

Naturalistic Arguments for Ethical Hedonism

Neil Sinhababu

Guest essays represent only the views of the author(s).

Table of Contents

1. Introduction: From Naturalism to Hedonism
2. The Reliability Argument
 - R1. The reliability of a process is the probability that beliefs it generates are true
 - R2. Phenomenal introspection is reliable in generating belief that pleasure is good
 - R3. No other processes are independently reliable in generating moral belief
3. The Universality Argument
 - U1. x is good \leftrightarrow x should please all metaphysically possible moral perceivers
 - U2. x should please $Y \leftrightarrow x$ makes pleasure accurate for Y
 - U3. x makes pleasure accurate for all metaphysically possible moral perceivers $\leftrightarrow x$ is pleasure
4. Conclusion: Naturalistic Foundations for Utilitarianism
5. Further Reading

Introduction: From Naturalism to Hedonism

Ethical hedonism is the view that pleasure is the only good thing—the only thing with moral value.¹ This essay presents two arguments for ethical hedonism.

The **Reliability Argument** assesses the psychological processes generating moral belief to determine how reliably they produce truth. Phenomenal introspection is the process of looking inward at experience to see what it's like. It reliably generates true beliefs about experience, for example the nonmoral belief that sound-experience has volume, and the moral belief that pleasure is good. While the reliability of phenomenal introspection is demonstrated by the true nonmoral beliefs it generates, there is evidence that other processes generating moral beliefs are unreliable. Widespread moral disagreement across history on issues like whether genocide is permissible entails widespread error and unreliability. The evolutionary history of processes generating moral belief also makes it unlikely that they are reliable. Since a reliable process generates belief in pleasure's goodness, and unreliable processes generate other moral beliefs, pleasure is probably the only good thing.²

The **Universality Argument** explains how all possible moral perceivers should feel. All should regard good events with pleasant feelings like hope, and bad events with unpleasant feelings like horror. These pleasant and unpleasant feelings are experiences that represent reality, like visual experience and other sensory experiences, so they should correspond accurately with reality as sensory experiences should. How could pleasure accurately represent the same object for all who feel it? Not if pleasure represents what systematically causes it, or what one's culture or conscience suggest, as these differ between perceivers. Universality is achieved when sameness between representations and their objects constitutes accuracy, as in quotations, depiction, onomatopoeia, and empathy. Pleasant moral feelings can therefore be universally accurate about pleasure. So pleasure is goodness—what all should hope for.³

“Pleasure” refers to a feeling occurring in sensations like tasting delicious food and feelings like delight and admiration.⁴ While sensory and emotional pleasures differ in various ways, experiencing any of them involves feeling better. What unifies all pleasure is that experiencing it makes one feel better; what unifies all displeasure is that experiencing it makes one feel worse (holding all else fixed). The hedonic tone of experience—how good something feels—slides on a continuous scale, much like the goodness and badness of events. Both the Reliability and Universality Arguments for ethical hedonism equally support pleasure’s goodness and displeasure’s badness; for brevity, I’ll often refer only to pleasure’s goodness.

Both arguments support [naturalistic moral realism](#). This metaethical view treats moral value as an objective feature of natural reality—the realm of empirically observable things within spacetime. Pleasure is an objective feature of natural reality. So is goodness, if it’s identical to pleasure.⁵

The moral truths sought here are *objective* and *universal*. Objective truths are independent of anyone’s attitudes towards them. The laws of physics are objective. Truths about what is delicious or fashionable aren’t objective, as they depend on the tastes of individuals and social groups. Universal truths are the same for all metaphysically possible beings. In the case of morality, they tell all possible beings to hope for the same types of outcomes, feel guilty about the same types of actions, and admire the same types of character traits.

Many deny that objective and universal moral truth exists.⁶ Many more deny that it can be empirically discovered within natural reality.⁷ The arguments in this essay seek to empirically discover objective and universal moral truth in natural reality. This truth is that *pleasure is goodness*.

The Reliability Argument

The Reliability Argument has three premises (R1–R3) that together entail a conclusion (%PIG):

R1. The reliability of a process is the probability that beliefs it generates are true.

R2. Phenomenal introspection is reliable in generating belief that pleasure is good.

R3. No other processes are independently reliable in generating moral belief.

Therefore,

∴ %PIG: Pleasure is probably the only good thing.

R1. The reliability of a process is the probability that beliefs it generates are true

R1 defines reliability, the truth-conduciveness of belief-generating processes.

⁸ Visual perception is generally reliable, as it tends to generate true belief; wishful thinking is unreliable, as it doesn't. This difference suggests retaining beliefs generated by visual perception, and doubting beliefs generated by wishful thinking.

