

The epistemic argument for hedonism

Neil Sinhababu - neiladri@gmail.com - National University of Singapore

This paper argues that we should give up all of our moral beliefs, except for belief in pleasure's goodness, and anything further that it entails. ("Pleasure's goodness" will here also refer to displeasure's badness.) The resulting position is a version of ethical hedonism on which we doubt rather than reject nonhedonic moral facts.¹ Many regard hedonism as counterintuitive, and I agree. So I cannot claim that hedonism best fits our intuitions about general principles and concrete cases, and defend it using reflective equilibrium.²

Instead, I will first present an argument for moral skepticism and then show that hedonism has a unique way of refuting it. The first half of this paper presents the skeptical argument, which arises from the breadth of moral disagreement. Widespread disagreement requires us to posit so much false moral belief that we must regard the processes by which we form moral beliefs as unacceptably error-prone, and abandon our moral beliefs. The second half of this paper argues that belief in the goodness of pleasure escapes skeptical defeat. We know that pleasure is good through phenomenal introspection, which reliably produces true belief. If introspecting pleasure's goodness is the only answer to the skeptical argument, our only moral beliefs should be in pleasure's goodness and whatever it entails.

1.1 The argument from disagreement

¹ According to ethical hedonism, pleasure's goodness and whatever follows from it are the only moral facts. Ethical hedonism combined with suitable aggregative and consequentialist principles entails hedonic utilitarianism.

² Many cases suggest that hedonism is counterintuitive – for example, the Experience Machine from Robert Nozick, *Anarchy State and Utopia*, New York: Basic Books (1974), 42-45. Many other objections are canvassed in Chapter 3 of Fred Feldman, *Pleasure and the Good Life*, New York: Oxford University Press (2004).

The skeptical argument I offer is the *argument from disagreement*. 1-4 are premises:

1. In any moral disagreement, at least one party must be in error.
2. There is widespread moral disagreement.
3. If there is widespread error about a topic, we should retain only those beliefs about it formed through reliable processes.
4. If there is widespread error about morality, there are no reliable processes for forming moral beliefs.
5. There is widespread error about morality (from 1 and 2).
6. We should retain only those moral beliefs formed through reliable processes (from 3 and 5).
7. There are no reliable processes for forming moral beliefs (from 4 and 5).
8. We should give up all of our moral beliefs (from 6 and 7).

This section briefly explains how the argument works, and distinguishes it from other recent arguments involving disagreement. The next three sections defend 2-4 at length.

1 lays out how moral disagreement must work – at least one party must be in error. The objectivity of morality entails this. Parties to any moral disagreement assert genuinely contradictory propositions, some of which must be false. Noncognitivists and constructivists might deny 1, claiming that moral disagreement need not involve error on any side. I cannot do their views justice here. I will just assume that noncognitivist and constructivist construals of moral concepts are mistaken, and any genuine moral facts must be objective.³

2, which I will defend in section 1.2, says that there is widespread moral disagreement. 2 refers only to disagreement about the fundamental moral questions, not disagreement that

arises from agreement on these questions combined with differing opinions on nonmoral

³ Noncognitivism can account for inconsistency only by giving a flawed account of nonmoral discourse, as argued in Mark Schroeder, *Being For*, Oxford: Oxford University Press (2008). For an argument that noncognitivism cannot make sense of moral disagreement, see David Merli, “Expressivism and the Limits of Moral Disagreement,” *The Journal of Ethics* 12 (2008): 25-55.

matters. Disagreement arising from differing nonmoral beliefs does not advance the argument, as it does not bear on the reliability of our processes for forming moral beliefs. 1 and 2 imply 5, that there is widespread moral error. Since parties to a moral disagreement contradict each other about an objective fact, each instance of disagreement must involve false belief. Disagreement between many parties about many things, then, must involve many false beliefs about many things.

What should we do if there is widespread error about some topic, as 5 suggests? 3, which I will defend in section 1.3, describes how to proceed. The processes driving belief-formation on the topic leave people at a high risk of having false beliefs. If we hold beliefs on the topic and the processes that created them are of average reliability for the topic, they run a high risk of being false. If this risk is high enough, we should abandon our beliefs. But we need not abandon our beliefs if they can be supported by an unusually reliable process of belief-formation.

4, which I will partially defend in section 1.4, says that if there is widespread error about morality, there is no reliable process for forming moral beliefs. This is merely a material conditional, not a necessary truth. There is nothing inconsistent about widespread error coexisting with reliable processes. (In fact, the second half of this paper answers the skeptical argument by presenting what seems to be the only actual counterexample to 4 – phenomenal introspection, which reliably informs us of pleasure's moral value.) But widespread error implies that any one process through which we form all or nearly all of our moral beliefs must be error-prone enough to be unreliable, since it must have created a large share of the widespread error. 4 can still be false if some process which produces only a small fraction of our moral beliefs is reliable, allowing us to avoid the unreliability of our other ways of forming moral beliefs. To defend 4, then, one must argue separately against the reliability of every such process. Section 1.4 argues against many of the most-discussed contenders. 4 and 5 together

imply 7, the absence of any reliable processes of belief formation. When 6, which only allows moral beliefs formed through reliable processes, is combined with 7, we arrive at 8 and must abandon all our moral beliefs.

The argument from disagreement avoids unfortunate consequences of John Mackie's argument from relativity, the most famous argument involving disagreement in contemporary metaethics. The importance of disagreement for Mackie seems to be that moral belief is not best explained by objective moral value, since there are so many differing views. He writes that "the actual variations in the moral codes are more readily explained by the hypothesis that they reflect ways of life than by the hypothesis that they express perceptions, most of them seriously inadequate and badly distorted, of objective values" (37).⁴ Unlike Mackie's argument, the argument from disagreement does not suggest rejecting morality on grounds of explanatory inadequacy. Explanatory arguments against morality might be implausible or self-undermining if they also force us to reject color, numbers, or epistemic norms, which are similarly hard to place in our best explanations.⁵ The argument from disagreement avoids these problems, as it applies only when widespread disagreement implies so much error that our beliefs must be unreliably formed. As there is more agreement on color, mathematics, and epistemic norms, the argument from disagreement does not force skeptical conclusions there.

