SPECIAL ISSUE ARTICLE

The argument from agreement: How universal values undermine moral realism

Hanno Sauer 🆔

Department of Philosophy and Religious Studies, Utrecht University, Janskerhof 13, Utrecht, 3512 BL, The Netherlands

Correspondence
Hanno Sauer, Department of Philosophy and Religious Studies, Utrecht University, Janskerhof 13, Utrecht 3512 BL, The Netherlands.
Email: h.c.sauer@uu.nl

Abstract
The most popular argument against moral realism is the argument from disagreement: if there are mind-independent moral facts, then we would not expect to find as much moral disagreement as we in fact do; therefore, moral realism is false. In this paper, I develop the flipside of this argument. According to this *argument from agreement*, we would expect to find lots of moral *dis*agreement if there were mind-independent moral facts. But we do *not*, in fact, find much moral disagreement; therefore, moral realism is false. I defend the argument, explain the empirical evidence that supports it, and show what makes this challenge novel and powerful.

KEYWORDS
moral disagreement, moral realism, moral relativism, moral universalism

1 | INTRODUCTION

What if there were no moral disagreements? Should moral realists welcome this fact? In this paper, I will argue that they should not, and that the lack of moral disagreement spells trouble for the view that there are mind-independent moral facts. Widespread agreement on moral matters constitutes a challenge to realism, and there is excellent evidence for the claim that there is, a matter of empirical fact, almost no fundamental moral disagreement in the real world.

Moral realism is the view that there are mind-independent facts about moral norms and values.[1] Perhaps the most popular challenge to moral realism draws on the allegedly widespread phenomenon of moral disagreement.

---

[1] In addition to this metaphysical claim, it is usually thought to have an epistemological side as well (see Kahane, 2013), according to which we can have successful epistemic access to those objective normative facts. But for the purposes of this paper, I will focus on its metaphysical commitments. Recently, there have been a number of new and promising challenges to moral realism, see Street (2006), Barber (2013), McGrath (2011), or Sauer (2018b). This paper makes a contribution to this new wave of anti-realist arguments.

From Herodotus' *Histories* over Montaigne's *Of Cannibals* to early Twentieth century anthropologists, the idea that intra- and intercultural evaluative diversity is somehow threatening to the existence of objective moral facts has held a lot of sway.[2] This is known as the *argument from disagreement*.[3]

I will defend the anti-realist's conclusion, but on converse grounds. In what follows, I will argue that it is neither true that widespread disagreement challenges moral realism, nor that we find any widespread moral disagreement in the real world. What I will claim is that realism about any given domain, including the moral one, predicts widespread *dis*agreement, and that what we find is widespread *agreement* about basic moral norms and values at the metaethically relevant level. This puts new pressure on moral realism. Interestingly, we arrive at the same anti-realist conclusion by rejecting both premises of what many think is the most powerful argument in its favor. I call this the *argument from agreement*.

My paper has four sections. In the first (1), I will briefly sketch how the traditional argument from disagreement is supposed to work and how to get to the argument from agreement by using it as a starting point. In section (2), I will provide a wealth of empirical evidence for the claim that there actually is widespread moral agreement and almost no fundamental disagreement about values. In (3) I will argue that, contrary to what is commonly supposed, moral realism – or any kind of realism for that matter – predicts widespread and fundamental *dis*agreement about the respective domain at issue. In section (4), I consider some promising objections to the argument from agreement and show how to make it work. In particular, I will highlight how difficult it is for *both* sides of the debate to spell out the analogy *or disanalogy* between moral and scientific beliefs, and to decide which subsets of those domains to compare to each other to support or undermine the case for realism.

2 | THE ARGUMENT FROM (DIS)AGREEMENT

Here is the traditional

> *Argument from Disagreement*
> (D1) If moral realism is true, then we would not expect much moral disagreement.
> (D2) There is lots of moral disagreement.
>   Therefore,
> (D3) Moral realism is false.[4]

This (fairly standard) rendition of the argument raises several interesting questions regarding the *quantity* and *quality* of the disagreement at issue.

What would it mean to find a lot or little disagreement? It would be implausible to suggest that realism predicts *zero* disagreement – that is: no one ever disagrees with anyone else about any moral issue – or that non-realists predict *perfect* disagreement – that is, everyone always disagrees with everyone else about all moral issues. So what do realists and anti-realists predict about the quantity of moral disagreement? Here, I assume that there is

---

[2] For a helpful summary of the various branches of relativism, see Fricker (2013). Fricker makes the interesting observation that in fact, Montaigne is criticizing cultural hypocrisy rather than making a genuinely relativistic point.

[3] An excellent overview of the most promising versions of the argument from disagreement can be found in Enoch (2009). Useful starting points for the modern debate are Harman (1975) and Brink (1984). Realist perspectives on disagreement can be found in Huemer (2005) and Shafer-Landau (2003). For the recent empirical revival of the argument from disagreement, see Wong (2006), Prinz (2007) and Doris and Plakias (2008; see also Fitzpatrick, 2014; Fraser & Hauser, 2010; Leiter, 2008; Meyers, 2013; Sauer, 2018a, chapter 4; Sneddon, 2009).