The reliability of a process is expressed by its truth ratio, the proportion of the total beliefs it generates that are true. If a process generates 1,000 beliefs and 954 of them are true, its truth ratio is 0.954 or 95.4%. If all we know about a belief is the truth ratio of the process generating it, we should assign it a probability of truth equal to its truth ratio. Cognitive science can help in more precisely distinguishing processes and assessing their reliability.⁹ Observed differences in reliability provide inductive evidence about which processes are more likely to generate true belief.

To avoid circularity, the reliability of the processes generating moral belief must be assessed without assuming which moral beliefs are true. This can be done in three ways.

First, we can investigate processes' reliability in generating nonmoral belief, and inductively extend it to moral belief. Using phenomenal introspection to discover pleasure's goodness works this way. Since phenomenal introspection reliably generates nonmoral beliefs about experience, induction suggests that it reliably generates moral beliefs about experience.

Second, investigating whether processes systematically generate inconsistent moral beliefs can demonstrate their unreliability, as any jointly inconsistent set of beliefs must include at least one falsehood. Arguments from disagreement work this way. Any process responsible for widespread inconsistency between different cultures' moral beliefs generates great error, and is unreliable.¹⁰ Third, investigations into how the process was historically formed may suggest its reliability or unreliability. [Evolutionary debunking arguments](#) work this way. If the process has an evolutionary or sociological history that would make its reliability an improbable coincidence, it's probably unreliable.

R2. Phenomenal introspection is reliable in generating belief that pleasure is good

R2 presents phenomenal introspection as a reliable process generating belief in pleasure's goodness. Phenomenal introspection generates beliefs about experience as one attends to what it is like.¹¹ It can be reliable about what dreams and hallucinations are like, even though these experiences are unreliable about the outside world. It doesn't generate beliefs about anything beyond experience.

Phenomenal introspection generates belief that pleasure is good, just as it generates belief that sound-experience has volume.¹² Attending to pleasure or sound-experience reveals what it's like. The only moral beliefs phenomenal introspection generates are that pleasure is good and displeasure is bad. As introspection can only generate beliefs about experience itself, it doesn't

generate any moral beliefs about the rightness of actions, which character traits are virtues, or other matters extending beyond experience.

While one can apply phenomenal introspection only to one's own experience, it suggests that anyone's pleasure is good, just as it suggests that anyone's sound-experience has volume.¹³ After introspection on my sound-experience reveals its volume, I should expect that your sound-experiences have volume too. So if introspection on my pleasure reveals its goodness, I should expect your pleasure to be good too. It may be difficult to determine whether some nonhuman animals or other beings unlike us can have sound-experiences or pleasure. But if they can, their sound-experiences have volume and their pleasure is good.

Phenomenal introspection reliably reveals objective facts about experience. Denying its reliability would lead to a bizarre sort of skepticism about what one's own experiences are like. There are objective facts about which sensations or feelings we're having and what they're like. Phenomenal introspection reveals facts of this kind. Some of these experiences might be unreliable about the external reality they represent, but this doesn't diminish the reliability of phenomenal introspection, which generates beliefs only about experience and not about what it represents. That sound-experiences have volume and pleasure is good are objective facts about experience itself. Any culture would be wrong to contradict phenomenal introspection by claiming otherwise.

To discover the moral truth by phenomenal introspection is to extend a demonstrably reliable process into a new domain, as Galileo did in turning his telescope to the stars. Just as the telescope's reliability with faraway earthly things inductively suggested that it was right about Jupiter having moons, the reliability of phenomenal introspection with nonmoral experience inductively suggests that it's right about pleasure being good.

R3. No other processes are independently reliable in generating moral belief

R3 denies the independent reliability of processes other than phenomenal introspection in generating moral belief. It can be defended with general arguments for moral skepticism, such as arguments from disagreement and debunking arguments.

Processes that are independently reliable in generating true moral belief don't require true or accurate moral input for reliability.¹⁴ Phenomenal introspection is independently reliable, generating true belief about pleasure's goodness with only pleasure as input. Trusting others about morality is only dependently reliable, as it depends on the truth of what they say. The same is true of conceptual inferences about morality, like inferring that since an action is wrong, it can't be right. The truth of the conclusion depends on the action being wrong. If no independently reliable processes generate moral belief, dependently reliable processes will lack the systematically true or accurate moral inputs needed for reliability.