Much recent work in epistemology discusses disagreement between epistemic peers, especially disagreement between two people about a single proposition.⁶ This work does not provide a satisfactory model of moral disagreement, which involves many people with

⁴ John Mackie, *Ethics: Inventing Right And Wrong*, London: Penguin (1977). This is how Mackie's argument is understood by David Brink, *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press (1989), 197-209.

⁵ See Geoffrey Sayre-McCord, "Moral Theory and Explanatory Impotence," responding to Gilbert Harman, "Ethics and Observation," both in Geoffrey Sayre-McCord (ed.), *Essays on Moral Realism*, Ithaca: Cornell University Press, 1988.

competing moral theories consisting of many propositions. Even if we should not revise our beliefs because of disagreement with one epistemic peer about one proposition, perhaps we should revise our beliefs because of disagreement with many people about entire theories. The argument from disagreement provides a route to this conclusion, moving from widespread disagreement to widespread error to the unreliability of our processes of belief-formation.⁷

The argument from disagreement presented here is novel in the current literature on moral disagreement. A variety of other arguments from moral disagreement to skepticism or anti-realism have been nicely criticized by David Enoch and Michael Huemer.⁸ But Enoch and Huemer do not respond to arguments that use the presence of error in each moral disagreement to show that widespread moral disagreement involves widespread error, and which then use the presence of widespread moral error to demonstrate that our processes of moral belief-formation are unreliable. The argument presented here is the only one that does this. I hope it will be of independent interest, apart from the broader argument for hedonism.

Richard Joyce and Sharon Street present evolutionary debunking arguments that can support moral skepticism, just as the argument from disagreement does.⁹ They hold that our

⁶ For example, Thomas Kelly, "The Epistemic Significance of Disagreement," in Hawthorne and Szabo (eds.), *Oxford Studies in Epistemology Vol. 1*, Oxford: Oxford University Press (2005); and Adam Elga, "Reflection and Disagreement," *Noûs* 41:3 (2007): 478-502.

⁷ For a helpful discussion of multiparty intertheoretic disagreement that inspires the argument presented here, see Todd Stewart, "The Competing Practices Argument and Self-Defeat," *Episteme* (2005): 13-24.

⁸ David Enoch, "How is Moral Disagreement a Problem for Realism?" *Journal of Ethics* 13 (2009) 15-50; and Michael Huemer, *Ethical Intuitionism*, New York: Palgrave MacMillan (2005).

⁹ Sharon Street, "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127:1 (2006): 109-166, and Richard Joyce, *The Myth of Morality*, Cambridge: Cambridge University Press (2001).

moral beliefs are produced by evolutionary processes that are unlikely to have any systematic correlation with objective moral truth. Seeing this, we should abandon our beliefs in objective moral facts. This evolutionary debunking argument offers another route to the skeptical conclusion of the argument from disagreement. Section 2.1 will outline how hedonists' answer to the argument from disagreement answers the evolutionary debunking argument as well.

1.2 There is widespread moral disagreement

Many philosophers will contend that there is not enough moral disagreement to drive the argument from disagreement forward. Proving otherwise would require a great deal of detailed anthropological and historical work, and is well beyond the scope of what I can accomplish here. Here I can aim only to make the existence of that much moral disagreement plausible enough that readers will regard the argument as worthy of concern. First I will present some evidence that actual moral disagreement is this great. Then I will respond to the most obvious rejoinder – that differing opinions about moral questions really arise from nonmoral disagreement – by arguing that the different nonmoral beliefs are often themselves the products of underlying moral disagreement. Finally, I will argue that whatever moral agreement remains after taking this point into account can be handled in a way that lets the argument from disagreement go forward.

For a vivid illustration of how intensely we disagree about morality, consider a scene from the end of the *Odyssey*. Odysseus has returned home and killed the hundred suitors who sought his wife's hand. He and his son Telemachus bring out twelve of the servant women, and make them assist in cleaning up the dead bodies.

Then when they had made the whole place quite clean and orderly, they took the women out and hemmed them in the narrow space between the wall of the domed room and that of the yard, so that they could not get away: and Telemachus said to the other two, "I shall not

let these women die a clean death, for they were insolent to me and my mother, and used to sleep with the suitors."

So saying he made a ship's cable fast to one of the bearing-posts that supported the roof of the domed room, and secured it all around the building, at a good height, lest any of the women's feet should touch the ground; and as thrushes or doves beat against a net that has been set for them in a thicket just as they were getting to their nest, and a terrible fate awaits them, even so did the women have to put their heads in nooses one after the other and die most miserably. Their feet moved convulsively for a while, but not for very long.¹⁰

The *Odyssey's* treatment of these events demonstrates how ancient Greek moral intuitions differ from ours. It does not dwell on the brutality of Telemachus, who has killed twelve women for the trivial reasons he states, making sure they suffer as they die. While gods and men seek vengeance for many other great and small offenses in the *Odyssey*, no finds this mass murder worth avenging. It is a minor event in the denouement to a happy ending in which Odysseus (who first proposes killing the women) returns home and Telemachus becomes a man. That the Greeks could so easily regard these murders as part of a happy ending for heroes shows how deeply we disagree with them. It is as if we gave the Greeks a trolley problem with the 12 women on the side track and no one on the main track, and they judged it permissible for Telemachus to turn the trolley and kill them all. This all appears not in some esoteric text of a despised or short-lived sect, but one of the central works of a long-lived and influential culture.

Human history offers similarly striking examples of disagreement on a variety of topics. These include sexual morality; the treatment of animals; the treatment of other ethnicities, families, and social classes; the consumption of intoxicating substances; whether and how one may take vengeance; slavery; whether public celebrations are acceptable; and gender roles.¹¹

One can only view the human past and much of the present with horror at the depth of human

¹⁰ Homer, *The Odyssey*, Butler (trans.), (1900): <http://classics.mit.edu/Homer/odyssey.html>.

Readers of mythology will easily find similar examples from other ancient texts.

moral error and the harm that has resulted.

One might think to explain much of this disagreement away as the result of differing nonmoral beliefs. Those who disagree about nonmoral issues may disagree on the moral rightness of a particular action despite agreeing on the fundamental moral issues. For example, they may agree that it is right to heal the sick, but disagree about whether a particular medicine will heal or harm. This disagreement about whether it is right to prescribe the medicine will not be fundamentally about morality, and will not support the argument from disagreement.