[4] Mackie (1977), for instance, conceived of this argument as an inference to the best explanation: the best explanation for the widespread disagreement we find about moral matters is that there are no mind-independent moral facts for people to agree about. Morality is a matter of subjective preference, more akin to taste than mathematics (Gill, 2007; on 'rationalist' interpretations of the sentimentalist thesis, see Sauer, 2011, 2017, chapters 5 and 6). More recently, Richard Rowland (2017) has argued that realism is incompatible with disagreement under idealized conditions. McGrath (2008) makes this point as well.

some threshold of disagreement beyond which both parties to the debate would agree that realism becomes difficult to defend. Whatever that threshold is, that is the one I will rely on as well. I am not suggesting here that there is an agreed-upon quantity of disagreement within this debate. The point is, rather, that those who take sides in the debate surrounding the argument of disagreement must have some ballpark beyond zero and short of total disagreement in mind to engage with the issue at all. This ballpark is what I will assume as well. Of course, I will argue that the very assumption that diversity is a threat to realism is actually a mistake, but as a characterization of the *de facto* dialectical situation, it seems appropriate.

Most authors who defend moral realism against the argument from disagreement do so by rejecting D2. This is typically done by invoking so-called *defusing explanations* that are supposed to show that what seems to be genuine *moral* disagreement actually turns out to be superficial disagreement about non-moral facts or something else, thus rendering the given case of *prima facie* disagreement non-fundamental and thereby unthreatening to the convergentist moral realist. Fewer authors decide to bite the bullet and accept the existence of widespread moral disagreement, instead challenging the idea that moral realism is incompatible with disagreement in the first place. Divergentism is the claim thats 'disagreement poses no threat to realism of any stripe, and so, a fortiori, poses no threat to moral realism in particular' (Shafer-Landau, 2003, p. 228). As we will see, I agree with this statement, but aim to take it in an altogether different direction.

What *opponents* of moral realism haven't done, to my knowledge at least, is to reject *both* (D1) *and* (D2) at the same time, but to do so in order to *challenge*, rather than defend, moral realism. There are many realists who hold that disagreement is at least compatible with realism (thereby denying D1). And there are some who are willing to concede that the case for agreement and convergence is stronger than usually thought (thereby denying D2). But few, if any, authors endorse the *conjunction* of these two claims, which I refer to as the

> *Argument from Agreement*
> (A1) If moral realism is true, then we *would* expect a lot of moral disagreement.
> (A2) We do *not* see a lot of moral disagreement.
>   Therefore,
> (A3) Moral realism is false.

It is not without irony that both premises in the argument from agreement are separately endorsed by many adherents of realism, and an interesting and thus far underappreciated fact about the argument from *dis*agreement that denying both of its premises yields its original conclusion. To resist the argument from agreement, realists must show why their fellow realists are mistaken.

Why would anyone endorse this argument, which most have found too implausible to give serious consideration? At the very least, it should trouble realists that ethical experts have discovered very little, if anything, that is not already in some shape or form known to the folk. The important 'discoveries' ethicists have made (think: gender equality, the wrongness of slavery, the 'innocent until proven guilty' principle) either draw on insights which are perfectly available to lay people, or have been made at pretty much the same time that the rest of society has made them as well.[5] This is an oddity in its own right. Prima facie, realism often predicts disagreement.

Some authors come close to endorsing this view. McGrath (2008), for instance, also argues that moral realism does not require convergence. (Since (A1) holds that moral realism requires divergence, I also accept her weaker thesis.) However, she, like most people, uses this compatibility claim as a line of *defense* of realism. She does not go one step further, saying that realism predicts disagreement and that we actually find lots of agreement: 'It is *obvious* that there is a great deal of actual moral disagreement' (McGrath, 2008, p. 60, my emphasis). In section (3), I will show that this is not just not obvious, but that the empirical record suggests that there is virtually no fundamental moral disagreement at all.

---

[5] Thanks to Michael Huemer for helpful comments on this issue.

My argument does not require that all forms of moral realism be challenged by agreement. Who, then, is the target? Most obviously, non-naturalist forms of realism seem to qualify. Realists such as Huemer (2008) hold that their intuitionist epistemology is a revisionary one. Competent moral judgment is difficult to do, and has the psychological deck stacked against it. This means that only few people will wind up appreciating the moral facts. These few, then, will disagree with the rest. It would be evidence against this position if it turned out that as far as fundamental moral beliefs are concerned, there is almost no revision, or that the revisionary aspects of modern moral outlooks are everywhere explained by the discovery of non-moral facts. Naturalists also seem to qualify. Some versions of moral naturalism hold that the acquisition of moral knowledge should be modeled after the discovery of empirical facts (Tropman, 2012, 2014). Moral facts are *discovered*, and if this is so, then it seems plausible to suggest that not everyone discovered them automatically. Other types of realism may qualify as a target of my argument as well, though some may not.[6]

As mentioned before, by far the most popular strategy for realists has been to show that cases of apparent moral disagreement can somehow be explained away in terms of disagreement about the non-moral facts, failures of interpretation, irrationality, special pleading, and so on. The lesson to take away from this first line of resistance is that how to zoom in on the 'right kind' of disagreement is indeed a tricky issue (Sauer, 2018a). The important thing to file away at this point is that virtually everyone agrees that relevant disagreements must not be 'superficial'. That is, they must not boil down to disagreements about the supervenience basis of moral facts. Other defusing explanations of purported moral disagreements must be blocked as well. The argument from agreement is compatible with this claim. It is indeed obvious that there is lots of non-fundamental moral disagreement. (A2) should thus be understood as the claim that there is not a lot of *fundamental* moral disagreement. I will argue that in fact, there is hardly any at all.

3 | EVIDENCE FOR AGREEMENT

What would it take for people to be in moral agreement? Firstly, moral agreements would have to be just as fundamental as the disagreements put forward by advocates of D2, i.e. not explainable by superficial agreement on the non-moral supervenience basis of moral truths. Suppose that there are two people who both think that there should be no differences in average income between two given groups. However, one person actually thinks that income should be proportional to intelligence, while the other does not. Both people merely happen to agree that there are no significant differences in average intelligence between the members of both groups. This is merely superficial moral agreement. Secondly, we would not expect *perfect* agreement: divergence in emphasis, such as the relative importance of promise-breaking or well-being, is allowed. Thirdly, the metaethically interesting level of disagreement concerns basic moral norms and values. People may hate each other's guts because one believes in socialism and the other in capitalism. Metaethically, this disagreement, however passionate, may turn out to be thoroughly irrelevant if both disagreeing parties care about the well-being of the poor, while simply disagreeing about which economic system has the best track record when it comes to realizing this value.

