclusion

In conclusion, the evolutionary hypothesis does not show that there is relevant hypothetical disagreement that defeats the justification of all our objectivist moral beliefs. As a result, the disagreement view fails. If all alternative interpretations of the epistemic significance of the evolutionary hypothesis fail, as some proponents of the disagreement view claim, then evolutionary debunking arguments fail to have skeptical consequences for moral objectivists.

However, even without assuming the radical claim that the evolutionary hypothesis is relevant only if it implies hypothetical moral peer disagreement, this paper shows that appeals to disagreement do not help the debunker's case, which takes away one possible route for debunkers to press their skeptical conclusion. For defenders of moral objectivity, this means relief on one front. For their opponents, this means that they need to reinforce efforts to find another epistemic phenomenon to undergird evolutionary debunking arguments against the objectivity of morality. In the end, there is strong reason to believe that hypothetical disagreement, no matter how far we rewind the tape of life, does not help the debunker in the case against moral objectivism.[87]

*Delft University of Technology*
*m.b.o.t.klenk@tudelft.nl*

87 I am grateful to Jeff Behrends, Tomas Bogardus, Andreas Mogensen, Hanno Sauer, Folke Tersman, and Joeri Witteveen as well audiences at the Harvard/MIT philosophy of psychology group and the 9th European Congress of Analytic Philosophy for their comments and questions on earlier versions of this paper. I also thank two anonymous referees of this journal for their helpful suggestions.

## REFERENCES

Ballantyne, Nathan. "The Problem of Historical Variability." In *Disagreement and Skepticism*, edited by Diego E. Machuca, 239–58. New York: Routledge, 2013.

Barkhausen, Max. "Reductionist Moral Realism and the Contingency of Moral Evolution." *Ethics* 126, no. 3 (April 2016): 662–89.

Bengson, John. "Grasping the Third Realm." In *Oxford Studies in Epistemology*, vol. 5, edited by Tamar Szabo Gendler and John Hawthorne, 1–38. Oxford: Oxford University Press, 2015.

Bishop, Michael A. "Why the Generality Problem Is Everybody's Problem." *Philosophical Studies* 151, no. 2 (November 2010): 285–98.

Bogardus, Tomas. "Only All Naturalists Should Worry about Only One Evolutionary Debunking Argument." *Ethics* 126, no. 3 (April 2016): 636–61.

Christensen, David, and Jennifer Lackey, eds. *The Epistemology of Disagreement: New Essays*. Oxford: Oxford University Press, 2013

Chudnoff, Elijah. "What Intuitions Are Like." *Philosophy and Phenomenological Research* 82, no. 3 (May 2011): 625–54.

Clarke-Doane, Justin. "Justification and Explanation in Mathematics and Morality." In *Oxford Studies in Metaethics*, vol. 10, edited by Russ Shafer-Landau, 80–103. Oxford: Oxford University Press, 2015.

———. "Objectivity and Reliability." *Canadian Journal of Philosophy* 47, no. 6 (2017): 841–55.

Climenhaga, Nevin. "Intuitions Are Used as Evidence in Philosophy." *Mind* 127, no. 505 (January 2018): 69–104.

Cohen, Stewart. "A Defense of the (Almost) Equal Weight View." In Christensen and Lackey, *The Epistemology of Disagreement*, 98–117.

Curry, Oliver Scott. "Morality as Cooperation. A Problem-Centred Approach." In *The Evolution of Morality*, edited by Todd K. Shackelford and Ranald D. Hansen, 27–51. Dordrecht: Springer, 2016.

Curry, Oliver Scott, Daniel Austin Mullins, and Harvey Whitehouse. "Is It Good to Cooperate? Testing the Theory of Morality-as-Cooperation in 60 Societies." *Current Anthropology* (forthcoming).

Darwin, Charles. *The Descent of Man, and Selection in Relation to Sex*. 1871. Edited by Adrian Desmond and James R. Moore. London: Penguin Books, 2004.

Davidson, Donald. *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press, 1984.

Elga, Adam. "Reflection and Disagreement." *Noûs* 41, no. 3 (September 2007): 478–502.

Enoch, David. "The Epistemological Challenge to Metanormative Realism:

How Best to Understand It, and How to Cope with It." *Philosophical Studies* 148, no. 3 (April 2010): 413–38.