I do not think the moral disagreements listed above are explained by differences in nonmoral belief. This is not because sexists, racists, and bigots share the nonmoral views of those enlightened by feminism and similar egalitarian doctrines – they do not. Rather, their differing views on nonmoral topics often are rationalizations of moral beliefs that fundamentally disagree with ours.¹² Those whose fundamental moral judgments include commitments to the authority of men over women, or of one race over another, will find it easy to accept descriptive psychological views that attribute less intelligence or rationality to women or the subjugated race.¹³ Moral disagreement supposedly arising from moral views in religious

¹¹ See John Doris and Alexandra Plakias, “How to Argue about Disagreement: Evaluative Diversity and Moral Realism,” in Sinnott-Armstrong (ed.), *Moral Psychology Vol. 2*, Cambridge MA: MIT Press (2008): 303-331.

¹² For a good discussion, see Nomy Arpaly, *Unprincipled Virtue*, New York: Oxford University Press (2003). For striking psychological evidence about rationalization, see Jonathan Haidt, “The Emotional Dog and the Rational Tail”, *Psychological Review* 108:4 (2001): 814-834.

¹³ As system justification theorists in psychology argue, even subjugated groups form a structure of beliefs according to which their subjugation is justified. John Jost, Mahzarin Banaji, and Brian Nosek, “A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo”, *Political Psychology* 25:5 (2004), 881-919.

texts is similar. Given how rich and many-stranded most religious texts are, interpretive claims about their moral teachings often tell us more about the antecedent moral beliefs of the interpreter than about the text itself. This is why the same texts are interpreted to support a wide variety of differing moral views. Similar phenomena occur with most moral beliefs. Environmentalists who value a lovely patch of wilderness will easily believe that disaster will result from its destruction, those who feel justified in eating meat will easily believe that the animals they eat do not suffer greatly, and libertarians who feel that redistributing wealth is unjust will easily believe that it raises unemployment. We should not assume that differing moral beliefs on practical questions result from fundamental moral agreement combined with differing nonmoral beliefs. They often result from fundamental moral disagreements generating the differing nonmoral beliefs, as 2 claims.

As we have no precise way of quantifying the breadth of disagreement or determining its epistemic consequences, it is unclear exactly how much disagreement the argument requires. While this makes the argument difficult to evaluate, it should not stop us from proceeding, as the unclear notion of widespread disagreement is one that we must employ in ordinary epistemic practice. If 99.9% of botanists agree on some issue about plants, non-botanists should defer to their authority and believe as the vast majority does. But if disagreement between botanists is suitably widespread, non-botanists should remain agnostic. It would help to have a more precise and systematic account of when disagreement is widespread enough to generate particular epistemic consequences. Until we do, we must employ the unclear notion of widespread disagreement, or some similar notion, throughout epistemic practice.

Against the background of widespread moral disagreement, there may still be universal or near-universal agreement on some moral questions. For example, perhaps all cultures agree that one should provide for one's elderly parents, even though they generally disagree elsewhere. How do these narrow areas of moral agreement affect the argument?

Narrow agreement can defeat the argument from disagreement if it results from a reliable process of belief-formation that lets us avoid error. But we should beware of retaining agreed-upon moral beliefs produced by powerful and unreliable processes that lead everyone to the same errors. While there might be pressure to explain agreement in terms of reliable processes in some cases, this does not hold when disagreement is widespread. Explaining agreement in terms of reliable processes would be preferable if moral beliefs were generally true. Then we would want to understand cases of agreement in line with the general reliability of processes producing moral belief. But if disagreement and thus error are widespread, we will be more open to explaining agreement in terms of unreliable processes, as moral beliefs are so often formed through processes that lead to error.

We have many plausible explanations of narrow agreement on which moral beliefs are unreliably caused. Evolutionary and sociological explanations of why particular moral beliefs are widely accepted often invoke unreliable mechanisms.¹⁴ According to these explanations, agreement results from some moral beliefs having been so essential to reproductive fitness that natural selection made them innate in us, or so important to the interests that control moral education in each human culture that they were inculcated in everyone. For example, parental influence over their children's moral education would explain agreement that one ought to provide for one's elderly parents. On any plausible normative ethical theory, these evolutionary and sociological explanations will not systematically connect moral beliefs with the moral facts. If there is widespread error, they will gain plausibility, allowing us to reconcile unusual cases of widespread agreement with the general unreliability of the processes producing moral belief.

¹⁴ For evolutionary explanations, see Street (2006) and Joyce (2001). For classic sociological explanations, see Michel Foucault, *Discipline and Punish: The Birth of the Prison*, Sheridan-Smith (trans.), New York: Vintage (1977); and Friedrich Nietzsche, *On the Genealogy of Morality*, Clark and Swensen (trans.), Indianapolis: Hackett, 1998.

1.3 If there is widespread error about a topic, we should retain only those beliefs about it formed through reliable processes

Now I will defend 3. First, I will show how the falsity of others' beliefs undermines one's own belief. Then I will clarify the notion of a reliable process. I will consider a modification to 3 that epistemic internalists might favor, and show that the argument accommodates it. I will illustrate 3's plausibility by considering cases in which it correctly guides our reasoning. Finally, I will show how 3 is grounded in the intuitive response to grave moral error.

First, a simple objection: "Why should I care whether other people have false beliefs? That is a fact about other people, and not about me. Even if most people are wrong about some topic, I may be one of the few right ones, even if there is no apparent reason to think that my way of forming beliefs is any more reliable."

While widespread error leaves open the possibility that one has true beliefs, it reduces the probability that my beliefs are true. Consider a parallel case. I have no direct evidence that I have an appendix, but I know that previous investigations have revealed appendixes in people. So induction supports believing that I have an appendix. Similarly, I know on the basis of 1 and 2 that people's moral beliefs are, in general, rife with error. So even if I have no direct evidence of error in my moral beliefs, induction suggests that they are rife with error as well.