It is crucially important to have a steady grasp of what constitutes fundamental rather than superficial disagreement. Unfortunately, such as grasp is very difficult to come by. For one thing, it seems that the argument from agreement runs into something of a dilemma.[7] Disagreement or agreement, it seems, can be fundamental in one of two ways. A moral belief can be *epistemically* fundamental in virtue of being non-inferentially justified. On this notion, it seems plausible to suggest that there is little fundamental moral disagreement. However, since my

---

[6] *A priorists* like Cuneo and Shafer-Landau (2014) may not qualify. But they may not count as realist in my sense, because if the moral fixed points are conceptual truths, and conceptual truths are mind-dependent, then they are not realist, at least not in the sense I am interested here. I will remain agnostic about this issue.

[7] The following discussion draws heavily on comments I received from Michael Huemer.

argument is largely based on exploiting disanalogies, it now becomes a problem that the fundamental (so under-stood) non-moral beliefs do not encounter much disagreement either, because these non-inferentially justified beliefs mostly consist of observations. Or, one could understand fundamental moral beliefs in an explanatory sense, such that the fundamental moral truths are the ones that explain why the less fundamental moral truths hold. On that notion, there clearly is a sufficient amount of non-moral disagreement – the scientific experts have come up with explanations of the phenomena deeply at odds with what lay people believe. The problem, then, is that the same seems to hold for the moral domain. Here, the ethical experts have devised theories regarding what grounds rightness and wrongness which are in conflict with lay people's opinions on these matters. The funda-mental moral truths would be adequately captured by esoteric theories such as utilitarianism or deontology.

What level of disagreement is there *at the level of moral theory*? For instance, there seems to be genuine dis-agreement between consequentialists and Kantians, or Kantians and virtue theorists, or virtue theorists and con-sequentialists.[8] However, it would be misleading to suggest that these theorists are either a) insensitive to the moral concerns of the others or b) endorse fundamentally different deontic outputs. Different moral theorists simply think that one theory provides a more elegant and simple account of whatever our moral duties are (about which they largely agree). Indeed, this point also ricochets, because it should be considered surprising that we only find this narrow space of options of three basic normative theories, and have found essentially only those for millennia. For why, if realism is true, would that be so?[9]

If individuals and cultures actually agree about what morality demands, where does the *appearance* of dis-agreement – on which proponents as well as critics of the argument of disagreement tend to rely – come from? Here, I think the distinction between fundamental and superficial disagreement starts to play an unexpected role. Traditionally, this distinction has been used by defenders of realism to show that most apparent disagreements can actually be explained away in realism-friendly terms. The idea was that upon closer scrutiny, most candidates for genuine moral disagreement actually turn out to be disagreements about non-moral facts or simply misinter-pretations of other cultures (Moody-Adams, 2009) and therefore fail to establish any relativistic conclusions. The suggestion I wish to make is that the distinction between superficial and fundamental disagreements does not explain away disagreement, but explains the *appearance* of widespread moral disagreement. It is correct that peo-ple and cultures *seem* to disagree about morality all the time; but many, if not most, of these disagreements turn out to be superficial and defusable: this means that actually, there is a lot of *fundamental moral agreement* around.

This means that defusing explanations of moral disagreement come back to haunt the moral realist. Originally, these explanations were supposed to show that most cases of moral disagreement actually turn out to be super-ficial and thus unthreatening to the realist. And this is exactly right: most disagreements *are* superficial and non-fundamental. The examples mentioned above – pain is bad, cooperation is good, cheaters ought to be sanctioned, morality is about harm, fairness, loyalty, respect and purity, and so forth – identify fundamental moral issues, and on those issues, there is overwhelming agreement among virtually all cultures and individuals. But this be-comes a problem for moral realists once we realize that moral realism never actually required fundamental moral agreement in the first place. It requires *dis*agreement, so that the aforementioned defusing explanations end up infecting, rather than immunizing, moral realism.

There is no principled way of classifying an instance of moral (dis)agreement as fundamental or superficial. In the end, I propose to let the empirical evidence speak largely for itself. Do the following examples for moral values strike us as being due to disagreement about a non-moral supervenience basis, or irrationality in the application of moral values? And if they do not, then by the standards deployed by friends and foes of the argument from dis-agreement, the examples mentioned below are sufficiently fundamental to count as metaethically relevant. What is the evidence for moral agreement, so understood?

---

[8] Thanks to Nicholas Laskowski for helping me phrase this point.

[9] This, too, can be explained evolutionarily, see Heath (2009).

(1) *Basic evaluative dispositions*. Selective pressures have bequeathed to us a set of hardwired basic evaluative dispositions. Street (2006) lists the following:

1. The fact that something would promote one's survival is a reason in favor of it.
2. The fact that something would promote the interests of a family member is a reason to do it.
3. We have greater obligations to help our own children than we do to help complete strangers.
4. The fact that someone has treated one well is a reason to treat that person well in return.
5. The fact that someone is altruistic is a reason to admire, praise, and reward him or her.
6. The fact that someone has done one deliberate harm is a reason to shun that person or seek his or her punishment (115).