Arguments from disagreement find widespread inconsistencies in human moral belief, showing that any process generating them is unreliable.¹⁵ Opinions about the permissibility of genocide serve as an example, differing dramatically throughout history.¹⁶ Consider the commands of the Old Testament's divine lawgiver, in a book considered holy by most alive today: "You shall save alive nothing that breathes, but you shall utterly destroy them, the Hittites and the Amorites, the Canaanites and the Perizzites, the Hivites and the Jebusites, as the Lord your God has commanded."¹⁷ Subsequent millennia saw countless societies proudly commit genocide. Even otherwise advanced societies of the mid-20th century thought it right to kill millions. Some hoped to kill millions more. Given how many human cultures thought it right to slaughter neighboring cultures and take their resources, we who regard genocide as impermissible may be in the historical minority.

Moral disagreement is similarly broad on many other important issues, including animal treatment, slavery, sexual morality, distributive justice, and gender roles. Disagreement extends to seemingly easy moral questions where error would lead to countless deaths—is genocide wrong? Any process generating most human moral judgments can't be reliable, as it produces enough disagreement to entail widespread error and therefore unreliability.

Evolutionary debunking arguments suggest that the processes generating our moral beliefs are optimized not for reliability, but for evolutionary fitness—survival and reproduction in our ancestors' environment.¹⁸ While processes optimized for reliability tend to generate true belief, processes optimized for fitness help those who have them survive and reproduce. Since we inherited genes and traits from those who survived and reproduced, our processes are likely to be optimized for fitness but not for reliability. It would be unlikely and coincidental for such processes to reliably generate true moral belief.

(Debunking arguments can likewise explain unreliability psychologically or sociologically rather than evolutionarily. What matters is that the historical forces shaping the processes suggest unreliability, often by suggesting that reliability would be unlikely and coincidental.)

It's important for evolutionary debunking arguments that fitness and reliability don't coincide, or else fitness-promoting processes would likely be reliable. Fitness does coincide with reliability about other topics like locating food, because knowing where food is promotes survival and reproduction. But there's no similarly obvious explanation of why reliability about morality would enhance fitness. Sharing the moral errors of one's society might improve one's social relations, increasing one's fitness. Assuming commonsense moral truths and noting that they're somewhat fitness-enhancing suggests that fitness and reliability coincide. But investigations into which ethical theory is true can't assume commonsense moral truths, as this would be circular.¹⁹ The widespread unreliability suggested by these skeptical arguments doesn't extend to phenomenal introspection. Even if unreliability is

widespread, processes generating only a small share of belief can be reliable. Phenomenal introspection is reliable, as demonstrated by the nonmoral beliefs it generates. Arguments from disagreement and debunking arguments can't block the inductive inference that it's reliable about morality.

A final objection to R3 notes that phenomenal introspection might not be the only reliable process accounting for a small share of moral belief. Another small reliable process might exist alongside the widespread error suggested by the skeptical arguments. Identifying another reliable process amidst widespread error would require evidence of its reliability, perhaps from true nonmoral beliefs it generates. Discovering one would be a breakthrough in moral epistemology, as present evidence suggests that phenomenal introspection stands alone.

If the Reliability Argument correctly describes our epistemic situation, phenomenal introspection is the only independently reliable process generating moral belief. It reveals pleasure's goodness. Since all other moral beliefs originate in unreliable processes, there is insufficient evidence for any moral truths other than pleasure's goodness. So pleasure is probably the only good thing.

Simplicity considerations suggest the further conclusion that pleasure and goodness are the same thing, rather than being two different things that happen to coincide. Pleasure is the sole form of goodness discovered through reliable processes, just as H₂O is the only form of water discovered by reliable processes. Even before testing all known water, science concludes that water is H₂O, as this is simpler than regarding them as separate but coexisting. Ethicists might draw the similarly simplifying conclusion that pleasure is goodness.

The Universality Argument

The Universality Argument has three premises (U1-U3) that together entail a conclusion (PIG):

U1. x is good $\leftrightarrow x$ should please all metaphysically possible moral perceivers

U2. x should please $Y \leftrightarrow x$ makes pleasure accurate for Y

U3. x makes pleasure accurate for all metaphysically possible moral perceivers
 $\leftrightarrow x$ is pleasure

Therefore,

\therefore PIG: Pleasure is goodness.

U1. x is good $\leftrightarrow x$ should please all metaphysically possible moral perceivers

U1 defines goodness as whatever should please all metaphysically possible moral perceivers. Other moral concepts like rightness and virtue also have hedonic analyses in terms of how all should feel. The point of such analyses is to explain how moral concepts relate to other concepts and to our feelings. I'll defend U1 by showing that it and similar hedonic analyses explain these relations.