3 invokes the reliability of the processes that produce our beliefs. Assessing processes of belief-formation for reliability is an essential part of our epistemic practices. If someone tells me that my belief is entirely produced by wishful thinking, I cannot simply accept that and maintain the belief. Knowing that wishful thinking is unreliable, I must either deny that my belief is entirely caused by wishful thinking or abandon the belief. But if someone tells me that my belief is entirely the result of visual perception, I will maintain it, assuming that it concerns sizable nearby objects or some other topic on which visual perception is reliable. While it is hard to provide precise criteria for individuating processes of belief-formation, as the literature on

the generality problem for reliabilism attests, individuating them somehow is indispensable to our epistemic practices.¹⁵ Following Alvin Goldman's remark that "It is clear that our ordinary thought about process types slices them broadly" (346), I will treat cognitive process types like wishful thinking and visual perception as appropriately broad.¹⁶ Trusting particular people and texts, meanwhile, are too narrow. Cognitive science may eventually help us better individuate cognitive process types for the purposes of reliability assessments and discover which processes produce which beliefs.

Epistemic internalists might reject 3 as stated, claiming that it is not widespread error that would justify giving up our beliefs, but our having reason to believe that there is widespread error. They might also claim that our justification for believing the outputs of some process depends not on its reliability, but on what we have reason to believe about its reliability. The argument will still go forward if 3 is modified to suit internalist tastes, changing its antecedent to "If we have reason to believe that there is widespread error about a topic" and/or changing its consequent to "we should retain only those beliefs about it that we have reason to believe were formed through reliable processes." While 3's antecedent might itself seem unnecessary on the original formulation, it is required for 3 to remain plausible on the internalist modification. Requiring us to have reason to believe that any of our belief-formation processes are reliable before retaining their outputs might lead to skepticism. The antecedent limits the scope of the requirement to cases of widespread error, averting general skeptical conclusions. The argument will still attain its conclusion under these modifications. Successfully

¹⁵ Earl Conee and Richard Feldman, "The Generality Problem for Reliabilism," *Philosophical Studies* 89 (1998):1-29. The generality problem seems to apply to all plausible epistemological views, as argued by Juan Comesaña, "A Well-Founded Solution to the Generality Problem," *Philosophical Studies* 129 (2006): 27-47.

¹⁶ Alvin Goldman, "What is Justified Belief?" in Sosa and Kim (eds.), *Epistemology: An Anthology*, Massachusetts: Blackwell (2000).

defending the premises of the argument and deriving widespread error (5) and unreliability (7) gives those of us who have heard the defense and derivation reason to believe 5 and 7. This allows us to derive 8. (Thus the pronoun 'we' in 3, 6, and 8.)

3 describes the right response to widespread error in many actual cases. Someone in the 12th century, especially upon hearing the disagreeing views of many cultures regarding the origins of the universe, would do well to recognize that error on this topic was widespread and retreat to agnosticism about it. Only when modern astrophysics extended reliable empirical methods to cosmology would it be rational to move forward from agnosticism and accept a particular account of how the universe began. Similarly, there is usually widespread disagreement among investors about which stocks will perform better than average, suggesting that one's beliefs on the matter have a high likelihood of error. It is wise to remain agnostic about the stock market without a way of forming beliefs with above-average reliability – for example, the sort of secret insider information that it is illegal to trade on.

3 permits us to hold fast to our moral beliefs in individual cases of moral disagreement, suggesting skeptical conclusions only when moral disagreement is widespread. When we consider a single culture's abhorrent moral views, like the Greeks' acceptance of Telemachus and Odysseus' murders of the servant women, our immediate thought is not that perhaps the Greeks were right to see nothing wrong and we should reconsider our outrage. Instead, we are horrified by their grave moral error. I think this is the right response. We are similarly horrified by the moral errors of Hindus who burned widows on their husbands' funeral pyres, American Southerners who supported slavery and segregation, our contemporaries who condemn homosexuality, and countless others. The sheer number of cases like this then requires us to regard moral error as a pervasive feature of the human condition. Humans typically form moral beliefs through unreliable processes and have appendixes. As we are human beings, this should reduce our confidence in our moral judgments. The prevalence of error in a world full of moral

disagreement demonstrates how bad humans are at forming true moral beliefs, undermining our own moral beliefs. Wary of the unreliable processes that so often lead humans to their moral beliefs, we will need our moral beliefs to issue from reliable processes.

1.4 If there is widespread error about morality, there are no reliable processes for forming moral beliefs

A reliable process for forming moral beliefs would avert skeptical conclusions. I will consider several processes and argue that they do not help us escape moral skepticism. Ordinary moral intuition, whether it involves a special rational faculty or our emotional responses, is shown to be unreliable by the existence of widespread error. The argument from disagreement either prevents reflective equilibrium from generating moral conclusions or undermines it. Conceptual analysis is reliable, but delivers the wrong kind of knowledge to avert skepticism. If all our processes for forming moral beliefs are unreliable, moral skepticism looms. 4 is false only because of one process – phenomenal introspection, which allows us to know of the goodness of pleasure, as the second half of this paper will discuss.

Widespread error guarantees the unreliability of any process by which we form all or almost all of our moral beliefs. While widespread error allows some processes responsible for a small share of our moral beliefs to predominantly create true beliefs, it implies that any process generating a very large share of moral belief must be highly error-prone. Since the process produced so many of our moral beliefs, and so many of them are erroneous, it must be responsible for a large share of the error. If more of people's moral beliefs were true, things would be otherwise. Widespread truth would support the reliability of any process that produced most or all of our moral beliefs, since that process would be responsible for so much true belief. But given widespread error, ordinary moral intuition must be unreliable. This point provides a forceful response to Moorean opponents who insist that we can't give up the

reliability of a process by which we form all or nearly all of our beliefs on an important topic, since this would open the door to counterintuitive skeptical conclusions. Even if this Moorean response works against external world skeptics who employ counterfactual thought experiments involving brains in vats, it doesn't work against moral skeptics who use 1 and 2 to derive widespread actual error. If we accept that widespread error actually obtains, we have already been forced to a counterintuitive skeptical conclusion. Insisting on the reliability of the process will then be both implausible and pointless. I will briefly consider two conceptions of moral intuition – as a special rational faculty by which we grasp non-natural moral facts, and as a process by which our emotions lead us to form moral beliefs – and show how widespread error guarantees their unreliability.