Street frames her list in terms of what people have reason to do, but nothing hangs on this. For my purposes, it is enough to recognize that, for instance, human beings treat the fact that something would promote one's survival as something that favors doing it. Human beings have a basic disposition to value survival. Moreover, people overwhelmingly agree that pain is bad, cheaters should be punished, cooperation rewarded, and so forth. Let me emphasize already that there is widespread agreement between those people who could claim to be moral 'experts' of some sort *and* laypeople on such basic evaluative facts. Because of their evolutionary rationale – the people whose ancestors did not have the above dispositions simply are not around anymore – these dispositions are likely to be deeply entrenched and thus widespread.

(2) *Cooperative strategies*. Recent evidence suggests that cooperation and the various strategies that generate cooperative gains are universally considered morally good. Oliver Curry (2016), who explicitly sets up his 'morality as cooperation' theory as a challenge to descriptive moral relativism, distinguishes seven such strategies:

1. Allocation of resources to kin
2. Coordination to mutual advantage
3. Social exchange
4. Hawkish traits (dominance as conflict resolution)
5. Dove-ish traits (submission as conflict resolution)
6. Division/Fairness
7. Possession/Property[10]

This theory, according to which cooperation is the unifying thread within morality and allows us to distinguish genuinely moral concerns from others, has two main advantages. Firstly, it is based on a more systematic foundation than competing approaches such as moral foundations theory (Graham, Haidt, & Nosek, 2009), whose list of moral foundations, according to Curry, is compiled in an '*ad hoc*' (10) fashion without much predictive power.[11] Second, it is supported by impressive cross-cultural evidence. Curry and his colleagues looked at the ethnographic record of 60 societies specifically selected to be representative of all of humanity, including samples from all continents and forms of social organization, spanning over several centuries. The material they investigated was drawn from the so-called *Human Relations Area Files*, which they had independently coded for paragraphs pertaining to *Ethics* or *Norms*, yielding around 400 paragraphs from over 500 documents from the aforementioned 60 societies. They operationalized the cooperative strategies 1.–7. (for example, kinship was operationalized in terms of instances where (only) family or in group members were helped, and so forth) and coded the strategies so operationalized in terms of whether they were

---

[10] Some items on this list, particularly 4. and 5., may not sound very moral to some. But the traits Curry has in mind here are close to traditional virtues such as courage and skill (4.) and humility and respect (5.). Conversely, the theory predicts that uncooperative strategies – 'neglecting kin, betraying one's group, free-riding, cowardice, disrespect, unfairness and theft' (11) will be universally regarded as morally bad, which is very plausible.

[11] However, disgust as a cooperative strategy is missing from his account, cf. Leeson (2013).

considered good or bad (for example, by being described as ethical, virtuous, and so forth). The strength of their results is hard to overstate: in 99.9% of cases, the aforementioned types of cooperation were considered morally good. These findings provide very strong evidence that fundamental moral agreement is extremely widespread and robust.

(3) Cross-cultural values. Cross-cultural evidence suggests that all people all over the world basically agree on which values there are (Alfano 2016). The Schwartz' (1994) Value Survey indicates that the following values are universally shared:

> Power: Social status and prestige, control or dominance over people and resources
>
> Achievement: Personal success through demonstrating competence according to social standards.
>
> Hedonism: Pleasure and sensuous gratification for oneself.
>
> Stimulation: Excitement, novelty, and challenge in life.
>
> Self-direction: Independent thought and action-choosing, creating, exploring.
>
> Universalism: Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature.
>
> Benevolence: Preservation and enhancement of the welfare of people with whom one is in frequent personal contact.
>
> Tradition: Respect, commitment, and acceptance of the customs and ideas that traditional culture or religion provide.
>
> Conformity: Restraint of actions, inclinations and impulses likely to upset or harm others and violate social expectations or norms.
>
> Security: Safety, harmony, and stability of society, of relationships, and of self. (Schwartz, 1994, p. 22)

Here, too, the material was sampled from 25.863 socioeconomically diverse adults from 44 countries from all continents. Respondents rated the values presented to them on a 9-point scale ranging from 'of supreme importance' (=7) to 'opposed to my values' (=−1). Here, the point is not that these values are not emphasized differently. They are. But for the issue of whether there is fundamental moral disagreement, such differences in emphasis do not matter. What matters is that the above values are universally endorsed in all cultures included in the sample.

(4) *Political Values*. Political psychology suggests that people's moral beliefs can be traced back to a number of moral foundations, a general sensitivity to which is shared (Haidt, 2012). Haidt distinguishes six of them:

1. Care/harm
2. Fairness/cheating
3. Loyalty/betrayal
4. Authority/subversion
5. Sanctity/degradation
6. Liberty/oppression[12]

Virtually everyone is at least somewhat sensitive to the moral relevance of all six foundations, even though different people – and, in particular, different political groups – assign more (or less) importance to different foundations. Liberals famously emphasize the first two foundations at the expense of the others; conservatives have a more balanced appreciation of all six foundations (Sauer, 2015). This, however, does not support moral relativism; it merely explains why there are differences between basic political outlooks such as liberalism and conservatism. Whether or not Haidt's Moral Foundations Theory is the correct account of the origin of moral intuitions does not matter for my purposes. What matters here is that *something like* a theory of moral foundations is probably correct, regardless of

---

[12] Haidt's *The Righteous Mind* still operates within his earlier five moral foundations framework. The liberty/oppression foundation was added more recently, see Iyer, Koleva, Graham, Ditto, and Haidt (2012).

whether it is Haidt's five (or six) moral foundations, Rozin's CAD triad/Shweder's 'Big Three', Rai and Fiske's (2011) four forms of relationship regulation or indeed Curry's seven pillars of morality as cooperation.