———. "Not Just a Truthometer: Taking Oneself Seriously (but Not Too Seriously) in Cases of Peer Disagreement." *Mind* 119, no. 476 (October 2011): 953–97.

———. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press, 2011.

Feldman, Richard. "Epistemological Puzzles about Disagreement." In *Epistemology Futures*, edited by Stephen C. Hetherington, 216–36. Oxford: Clarendon Press, 2006.

Feldman, Richard, and Ted A. Warfield, eds. *Disagreement*. Oxford: Oxford University Press, 2010.

FitzPatrick, William J. "Debunking Evolutionary Debunking of Ethical Realism." *Philosophical Studies* 172, no. 4 (April 2015): 883–904.

Foley, Richard. *Intellectual Trust in Oneself and Others*. Cambridge: Cambridge University Press, 2001.

Gelfert, Axel. "Who Is an Epistemic Peer?" *Logos and Episteme* 2, no. 4 (2011): 507–14.

Goldberg, Sanford C. "Disagreement, Defeat, and Assertion." In Christensen and Lackey, *The Epistemology of Disagreement*, 167–89.

Handfield, Toby. "Genealogical Explanations of Chance and Morals." In *Explanation in Ethics and Mathematics*, edited by Uri D. Leibowitz and Neil Sinclair, 58–82. Oxford: Oxford University Press, 2016.

Harman, Gilbert. *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press, 1988.

Harms, William F. *Information and Meaning in Evolutionary Processes*. Cambridge: Cambridge University Press, 2004.

Henrich, Joseph, and Natalie Henrich. "The Evolution of Cultural Adaptations: Fijian Food Taboos Protect against Dangerous Marine Toxins." *Proceedings of the Royal Society of London B: Biological Sciences* 277, no. 1701 (2010): 3715–24.

Horn, Justin. "Moral Realism, Fundamental Moral Disagreement, and Moral Reliability." *Journal of Value Inquiry* 51, no. 3 (September 2017): 363–81.

Huemer, Michael. *Ethical Intuitionism*. Basingstoke: Palgrave Macmillan, 2005.

———. "A Liberal Realist Answer to Debunking Skeptics: The Empirical Case for Realism." *Philosophical Studies* 173, no. 7 (July 2016): 1983–2010.

———. "The Problem of Defeasible Justification." *Erkenntnis* 54, no. 3 (May 2001): 375–97.

Jackson, Frank, and Philip Pettit. "Moral Functionalism and Moral Motivation." *Philosophical Quarterly* 45, no. 178 (January 1995): 20–40.

Joyce, Richard. "Arguments from Moral Disagreement to Moral Skepticism." In *Moral Skepticism: New Essays*, edited by Diego E. Machuca, 141–62. New York: Routledge, 2018.

———. *The Evolution of Morality*. Cambridge, MA: MIT Press, 2006.

Kelly, Tom. "The Epistemic Significance of Disagreement." In *Oxford Studies in Epistemology*, vol. 1, edited by Tamar S. Gendler and John P. Hawthorne, 167–96. Oxford: Clarendon Press, 2005.

King, Nathan L. "Disagreement: What's the Problem? or A Good Peer Is Hard to Find." *Philosophy and Phenomenological Research* 85, no. 2 (September 2012): 249–72.

Klenk, Michael. "Can Moral Realists Deflect Defeat Due to Evolutionary Explanations of Morality?" *Pacific Philosophical Quarterly* 98, no. 19 (December 2017): 227–48.

———. Review of *Robust Ethics: The Metaphysics and Epistemology of Godless Normative Realism*, by Erik Wielenberg. *Dialectica* 70, no. 3 (September 2016): 482–88.

———. "Survival of Defeat: Evolution, Moral Objectivity, and Undercutting." PhD Diss., Utrecht University, 2018.

———. "Third Factor Explanations and Disagreement in Metaethics." *Synthese*. Published ahead of print, July 12, 2018. https://link.springer.com/article/10.1007/s11229-018-1875-8.

Kohlberg, Lawrence, and Richard H. Hersh. "Moral Development: A Review of the Theory." *Theory Into Practice* 16, no. 2 (April 1977): 53–59.

Kölbel, Max. "Faultless Disagreement." *Proceedings of the Aristotelian Society* 104 (2004): 53–73.

Kornblith, Hilary. "Belief in the Face of Controversy." In Feldman and Warfield, *Disagreement*, 29–52.

Leiter, Brian, ed. *The Future of Philosophy*. Oxford: Clarendon Press, 2004.