Some philosophers regard moral intuition as a process in which we use a special rational faculty to know non-natural moral facts.¹⁷ They argue that knowledge on many topics including mathematics, logic, and modality involves this rational faculty, so moral knowledge might operate similarly. This suggests a way they might defend the reliability of moral intuition in the face of widespread error. Assuming that intuition truly is reliable in these other areas, they might claim, its overall reliability across moral and nonmoral areas allows us to reliably form moral beliefs by using it. This defense should be rejected. When an epistemic process is manifestly unreliable on some topic, as widespread error shows any process responsible for most of our moral beliefs to be, the reliability of that process elsewhere cannot save it on that topic. Even if testimony is generally a reliable way of forming beliefs, it is not generally a reliable way of forming beliefs about the next spin of the roulette wheel. The general reliability of testimony does not imply the reliability of a compulsive gambler's testimony on such a topic. Even if intuition remains reliable elsewhere, widespread disagreement still renders it unreliable in ethics.

Following recent psychological research, I regard ordinary moral intuition as a process

¹⁷ For example, Huemer (2005).

of emotional perception in which we form moral beliefs on the basis of our emotional responses.¹⁸ This is like how we form beliefs about the colors of objects on the basis of visual experience. Pleasant emotions like approval, admiration, or hope as we consider actions, persons, or states of affairs lead us to believe they are right, virtuous or good. Unpleasant emotions like guilt, disgust, or horror as we consider actions, persons, or states of affairs lead us to believe they are wrong, vicious, or bad. We might have regarded this as a reliable empirical means for discovering moral facts, like visual perception of color, if not for widespread error. But because of widespread error, we can only see it as an unreliable process responsible for our dismal epistemic situation.

Reflective equilibrium is the prevailing methodology in normative ethics today. It involves seeking maximal coherence between our beliefs about particular cases and general principles. Whether or not nonmoral propositions like the premises of the argument from disagreement are admissible in reflective equilibrium, widespread error prevents reflective equilibrium from reliably generating a true moral theory, as I will explain.

If the premises of the argument from disagreement are admitted into reflective equilibrium, the argument can be reconstructed there, and reflective equilibrium will dictate

¹⁸ For psychological research describing how emotion produces moral judgments, see Simone Schnall *et al*, "Disgust as Embodied Moral Judgment", *Personality and Social Psychology Bulletin* 34:8 (2008): 1096-1109; Simone Schnall, Jennifer Benton, and Sophie Harvey, "With a Clean Conscience: Cleanliness Reduces the Severity of Moral Judgments", *Psychological Science* 19:12 (2008): 1219-1222; and Haight (2001). For research on psychopaths demonstrating their failure to properly form moral judgments in the absence of emotion, see Hervey Cleckley, *The Mask of Sanity*, Augusta GA: Emily S. Cleckley (1988); RJR Blair, "A cognitive developmental approach to morality: investigating the psychopath," *Cognition* 57 (1995): 1-29; and Kent Kiehl, "A cognitive neuroscience perspective on psychopathy: Evidence for paralimbic system dysfunction," *Psychiatry Research* 142:2 (2006): 107-128.

that we give up all of our moral beliefs. To avoid this conclusion, the premises of the argument from disagreement would have to be revised away on moral grounds. These premises are a metaethical claim about the objectivity of morality which seems to be a conceptual truth, an anthropological claim about the existence of disagreement, a very general epistemic claim about when we should revise our beliefs, and a more empirically grounded epistemic claim about the processes of belief-formation that we have discovered and their reliability. While reflective equilibrium may move us to revise substantive moral beliefs in view of other substantive moral beliefs, claims of these other kinds are less amenable to such revision. Unless ambitious arguments for revising these claims away succeed, we must follow the argument to its conclusion and accept that reflective equilibrium makes moral skeptics of us.

If only moral principles and judgments are considered in reflective equilibrium, it will not make moral skeptics of us, but the argument from disagreement will undermine its conclusions. This is because the argument suggests giving up the pre-existing moral beliefs against which we test various moral propositions in reflective equilibrium. While we may be justified in believing something because it coheres with our other beliefs, this justification goes away once we see that those beliefs should be abandoned. Coherence with beliefs that we know we should give up does not confer justification.

Now I turn to conceptual analysis, which provides us with some moral beliefs – for example, that the moral supervenes on the nonmoral, and that morality is objective. It also may provide judgments about relations between different moral concepts – perhaps, that if the only moral difference between two actions is that one would produce morally better consequences than the other, it is right to do what produces better consequences. I regard conceptual analysis as reliable, so that the argument from disagreement does not force us to give up the beliefs about morality it produces. Unfortunately, if analytic naturalism is false, as has been widely held in metaethics since G. E. Moore, conceptual analysis cannot provide all the knowledge we

need to build a normative ethical theory.¹⁹ It provides only very general moral truths. Even when it relates moral concepts like goodness and rightness to each other, it does not tell us that anything is good or right to begin with. We need that kind of knowledge to avoid global moral skepticism.

So far I have argued that our epistemic and anthropological situation, combined with plausible metaethical and epistemic principles, forces us to abandon our moral beliefs. But if a reliable process of moral belief-formation exists, 4 is false, and we can retain some of our moral beliefs. Now I will discuss the only such reliable process of which I know.

2.1 Phenomenal introspection reveals pleasure's goodness

Phenomenal introspection, a reliable way of forming true beliefs about our experiences, tells us that pleasure is good and displeasure is bad. Even as our other processes of moral belief-formation prove unreliable, it provides reliable access to pleasure's goodness, justifying the positive claims of hedonism. This section clarifies what phenomenal introspection and pleasure are, and explains how phenomenal introspection provides reliable access to pleasure's value. Section 2.2 argues that pleasure's goodness is genuine moral value, rather than value of some other kind.

To use phenomenal introspection is to look inward at one's subjective experience, or phenomenology, and determine what it is like. One can use phenomenal introspection reliably while dreaming or hallucinating, as long as one can determine what the dream or hallucination is like. By itself, phenomenal introspection produces no beliefs about things outside experience, or about relations between our experiences and non-experiential things. It cannot by itself produce judgments about the rightness of actions or the goodness of non-experiential things, as

¹⁹ G. E. Moore, *Principia Ethica*, Cambridge: Cambridge University Press (1903). For dissent, see Frank Jackson, *From Metaphysics to Ethics*, Oxford: Oxford University Press (1998). Even if Jackson is right, moral concepts may be complex enough that epistemic difficulties arise again.

these are located outside of experience.