An interesting final data point for the idea that there is not actually a lot of moral disagreement is this: consider the people with the most radically divergent moral views you can think of. Now consider how surprisingly non-radical these disagreements are. In particular, compare how radical they *could* be, and how radical disagreements can in fact become when we look at other domains about which we are confident realists. Moral disagreements often have the following flavor: the Maragoli think fairness requires that one give 25% in the Ultimatum Game, whereas US citizens believe that one ought to give around 45% (Henrich, Heine, & Norenzayan, 2010). But in principle, it would not be inconceivable for one culture to think that harming others is wrong and that cooperation should be rewarded, for another culture to think that cooperation is wrong bad and that harming others is good, and for another culture to think that counting blades of grass is the best way to spend one's time or to jump up and down and scream at purple things (Street, 2006). But in reality, that is, in lived societies populated by people of flesh-and-blood, we *never find any disagreements anywhere near this strong*.

Now consider the radically divergent views people actually hold about realistic domains such as scientific truths, for instance about the origin of the cosmos. Some believe that it was created in six days by an omnipotent immaterial consciousness, some believe it had no origin but that the cosmos is a cyclical spiral, some hold that it was given birth to by a giant turtle. The range of moral beliefs actually held by people both syn- and diachronically is nowhere near this wide. This fact alone calls for an explanation, and it calls especially loudly on the moral realist.[13]

4 | REALISM REQUIRES DISAGREEMENT

Realism about a given domain predicts considerable disagreement within that domain rather than agreement, consensus or convergence. Indeed, *agreement* should be considered suspicious, especially when there is a non-truth-tracking explanation for why there would be agreement even if there were not any mind-independent facts to bring it about.

The alleged connection between agreement within a domain and realism about that domain is explicitly established by many advocates of moral realism. Here is a reconstruction by David Enoch (2011):

> [...] (1) In many discourses there is wide-ranging agreement about the truths central to the relevant discourse.
> [...] (2) What best explains such wide-ranging agreement is that there are objective truths the discourse answers to, truths on which opinions gradually converge.
> [...] (3) Therefore, there are objective truths the relevant discourse answers to. (From (1) and (2), by inference to the best explanation.) (196)

The best explanation, Enoch suggests, for wide-ranging agreement about a given domain is that this domain contains objective facts.

According to my argument, Enoch's premise (2) is false. What best explains agreement is precisely the *lack* of objective truths, because when there are such truths, *dis*agreement is to be expected. In general, the reason for this is that most of the interesting truths within a domain will be unobvious. Discovering them requires often painstaking inquiry and methodical reasoning, frequently conducted by professional investigators. Clearly, many philosophers think that this is precisely what is not required for obtaining moral knowledge. I will argue below

---

[13] On the other hand, the empirically best-supported explanations given for patterns of disagreement almost always rely on differences in economic circumstances that explain differences in fairness intuitions or sharing dispositions. Even the most outlandish beliefs and practices, such as wife selling in 19th century England or medieval trial by ordeal ultimately do turn out not to be based on fundamental evaluative divergence, but on limited information and rational incentives (Leeson, 2017).

that this assumption is unwarranted. Given that most truths within a given realist domain should not be readily accessible, widespread agreement about the truths of that domain should make us suspicious.

The claim that realism does not predict agreement becomes especially plausible when there are other, non-truth tracking explanations for why people would come to converge on a set of domain-specific beliefs even if there were no mind-independent facts those beliefs answer to. For instance, if we find that the members of a community agree on certain beliefs because of the pernicious influence of some charismatic guru, the envisioned inference to the best explanation is blocked. In the case of morality, such a non-truth-tracking explanation could, for instance, be supplied by selective pressures of biological or cultural evolution, more on which below.

Note that I am not committed to denying that realism predicts convergence over time. Convergence over time and widespread agreement are separate issues. Realism predicts merely that there will be (some) convergence among well-informed people, i.e. the experts. But typically, experts then find themselves disagreeing with laypeople, who were not swept up in the same dynamic of evidence-driven consensus formation which leads to widespread convergence among the informed, combined with widespread divergence between the informed and the uninformed. The claim that realism predicts both widespread convergence *and agreement* on moral matters thus seems to assume that there could be no laypeople/expert divide within morality. But this, too, sits uncomfortably with realism (McGrath, 2011; cf. Enoch, 2014). For why, if there are mind-independent moral facts, would there be no one better positioned to know them than others?

The main reason for thinking that realism about a given domain predicts disagreement is that when we look at domains about which most people are confident realists – such as science – we find lots of disagreement. By this, I do not mean to suggest that scientific experts do not converge on the basic tenets of their respective disciplines. What I mean is that when comparing scientific experts and laypeople outside of a given discipline, we find that almost no one, globally speaking, believes the basic truths of science, be it physics, biology, economics, or cognitive science. Almost all of the foundational insights of these disciplines are only believed by an extremely tiny minority of experts or people who have received some formal training in them. All others either believe these claims to be untrue or are, at the very least, thoroughly unaware of what these claims are in the first place.

Most scientific truths are deeply counterintuitive, and fairly recalcitrantly so. Counterintuitive claims are, by their very nature, unlikely to be believed by many people who haven't received some sort of training (or indeed indoctrination, as with the counterintuitive teachings of many cults and religions). Consider physics: there is nothing intuitive about the idea of inertia, the relativity of simultaneity, or the mysteries of quantum mechanics. Folk physics, on the other hand, is intuitively compelling but gets it all wrong (McCloskey, Washburn, & Felch, 1983). Consider biology: even today, the idea that natural selection (and other evolutionary pressures) instead of the vastly more viscerally appealing ideas of intelligent design or Lamarckianism remains deeply counterintuitive (Medin & Atran, 2004). Finally, consider economics: economists routinely complain about the fact that the public as well as elected officials fail to grasp the basic workings of the price mechanism, comparative advantage or the nature of public goods. That is because prices and global trade are strange, and difficult to comprehend (Kahneman, Knetsch, & Thaler, 1986).