Morality is universal in the sense that it explains how all metaphysically possible moral perceivers should feel. (For "all metaphysically possible moral perceivers", I'll often just use "all".) A metaphysically possible moral perceiver is anyone who could have pleasant or unpleasant feelings—like hope, pride, admiration, horror, guilt, and hatred—and take these feelings to represent objective features of reality. Morality therefore concerns how people of all existing cultures should feel, as well as any other beings who can have similar feelings and take them to represent reality, including fictional beings like Tolkien's hobbits or Princess Leia from *Star Wars*.

To say something has moral value is to say that it should please this vast range of possible beings. That isn't to say that they are all pleased. Actual humans differ dramatically in what pleases them, and metaphysically possible moral perceivers differ even more. We however seek to discover how they all should

feel. The Universality Argument is largely concerned with better understanding this *should*. The next two premises identify it with a relation in natural reality, and explain how it can be universal.

Other moral concepts applying to events, actions, character traits, and societies similarly concern how all should feel. “You should feel guilty about what you did” is a way to call someone’s action wrong. Calling character traits “admirable” is saying they’re virtues; calling them “contemptible” is calling them vices. The hedonic analyses of moral concepts below show that this wide range of moral concepts can be analyzed in terms of norms for pleasant and unpleasant feelings.

Concept	Hedonic analysis	Examples of moral feelings
GOOD	events that should please all	hope (if uncertain), delight
RIGHT	actions that should please all	pride (for oneself)
VIRTUOUS	character traits that should please all	admiration
BAD	events that should displease all	horror, sadness
WRONG	actions that should displease all	guilt (for oneself), anger (for others)
EVIL / VICE	character traits that should displease all	hatred (for evil), contempt (for vice)

Moral concepts apply to whatever should arouse moral feelings in all.²⁰

Pleasure and displeasure are the basic moral feelings. More complex moral feelings such as hope, guilt, and admiration have additional phenomenal character or representational content, as well as being pleasant or unpleasant.

The analysis excludes motivational, deliberative, or physiological effects accompanying moral feelings, which aren't essential to moral concepts.

Feelings are experiences, like sensations. Feeling guilt about one's action is perceiving it as wrong, just as sensing a three-sided shape is perceiving a triangle. "Wrong" applies to actions that anyone should feel guilty about after doing, much as "three-sided" applies to triangles. Misperception is possible with both feelings and sensations. One might mistakenly feel guilt about homosexual activity due to a conservative upbringing, or mistakenly see a trapezoid as a triangle due to blurry vision.

Moral concepts trace the outlines of moral feelings. The feelings' hedonic character explains the concepts' moral valence. All should be pleased about morally positive things, hoping for good events, taking pride in right action, and admiring virtue. All should be displeased about morally negative things, being horrified by horrible events, feeling guilty about their wrong actions, and hating evil. Generally speaking, all should feel good about the good, and feel bad about the bad.

Relations between feelings explain why some combinations of these concepts are incoherent. It can be incoherent to call an event "good and bad" because one can't feel better and worse about something simultaneously. It can be similarly incoherent to call something yellowish-purple, as one can't see both colors in the same place simultaneously. In both cases, separating the objects of experience restores coherence. Just as one can see yellow lines on a purple background, one can admire virtue while feeling contempt for one's own vice. One might be pleased by something's good features while being displeased by its bad features. But as one thinks specifically of a single feature of something—perhaps only the kindness, or only the dishonesty, of a kind but dishonest action—one's pleasure or displeasure becomes purer.

Intentional relations connecting hope, pride, and admiration explain conceptual relations between goodness, rightness, and virtue. All should be proud to make the world more like all should hope for it to be. Desiring to make

the world that way is admirable. Hedonic analyses translate these relations between feelings into relations between concepts: It's right to make the world a better place, and desiring to do so is virtuous.

Hedonic analyses leave substantive moral questions open, allowing conceptual coherence for false moral theories.²¹ One can coherently believe that wearing shoes is wrong—that all should be displeased at those who wear shoes. This moral claim might be ridiculous, but it isn't incoherent like “wearing shoes is being barefoot.” The analyses don't entail substantive moral claims, so they don't render any substantive moral theories incoherent. Further premises are needed to derive a substantive moral theory.

U2. x should please Y ↔ x makes pleasure accurate for Y

U2 identifies *should* with accuracy. Together with U1, it entails that goodness makes pleasure accurate for all. The concept of accuracy is familiar from sensory perception—looking at a triangle makes visual experience of a triangle accurate. Hedonic analyses implicitly suggest U2 by applying *should* to feelings. Since moral feelings are experiences representing reality, like sensations, they should be accurate as sensations should. Longstanding orthodoxy instead analyzes moral concepts in terms of how all should act.²² *Should* then can't be identified with accuracy, as actions aren't experiences representing reality. After explaining what accuracy is, I'll defend U2 with empirical evidence that moral concepts apply fundamentally to feeling rather than action.