Phenomenal introspection can be wrong, but is generally reliable. As experience is rich in detail, one could get some of the details wrong in one's belief. Under adverse conditions when one has false expectations about what one's experiences will be, or when one is in an extreme emotional state, one might make larger errors. Paradigmatically reliable processes like vision share these failings. Vision sometimes produces false beliefs under adverse conditions, or when we are looking at complex things. It is, nevertheless, fairly reliable. The view that phenomenal introspection is unreliable about our phenomenal states is about as radical as skepticism about the reliability of vision. While contemporary psychologists reject introspection into one's motivations and other causal processes as unreliable, phenomenal introspection fares better. Daniel Kahneman, for example, writes that "experienced utility is best measured by moment-based methods that assess the experience of the present."²⁰ Even those most skeptical about the reliability of phenomenal introspection, like Eric Schwitzgebel, concede that we can reliably introspect whether we are in serious pain.²¹ Then we should be able to introspectively determine what pain is like. I assume the reliability of phenomenal introspection in what follows.

One can form a variety of beliefs using phenomenal introspection. For example, one can believe that one is having sound experiences of particular noises and visual experiences of different shades of color. When looking at a lemon and considering the phenomenal states that are yellow experiences, one can form some beliefs about their intrinsic features – for example, that they are bright experiences. And when considering experiences of pleasure, one can make

²⁰ Daniel Kahneman, "Experienced Utility and Objective Happiness: A Moment-Based Approach," in Kahneman and Tversky (eds.), *Choices, Values and Frames*, New York: Cambridge University Press (2000).

²¹ Eric Schwitzgebel, "The Unreliability of Naive Introspection," *Philosophical Review* 117:2 (2008) 245-274.

some judgments about their intrinsic features – for example, that they are good experiences. Just as one can look inward at one's experience of lemon yellow and appreciate its brightness, one can look inward at one's experience of pleasure and appreciate its goodness.²² When I consider a situation of increasing pleasure, I can form the belief that things are better than they were before, in the same way I form the belief that there is more brightness in my visual field as lemon yellow replaces black. And when I suddenly experience pain, I can form the belief that things are worse in my experience than they were before.

"Pleasure" here refers to the hedonic tone of experience. Having pleasure consists in one's experience having this hedonic tone. Without descending into metaphor, it is hard to give a further account of what pleasure is like than to say that when one has it, one feels good. As Aaron Smuts writes in defending the view of pleasure as hedonic tone, "to 'feel good' is about as close to an experiential primitive as we get."²³ Some philosophers, like Fred Feldman, see pleasure as fundamentally an attitude rather than a hedonic tone.²⁴ But as long as hedonic tones – good and bad feelings – are real components of experience, phenomenal introspection will reveal pleasure's goodness. Opponents of the hedonic tone account of pleasure usually concede that hedonic tones exist, as Feldman seems to in discussing "sensory pleasures," which he thinks his view helps us understand. Even on his view of pleasure, phenomenal introspection can produce the belief that some hedonic tones are good while others are bad.

There are many different kinds of pleasant experiences. There are sensory pleasures, like

²² Occasionally I doubt that pleasure is really good, and need to test this claim. So I eat or drink something with a pleasant taste, consider my experience, and become convinced again.

²³ Aaron Smuts, "The Feels Good Theory of Pleasure," forthcoming, *Philosophical Studies*. For similar views, see John Locke, *Essay concerning Human Understanding* (1690) 2.2.1; David Hume, *Treatise of Human Nature*, L. A. Selby-Bigge (ed), Oxford: Clarendon Press (1888) 1.1.2; Roger Crisp, *Reasons and the Good*, New York: Oxford University Press (2006): 107-109.

²⁴ Fred Feldman, *Pleasure and the Good Life*, New York: Oxford University Press (2004).

the pleasure of tasting delicious food, receiving a massage, or resting your tired limbs in a soft bed after a hard day. There are the pleasures of seeing that our desires are satisfied, like the pleasure of winning a game, getting a promotion, or seeing a friend succeed. These experiences differ in many ways, just as the experiences we have when looking at lemons and the sky on a sunny day differ. It is easy to see the appeal of Feldman's view that pleasures "have just about nothing in common phenomenologically" (79). But just as our experiences in looking at lemons and the sky on a sunny day have brightness in common, pleasant experiences all have "a certain common quality – feeling good," as Roger Crisp argues (109).²⁵ As the analogy with brightness suggests, hedonic tone is phenomenologically very thin, and usually mixed with a variety of other experiences.²⁶ Pleasure of any kind feels good, and displeasure of any kind feels bad. These feelings may or may not have bodily location or be combined with other sensory states like warmth or pressure. "Pleasure" and "displeasure" mean these thin phenomenal states of feeling good and feeling bad. As Joseph Mendola writes, "the pleasantness of physical pleasure is a kind of hedonic value, a single homogenous sensory property, differing merely in intensity as well as in extent and duration, which is yet a kind of goodness" (442).²⁷

²⁵ Crisp (2006).

²⁶ For a similar analogy, but with the volume of sounds rather than the brightness of colors, see Shelly Kagan, "The Limits of Well-Being," in Paul, Miller, Jr., and Paul (eds.), *The Good Life and the Human Good*, Cambridge University Press (1992).

²⁷ Joseph Mendola, "Intuitive Hedonism," *Philosophical Studies* 128 (2006): 441-477. Also: "Each moment of phenomenal experience presents itself as positively or negatively valenced to some degree. The valence is either (a) a lesser or greater degree of intrinsic value, as in the phenomenal component of happiness and pleasure; (b) a lesser or greater degree of intrinsic disvalue, as in the phenomenal component of unhappiness and pain; or (c) a null valence, which is neither intrinsic value nor disvalue. That valence of the experience of a moment is an objective phenomenal property." Joseph Mendola, "Objective Value and Subjective States,"

What if Feldman is right and hedonic states feel good in fundamentally different kinds of ways? Then phenomenal introspection will suggest a pluralist variety of hedonism. Each fundamental flavor of pleasure will have a fundamentally different kind of goodness, as phenomenal introspection that is more accurate than mine will reveal. This is not my view, but I suggest it to those convinced that hedonic tones are fundamentally heterogeneous.