Widespread agreement across the board is likely due to the fact that there are no counterintuitive truths in a domain for experts to discover over time, and for non-experts to remain unaware of. Which is to say that realism about that domain is probably false.

5 | BUT *DOES* IT?

The are several possible responses to my earlier claim that realism requires divergence. For one thing, realists can reject *invidious comparisons* like the analogy between scientific and moral knowledge, so that the disagreement we find with regard to the former ceases to matter for the lack of disagreement we find in the latter. For another, realists can invite *comforting comparisons*, by seeking an analogy between everyday non-scientific knowledge

and moral knowledge, such that the agreement we find in the former, combined with our inclination to construe this domain realistically, supports realism about the latter. Thirdly, realists could argue that the *a priori* character of moral knowledge explains how widespread agreement can be compatible with realism. Fourth, realists could object that the argument from disagreement is committed to the purportedly implausible idea of moral expertise. Fifth, realists can launch a counterattack, arguing that the pattern of moral agreement and disagreement we find is exactly what we would expect if moral realism were true. I will take up these options in turn.

Scientific and Moral Knowledge. One may be inclined to argue that the analogy with science is flawed, as it stacks the deck against those who would want to defend realism from the argument from agreement. But notice that the analogy with science is one that is frequently sought by realists themselves (Huemer, 2005). It then seems dialectically unfair to exploit this analogy when it seems to help realism, but reject it when it turns out not to.[14]

However, there *is* a legitimate point here, which is that what should be compared to what to make a relevant analogy (or disanalogy) work is a difficult question. One could say that even physicists and laypeople agree about a lot, namely observational reports (e. g. the red light is blinking, this feather falls at a certain speed in a vacuum under the current experimental conditions). The experts will merely be better able to explain, in more fundamental terms (such as the equivalence of inert and heavy mass), why these observations occur. Considering this analogy, should fundamental values be compared to observational reports – about which the experts and laypeople purportedly agree – or fundamental explanatory principles – about which they do not? The second option seems preferable because it compares like to like. And if we compare fundamental and unifying principles to each other, we find lots of disagreement between experts and laypeople in the sciences but hardly any disagreement between ethical experts and ethical laypeople.

This point brings us back to the earlier discussion of how to draw the distinction between fundamental and superficial (dis)agreements. Above, I suggested sticking to a criterion that disqualifies moral disagreements that can be defused as mere disagreements about a non-moral supervenience base. But if that is the notion of fundamentality assumed here, then the analogy with disagreement in physics seems to break down, because there is no supervenience base for physical properties besides those same physical properties. However, the analogy between ethics and physics (or other domains for that matter) remains instructive, because in both cases, we scrutinize the level of fundamental explanatory truths (the basic laws of nature/the basic moral values) for patterns of agreement and disagreement. In the case of ethics, we arrive at that level by subtracting non-moral and other defusable sources of disagreement, in the case of physics, we do so by looking at the 'Quinean' core of theory construction.

To be sure, ethical experts such as moral philosophers defend lots of strange and counterintuitive views – such as the morality of open borders, or infanticide, or human enhancement – but they do not defend them because they endorse different values. They merely have a better, or at least different, appreciation of what these values entail and how to implement them.

Everyday Knowledge and Moral Knowledge. Another objection to (A1) above could be phrased as such:

(E1) If realism about everyday midsized objects is true, we would expect a lot of disagreement about them.

(E2) We do not see a lot of disagreement about everyday midsized objects.
Therefore,

(E3) Realism about nearby midsized objects is false.

This parody argument suggests that there is another analogy, namely one with knowledge of everyday truths, which purports to show that realism does not always predict disagreement. Clearly, (E1) is false. But then, by analogy, (A1)

---

[14] In a recent paper, Finur Dellsén (2017) shows that whether or not whe should expect agreement or disagreement on a given issue, and what the epistemic implications of it are, is frequently a counterintuitive question. Dellsén argues that even disagreement *between the* experts, and not just between experts and laypeople, supports realism.

seems to be false as well. At the very least, the parody shows that one needs additional premises to get from realism to expected disagreement which might not be plausible in the moral case.[15] Again, however, most contemporary moral realists – remember the targets of my argument narrowed down above – do not think that moral facts are like facts about ordinary midsize objects. They believe that acquiring moral knowledge requires empirical or rational investigation of the kind which is not necessary to gain access to mundane facts about apples and oranges.

Easy Moral Knowledge. One way to argue that (A1) is false is to hold that moral truths are *a priori*. Michael Smith, for instance, writes: 'It is agreed on nearly all sides that moral knowledge is relatively a priori, at least in the following sense: if you equip people with a full description of the circumstances in which someone acts, then they can figure out whether the person acted rightly or wrongly just by thinking about the case at hand' (2000, p. 203). Many others hold similar views: 'Morality does not require beliefs that are not known to all moral agents' (Gert, 2004, p. 90). Or: 'this book [...] contains no new information about what kinds of actions morality prohibits, requires, discourages, encourages, or allows. Anyone who is intelligent enough to read this book already has all of this information' (3). Or: 'Our moral life seems to presuppose that [moral] facts are in principle available to all; that no one in particular is better placed to discover them than anyone else' (Smith, 1994, p. 5). Or: '[i]t seems implausible to say that it would take a 'moral genius' to see through the wrongness of chattel slavery' (Guerrero, 2007, p. 71).[16] Typically, there is not much disagreement on conceptual trivialities.