Lillehammer, Hallvard. "Methods of Ethics and the Descent of Man: Darwin and Sidgwick on Ethics and Evolution." *Biology and Philosophy* 25, no. 3 (June 2010): 361–78.

Lutz, Matt. "What Makes Evolution a Defeater?" *Erkenntnis*. Published ahead of print, August 7, 2017. https://doi.org/10.1007/s10670-017-9931-1.

Matheson, Jonathan. *The Epistemic Significance of Disagreement*. Basingstoke: Palgrave Macmillan, 2015.

Mogensen, Andreas L. "Contingency Anxiety and the Epistemology of Disagreement." *Pacific Philosophical Quarterly* 97, no. 4 (December 2016): 590–611.

———. "Disagreements in Moral Intuition as Defeaters." *Philosophical Quarterly* 67, no. 267 (April 2017): 282–302.

Morris, Ian. *Foragers, Farmers, and Fossil Fuels: How Human Values Evolve*. Princeton: Princeton University Press, 2015.

Pritchard, Duncan. *Epistemic Luck*. Oxford: Oxford University Press, 2005.

Scanlon, T. M. *Being Realistic about Reasons*. Oxford: Oxford University Press, 2014.

Sen, Amartya. "Positional Objectivity." *Philosophy and Public Affairs* 22, no. 2 (Spring 1993): 126–45.

Shafer-Landau, Russ. *Moral Realism: A Defence*. Oxford: Clarendon Press, 2003.

Sher, George. "But I Could Be Wrong." *Social Philosophy and Policy* 18, no. 2 (Summer 2001): 64.

Sidgwick, Henry. *The Methods of Ethics*. 7th ed. 1874. Indianapolis, IN: Hackett Publishing, 1981.

Sinclair, Neil. "Belief-Pills and the Possibility of Moral Epistemology." In *Oxford Studies in Metaethics*, vol. 13, edited by Russ Shafer-Landau, 98–122. Oxford: Oxford University Press, 2018.

Sinnott-Armstrong, Walter. "Moral Disagreements with Psychopaths." In *Challenges to Moral and Religious Belief: Disagreement and Evolution*, edited by Michael Bergmann and Patrick Kain, 40–60. Oxford: Oxford University Press, 2014.

Skarsaune, Knut O. "Darwin and Moral Realism: Survival of the Iffiest." *Philosophical Studies* 152, no. 2 (January 2011): 229–43.

Street, Sharon. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127, no. 1 (January 2006): 109–66.

———. "Objectivity and Truth: You'd Better Rethink It." In *Oxford Studies in Metaethics*, vol. 11, edited by Russ Shafer-Landau, 293–334. Oxford: Oxford University Press, 2016.

Tersman, Folke. "Debunking and Disagreement." *Noûs* 51, no. 4 (December 2017): 754–74.

———. *Moral Disagreement*. Cambridge: Cambridge University Press, 2006.

———. "Moral Disagreement: Actual vs. Possible." In *Disagreement and Skepticism*, edited by Diego E. Machuca, 90–108. New York: Routledge, 2013

Vavova, Katia. "Evolutionary Debunking of Moral Realism." *Philosophy Compass* 10, no. 2 (February 2015): 104–16.

———. "Moral Disagreement and Moral Skepticism." *Philosophical Perspectives* 28, no. 1 (December 2014): 302–33.

Weatherson, Brian. "Disagreements, Philosophical and Otherwise." In Christensen and Lackey, *The Epistemology of Disagreement*, 1–17.

Wedgwood, Ralph. "The Moral Evil Demons." In Feldman and Warfield, *Disagreement*, 216–46.

White, Roger. "You Just Believe That Because …" *Philosophical Perspectives* 24, no. 1 (December 2010): 573–615.

Wielenberg, Erik J. "On the Evolutionary Debunking of Morality." *Ethics* 120, no. 3 (April 2010): 441–64.

———. *Robust Ethics: The Metaphysics and Epistemology of Godless Normative Realism*. Oxford: Oxford University Press, 2014.

Williamson, Timothy. *Knowledge and Its Limits*. Oxford: Oxford University Press, 2000.

Wright, Crispin. *Truth and Objectivity*. Cambridge, MA: Harvard University Press, 1992.

Wynn, Thomas, and Frederick L. Coolidge. *How to Think Like a Neandertal*. New York: Oxford University Press, 2013.