If phenomenal introspection reliably informs us that pleasure is good, how can anyone believe that their pleasures are bad? Hedonists can blame other processes of moral belief-formation for these beliefs. For example, someone who feels disgust or guilt about sex may not only regard sex as immoral, but the pleasure it produces as bad. Even if phenomenal introspection on pleasure disposes one to believe that it is good, stronger negative emotional responses to it may more strongly dispose one to believe that it is bad. This explanation of disagreement about pleasure's value lets hedonists deny that people believe that pleasure is bad on the basis of phenomenal introspection alone. As long as negative judgments of pleasure come from unreliable processes instead of phenomenal introspection, the argument from disagreement will eliminate them, while the reliable process of phenomenal introspection will univocally support pleasure's goodness.

The parallel between yellow's brightness and pleasure's goodness demonstrates the objectivity of the value detected in phenomenal introspection. Just as anyone's yellow experiences objectively are bright experiences, anyone's pleasure objectively is a good experience.²⁸ While one's phenomenology is often called one's "subjective experience", this does *Philosophy and Phenomenological Research* 50:4 (1990) 695-713. For a similar defense (with a graph of how the magnitude of this value fluctuates over a day) see Torbjorn Tannsjo, "Narrow Hedonism," *Journal of Happiness Studies* 8 (2007): 79-98. See also Crisp (2006).

²⁸ Here I use "goodness" and "value" interchangeably, as is customary. See for example Mark Schroeder, "Value Theory," Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <http://plato.stanford.edu/entries/value-theory/>.

facts about it are still objective. “Subjective” in “subjective experience” means “internal to the mind”, not “ontologically dependent on attitudes towards it.” My yellow-experiences are objectively bright, so that anyone who thought my yellow-experiences were not bright would be mistaken. Pleasure similarly is objectively good – it is true that anyone's pleasure is good, and anyone who denies this is mistaken. As Mendola writes, “In the phenomenal value of phenomenal experience, we have a plausible candidate for objective value” (712).

Even though phenomenal introspection only tells me about my own phenomenal states, I can know that others' pleasure is good. Of course, I cannot phenomenally introspect their pleasures any more than I can phenomenally introspect pleasures that I will experience next year. But if I consider my experiences of lemon yellow and ask what it would be like if others had the same experiences, I must think that they would be having bright experiences. Similarly, if in a pleasant moment I consider what it is like when others have exactly the experience I am having, I must think that they are having good experiences. If they have exactly the same experiences I am having, their experiences will have exactly the same intrinsic properties as mine. This is also how I know that if I have the same experience in the future, it will have the same intrinsic properties. Even though the only pleasure I can introspect is mine now, I should believe that pleasures experienced by others and myself at other times are good, just as I should believe that yellow experienced by others and myself at other times is bright. My argument thus favors the kind of universal hedonism that supports utilitarianism, not egoistic hedonism.

I can now outline hedonism's answer to Joyce and Street's evolutionary debunking arguments, as promised in section 1.1. Phenomenal introspection is likely to be a process of belief-formation that has evolved to be generally reliable, like visual perception. If knowing what one is experiencing is an important part of perception, creatures who could not accurately determine what their experiences were like might fail to form useful beliefs about their surroundings and die without reproducing. Or if direct realists are right about perception, such

creatures might still form perceptual beliefs, but be unable to discuss their experiences in a way that is useful to social beings like us. Either way, humans who could reliably form true beliefs about their phenomenal states would be more likely to survive and reproduce. Then we can defend hedonism in the face of evolutionary debunking arguments using what Street calls a “byproduct hypothesis.” Since belief in pleasure's goodness is a byproduct of processes selected for reliability, it is reliably caused even if other moral beliefs are not.²⁹ If all other moral beliefs are undermined by their origins in processes not selected for reliability, this will provide an evolutionary version of the epistemic argument for hedonism.³⁰

2.2 Pleasure's goodness is moral value

This section argues that the goodness of pleasure detected in phenomenal introspection is moral value. “Good” also expresses many nonmoral concepts of value, including aesthetic, functional, and prudential value.³¹ This suggests the objection that phenomenal introspection reveals pleasure to be good only in a nonmoral sense, which hedonists mistake for moral value. In light of Moore's point that analysis of moral concepts leaves normative ethical questions open, I cannot respond by analyzing the concept of moral value and showing that this analysis entails the moral value of pleasure. Instead, I will motivate the view that phenomenal introspection reveals pleasure's moral value first by outlining how it leads us to morally judge

²⁹ This would provide a targeted debunking argument of the sort considered in Guy Kahane, “Evolutionary Debunking Arguments” *Nous* (2011). For further discussion of how hedonists are poised to use targeted debunking arguments, see Adam Lerner, “Fine-Tuning Evolutionary Debunking Arguments” (unpublished).

³⁰ Street considers defenses of realism grounded in pleasure's goodness and pain's badness, but these arguments do not consider pleasure and displeasure to be hedonic tones.

³¹ Many of these are not plausible alternatives. Phenomenal introspection does not reveal that pleasure has functional value like a good knife, or that it has aesthetic value like *Macbeth*.

actions causing pleasure, then by discussing a kind of metaethical theory in which it could produce moral knowledge, and finally by describing how to develop a full moral theory from pleasure's value.

To see how phenomenal introspection on pleasure and displeasure reveals genuine moral properties, consider how these experiences relate to moral praise and blame. The phenomenally introspectible badness of pain is a bad thing that we can morally blame cruel people for intentionally causing and morally praise kind people for intentionally relieving. Victims of torture who suffer pain know how bad it is through their experiences of suffering it, and this badness is central to what makes torture wrong. Similarly, if nice people act to cause others pleasure, the experienced goodness of the pleasure is a good thing that justifies the actions and makes them praiseworthy. To see how phenomenal introspection reveals moral value, contrast how a purely neurological view of pleasure and pain fails to reveal the features of these situations justifying praise and blame. Understanding what happens to victims of torture merely in neurological terms does not reveal what is so bad about their situation, what justifies blaming those who put them in that situation, and what justifies praising those who alleviate their condition. Only when we consider how these neurological states relate to their experiences do we see what is good or bad in these situations, grounding our moral evaluations.