However, most realists continue to claim that moral truths are unobvious, so at the very least, this reply to my argument does not work for those theories. Moreover, even though I have argued that we do find widespread moral agreement, there is at least *some* moral disagreement. But if the 'easy knowledge' reply is correct, then we would not expect even this much disagreement. There is no disagreement on whether bachelors are married or unmarried. The reply backfires, so realists should seek no comfort in it.

Moral Expertise. The argument from agreement seems to rely on the claim that there could be moral expertise, which is controversial.[17] It may turn out that all the heavy lifting in the argument from disagreement is actually being done by the claim that there can be no one who fundamentally knows better about morality than other people. So it's never been about disagreement in the first place, and always about the idea that no one's fundamental moral views are superior. This would surprise no one who has ever taught relativism to undergraduates, who seem to take offense with non-relativism for precisely this reason: that it seems disrespectful or intolerant to describe someone else's moral views as mistaken.

Counterattack. Some realists may claim that we find a certain pattern – a mix, as it were – of agreement and disagreement, and that the pattern we find can comfortably be accounted for by realism. Consider Brink (1984). Brink may argue first that there's more agreement than Mackie and others allow, and that the rest can be explained by appealing to distorting factors via the usual defusing explanations. Some moral truths are easy and others are tough to know, just like some truths about the natural world are easily accessible and others are not. This predicts a pattern of agreement and disagreement, which the realist can maintain is what we observe.[18] However, this line of reply overlooks the fact that what we find is strong agreement about the *basic principles* of morality. If the analogy with science works, then we should not expect this (notice the huge disagreement between physicists and the folk on the fundamental laws of nature). So the pattern we in fact observe remains problematic for realism.

Three final points. Disagreements about how to draw the circle of moral concern seem to be of fundamental importance. But are they? Is it fundamental disagreement in the metaethically interesting sense that some people think that certain groups (Jewish people, people of color, other minorities) do not deserve full moral consideration? I suggest that it is not, because these are not cases of fundamental disagreements about *basic values* (i.e.

---

[15] Thanks to Antti Kauppinen for discussion on this point.

[16] Many of these quotations were compiled by McGrath (2011).

[17] Thanks to Jonathan Webber for helpful suggestions on this point.

[18] Thanks to Antti Kauppinen again for helping me address this issue.

some people thinking suffering is bad and others do not), but merely about membership. Criteria of membership are among the *normatively* most urgent cases of disagreement. Metaethically, they are less interesting.

Perhaps realism about a given domain predicts neither agreement nor disagreement. The only thing it predicts is *convergence over time*. That is to say that whatever degree of disagreement there is in the beginning, realism predicts that people's disagreements will decrease, i.e. that they converge on the truths in a domain. However, the suggestion that realism makes no predictions whatsoever about how much initial disagreement – disagreement before the dynamic of convergence kicked off by moral inquiry has even started, so to speak – there should be is misleading. Realism does predict high degrees of 'primordial' disagreement, because where there are objective facts, there are multiple ways of getting it wrong. It's incorrect that realism merely predicts a dynamic towards convergence but takes no stance on how much disagreement there is initially.

Finally, it is likely that almost all the convergence we actually observe is not domain-specific. It occurs because people's circumstances come to resemble each other, or because people converge on the non-moral facts, or for other reasons. Look at what explains the developments realists like to explain in terms of domain-specific moral convergence, i.e. developments towards liberalization, or increasing resistance to violence and discrimination (see Huemer, 2016; Pinker, 2011). In almost all of these cases, these increases in agreement are not due to a better appreciation of mind-independent moral facts, but other, non-rational (though perhaps very desirable) forces such as upheavals in social organization.

6 | CONCLUSION

If I am right, then there is almost no fundamental moral disagreement. But somehow, this fact feels unsatisfying, as if it were supposed to show that everything is fine, when evidently it is not. It may be that ultimately, my argument shows that fundamental moral disagreement of the kind metaethicists are interested in is of almost no consequence for normative matters.

Perhaps *non-fundamental* moral disagreement has always been the interesting problem. Why is it that people share their basic values, but still end up disagreeing so much about how to apply these values to the world, or to the question of how to live? This, I think, is where a lot of valuable work remains to be done. This work will inevitably have to be deeply empirically informed, and investigate the psychological and institutional mechanisms that lead to tension, conflict, and disengagement. The insight that people agree on basic values such as justice, cooperation or welfare does not help us adjudicate between competing interests in the Israel/Palestine conflict, and it does not show us how to reconcile the perspectives of those who stand on opposite sides on abortion, health care, or immigration.

ORCID

Hanno Sauer https://orcid.org/0000-0001-8039-0376

REFERENCES

Alfano, M. (2016). *Moral psychology: An introduction*. Cambridge, UK: Polity.

Barber, A. (2013). Science's immunity to moral refutation. *Australasian Journal of Philosophy*, *91*(4), 633–653.

Brink, D. O. (1984). Moral realism and the sceptical arguments from disagreement and queerness. *Australasian Journal of Philosophy*, *62*(2), 111–125.

Cuneo, T., & Shafer-Landau, R. (2014). The moral fixed points: New directions for moral nonnaturalism. *Philosophical Studies*, *171*(3), 399–443.

Curry, O. S. (2016). Morality as cooperation: A problem-centred approach. In *The evolution of morality* (pp. 27–51). Cham: Springer International Publishing.

Dellsén, F. (2017). When expert disagreement supports the consensus. *Australasian Journal of Philosophy*, *96*(1), 142–156.

Doris, J., & Plakias, A. (2008). How to argue about disagreement: Evaluative diversity and moral realism. In W. Sinnott-Armstrong (Ed.), *Moral psychology volume 2, the cognitive science of morality: Intuition and diversity* (pp. 303–331). Cambridge, MA: MIT Press.