Moral value could be located within our phenomenology on the naturalistic realist view suggested by Peter Railton, potentially making it accessible to phenomenal introspection. This reductionist form of moral realism treats moral properties as standing in synthetic *a posteriori* identity relations to the natural properties that they supervene on, just as water is identical to H₂O.³² So if pleasure is located within our experience and pleasure has moral value, moral value is located in our experience. This does not guarantee that phenomenal introspection could

³² Peter Railton, "Moral Realism," *Philosophical Review*, 95:2 (1986), 163-207; and "Naturalism and Prescriptivity," *Social Philosophy and Policy*, 7:1 (1989), 151-174. While Railton is not a hedonist, he is sympathetic enough to hedonism to describe a way of confirming it in the latter paper.

inform us of the moral facts. If X and Y are identical as a matter of synthetic *a posteriori* necessity, one can know that X is present without being in any position to know that Y is present. But reductionism metaphysically locates moral value in such a place that it is an empirical question whether phenomenal introspection can access it. To empirically determine whether phenomenal introspection reveals anything to be of moral value, we should try it out and check whether any components of our phenomenology seem to be morally valuable. We detect some kind of value in our experiences of pleasure, and apply the word 'good' to it. The case of torture above demonstrates how the role of this kind of value in justifying moral praise and blame of those whose actions promote or reduce it suggests categorizing it as moral value. While few philosophers have explicitly considered whether experience reveals pleasure's moral value, those I know of agree. In explaining the appeal of moral theories like hedonic utilitarianism, Robert Shaver writes that "The goodness or badness of pleasures or pains is evident from the point of view of anyone capable of experience" (248)³³ and Shelly Kagan writes that "The value of pleasure and the disvalue of pain seem virtually self-evident to anyone experiencing them" (30).³⁴ If simply experiencing pleasure provides evidence for theories on which it has moral value, as Shaver and Kagan argue, the value detected in introspection on pleasure must be moral value. Phenomenal introspection, then, reveals pleasure's moral value, just as naturalistic realism allows.

A full moral theory including accounts of rightness and virtue can be built from the deliverances of phenomenal introspection combined with conceptual analysis. Shaver, Kagan, and I suggest that phenomenal introspection reveals pleasure to have a kind of goodness that makes states of affairs better in consequentialist moral theories. A state of affairs thus is *pro tanto* better as there is more pleasure and *pro tanto* worse as there is more displeasure. More pleasure makes states of affairs better. Conceptual analysis here connects the concept of goodness with

³³ Robert Shaver, "The Appeal of Utilitarianism," *Utilitas* 16:3 (2004): 235-250.

³⁴ Shelly Kagan, *Normative Ethics*, Boulder: Westview Press (1998).

the concept of a better state of affairs, and with other moral concepts like rightness and virtue. Even if conceptual analysis cannot connect the moral and the nonmoral as a full normative ethical theory requires, it reveals connections between our moral concepts. For example, the following propositions or something like them seem to be conceptual truths: states of affairs are *pro tanto* better insofar as they include more goodness, an action is *pro tanto* better insofar as it causally contributes to better states of affairs, and agents are *pro tanto* more virtuous insofar as they desire that better states of affairs obtain. These putative conceptual truths about *pro tanto* relations do not contradict strong forms of deontology, as they allow that obligations may trump good consequences in determining right action. Utilitarians who build their theories along these lines can treat deontology as a conceptually coherent position whose substantive claims are in fact not favored by evidence from any reliable processes. So they need not treat utilitarianism itself as a conceptual truth and run afoul of Moore's open question argument. If the argument from disagreement forces us to abandon belief in all other moral facts, introspecting pleasure's goodness and following these conceptual *pro tanto* connections to conclusions involving other moral concepts may be the only way to develop a full moral theory through reliable processes.

3.0 Conclusion

First I argued that we should give up all our moral beliefs, as widespread moral disagreement reveals the general unreliability of the processes producing them. Then I argued that phenomenal introspection reliably informs us of pleasure's goodness. This is a counterexample to premise 4 of the argument from disagreement. However, premises 1-3 of that argument still stand, entailing 6: "We should retain only those moral beliefs formed through reliable processes." If phenomenal introspection is the only reliable process for discovering synthetic moral truths, our only moral beliefs should be in the goodness of pleasure and

whatever it entails. We should accept a version of ethical hedonism on which we doubt rather than deny nonhedonic moral facts.

This gives hedonists a powerful response to the intuitively compelling counterexamples that lead many to reject their view and believe in a wider range of moral facts. Hedonists can respond that while it intuitively seems that many kinds of moral goodness, badness, rightness, wrongness, virtue, and vice are not grounded in pleasure's goodness, beliefs in such moral facts are unacceptably likely to be mistaken. Accepting nonhedonic moral facts is intuitive to us only because we are afflicted by unreliable processes of belief formation. To avoid moral error, we should form moral beliefs only through reliable processes. If we do so, pleasure is the only thing we will believe is good. The conceptual connections between moral properties of things, states of affairs, actions, and characters will then provide our only reliable route to a full moral theory.

The argument offered here will not necessarily make the moral theory suggested by its conclusion feel right. The epistemic processes that have led so many people into error make the putative counterexamples to hedonism intuitively powerful, attracting us to belief in nonhedonic moral facts. Hedonism will feel wrong even to hedonists. We might be surprised that believing the true moral theory feels this way. But this is as we should expect, given how much moral error has historically existed and how deeply it is entrenched in our moral feelings. This is how it feels to face our strongest cognitive biases. Whatever the moral truth may be, it would feel wrong to most humans.

In many areas of inquiry, methodological advances have been essential in advancing human knowledge. Nothing short of a methodological revolution could take us from the widespread moral error that has afflicted humankind to knowledge of the moral truth. If my arguments are sound, the truth is that pleasure is the only thing we can know to be good. Relying solely on phenomenal introspection to reveal the nature of goodness is the revolution.