Enoch, D. (2009). How is moral disagreement a problem for realism? *The Journal of ethics*, *13*(1), 15–50.

Enoch, D. (2011). *Taking morality seriously: A defense of robust realism*. Oxford, UK: Oxford University Press.

Enoch, D. (2014). A defense of moral deference. *The Journal of Philosophy*, *111*(5), 229–258.

Fitzpatrick, S. (2014). Moral realism, moral disagreement, and moral psychology. *Philosophical Papers*, *43*(2), 161–190.

Fraser, B., & Hauser, M. (2010). The argument from disagreement and the role of cross-cultural empirical data. *Mind and Language*, *25*(5), 541–560.

Fricker, M. (2013). Styles of moral relativism. A critical family tree. In R. Crisp (Ed.), *The Oxford handbook of the history of ethics* (pp. 793–818). Oxford, UK: Oxford University Press.

Gert, B. (2004). *Common morality: Deciding what to do*. Oxford, UK: Oxford University Press.

Gill, M. B. (2007). Moral rationalism vs. moral sentimentalism: Is morality more like math or beauty? *Philosophy Compass*, *2*(1), 16–30.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, *96*(5), 1029.

Guerrero, A. A. (2007). Don't know, don't kill: Moral ignorance, culpability, and caution. *Philosophical Studies*, *136*(1), 59–97.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Vintage.

Harman, G. (1975). Moral relativism defended. *The Philosophical Review*, *84*(2), 3–22.

Heath, J. (2009). Three evolutionary precursors to morality. *Dialogue: Canadian Philosophical Review/Revue canadienne de philosophie*, *48*(4), 717–752.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world. *Behavioral and Brain Sciences*, *33*(2–3), 61–83.

Huemer, M. (2005). *Ethical intuitionism*. Basingstoke, UK: Palgrave Macmillan.

Huemer, M. (2008). Revisionary intuitionism. *Social Philosophy and Policy*, *25*(1), 368–392.

Huemer, M. (2016). A liberal realist answer to debunking skeptics: The empirical case for realism. *Philosophical Studies*, *173*(7), 1983–2010.

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PLoS ONE*, *7*(8), e42366.

Kahane, G. (2013). Must metaethical realism make a semantic claim? *Journal of Moral Philosophy*, *10*(2), 148–178.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, *76*(4), 728–741.

Leeson, P. T. (2013). Gypsy law. *Public Choice*, *155*(3–4), 273–292.

Leeson, P. T. (2017). *WTF?!: An economic tour of the weird*. Palo Alto, CA: Stanford University Press.

Leiter, B. (2008). Against convergent moral realism: The respective roles of philosophical argument and empirical evidence. In W. Sinnott-Armstrong (Ed.), *Moral psychology volume 2, the cognitive science of morality: Intuition and diversity* (pp. 333–337). Cambridge, MA: MIT Press.

Mackie, J. L. (1977). *Ethics: Inventing right and wrong*. Harmondsworth, UK: Penguin Books.

McCloskey, M., Washburn, A., & Felch, L. (1983). Intuitive physics: The straight-down belief and its origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 636.

McGrath, S. (2008). Moral disagreement and moral expertise. *Oxford Studies in Metaethics*, *3*(87), 107.

McGrath, S. (2011). Skepticism about moral expertise as a puzzle for moral realism. *The Journal of Philosophy*, *108*(3), 111–137.

Medin, D. L., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological review*, *111*(4), 960.

Meyers, C. D. (2013). Defending moral realism from empirical evidence of disagreement. *Social Theory and Practice*, *39*(3), 373–396.

Moody-Adams, M. M. (2009). *Fieldwork in familiar places: Morality, culture, and philosophy*. Cambridge, MA: Harvard University Press.

Pinker, S. (2011). *The better angels of our nature: The decline of violence in history and its causes*. London, UK: Penguin.

Prinz, J. (2007). *The emotional construction of morals*. Oxford, UK: Oxford University Press.

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological review*, *118*(1), 57.

Rowland, R. (2017). The significance of significant fundamental moral disagreement. *Noûs*, *51*, 802–831.

Sauer, H. (2011). The appropriateness of emotions. Moral judgment, moral emotions, and the conflation problem. *Ethical Perspectives*, *18*(1), 107–140.

Sauer, H. (2015). Can't we all disagree more constructively? Moral foundations, moral reasoning, and political disagreement. *Neuroethics*, *8*(2), 153–169.

Sauer, H. (2017). *Moral judgments as educated intuitions*. Cambridge, MA: MIT Press.

Sauer, H. (2018a). *Debunking arguments in ethics*. Cambridge: Cambridge University Press.

Sauer, H. (2018b). No excuses for moral realism. *Philosophical Studies*, *175*(3), 553–578.

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues*, *50*(4), 19–45.

Shafer-Landau, R. (2003). *Moral realism: A defence*. Oxford, UK: Oxford University Press.

Smith, M. (2000).Moral realism. In H. LaFollette (Ed.), *The Blackwell guide to ethical theory* (pp. 17–43). Oxford, UK: Blackwell.

Sneddon, A. (2009). Normative ethics and the prospects of an empirical contribution to assessment of moral disagreement and moral realism. *The Journal of Value Inquiry*, *43*(4), 447–455.

Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies*, *127*(1), 109–166.

Tropman, E. (2012). Can Cornell moral realism adequately account for moral knowledge? *Theoria*, *78*(1), 26–46.

Tropman, E. (2014). Why Cornell moral realism cannot provide an adequate account of moral knowledge. *Theoria*, *80*, 184–190.

Wong, D. (2006). *Natural moralities. A defense of pluralistic relativism*. Oxford, UK: Oxford University Press.