Accurate experiences correspond with reality, much as true beliefs do. Experiences should be accurate rather than inaccurate, just as beliefs should be true rather than false.²³ Having visual experiences of triangles when looking at triangles is accurate—then you're seeing the triangles as you should, and your perception corresponds with reality. Having visual experience of triangles when looking at trapezoids is inaccurate—then you aren't seeing the trapezoids as you should, and your perception doesn't correspond with reality.

Utilitarians about morality can accept norms of truth for belief and accuracy for experience, as these are nonmoral norms. False beliefs are wrong, but usually not morally wrong.

Empirical evidence suggests that pleasure and displeasure cause moral belief, just as sensations do. When experimental subjects fill out questionnaires asking how wrong various unsavory actions are, subtly giving them additional displeasure with annoying typography on the questionnaires or bad smells in the room causes harsher moral judgments.²⁴ Surrounding stimuli can affect beliefs formed from experience this way. Yellow light shining on a white wall can similarly cause belief that the wall is yellow. Feelings, like sensations, generate beliefs that share their representational content. For the sensation or feeling to be accurate is for reality to match its content.

Experience can give content to our concepts. Fully understanding the concept of yellow requires knowing what yellow-experience is like. If we couldn't have feelings like hope and guilt, goodness and rightness would be as hard for us to understand as yellow would be to someone who never had yellow-experiences.

The orthodox view that moral concepts should generally be analyzed in relation to action is empirically implausible, as moral belief is insufficient to motivate action or generate new motivation.²⁵ This is why people often do what they believe is wrong, eating meat despite believing they're wrongfully harming animals, or wasting time when they know there are more morally important things to do. Moral beliefs are similarly insufficient to generate desires or other stable motivational states. This is why attempts to convert homosexuals to heterosexuality have repeatedly failed, even when the participants themselves regarded a heterosexual orientation as morally better.²⁶ Since sexual orientation is a matter of desire, it can't be changed by moral belief, even when accompanied by religious conviction and social pressure over time.²⁷ While believing something is right and desiring to do what's right can motivate action and instrumentally generate desires, nonmoral beliefs combined with desires for nonmoral ends can do the same. So moral concepts

don't in general have any distinctive relation to action, contrary to metaethical orthodoxy.²⁸ Concepts of rightness and wrongness apply to action, but believing that they apply is insufficient to generate any motivation for or against it.

Experience is sufficient to cause belief that shares its content, but belief is insufficient to cause action or change desire. This fundamental psychological fact suggests analyzing moral concepts in terms of feeling rather than action. Experience is typically causing belief at a rapid pace, as a rapid flow of perceptual experience updates our beliefs about things around us. But even deeply held moral beliefs aligned with social and religious pressure over time fail to cause action or motivational states, unless desire is present. The psychological role of moral concepts requires analyzing them in terms of the feelings causing moral beliefs, rather than actions or motivational states that moral beliefs have no distinctive power to cause.²⁹ Moral feelings are experiences with representational content, so *should* applies a norm of accuracy to them. Accuracy is a non-normative relation observed to exist in natural reality. For visual experience of a triangle to be accurate doesn't require anything normative or non-natural—it just requires a triangle. Applying *should* to action doesn't allow this naturalistic reduction. Since actions and desires don't represent reality, they can't be accurate or inaccurate. This is why most theorists analyzing moral concepts in terms of action either regard morality as beyond natural reality, or deny its objectivity and universality. Identifying *should* with accuracy preserves hopes for a universal and objective morality contained within the natural world.

U3. x makes pleasure accurate for all metaphysically possible moral perceivers ↔ x is pleasure

U3 reveals what makes pleasure accurate for all—something else identical to itself, namely another experience of pleasure. Accuracy is [qualitative identity](#), or sameness between things. Identity constitutes accuracy for many other forms of representation. To defend U3, I'll show that treating accuracy as

identity explains the universality of morality, making the moral feelings of all accurate about the same things.

The universality of morality requires the same moral standard for all. U1 and U2 therefore treat universal moral value as making pleasure accurate for all. Relativist views treating moral value as determined by cultural standards or individual conscience deny universality, allowing different moral truths for different cultures or individuals.³⁰ If genocide is good in genocidal cultures and bad in peaceful cultures, what's good and bad isn't universal. Since moral truth must be universal, relative truth isn't genuine moral truth. Similar problems afflict recent naturalistic theories, which treat moral beliefs as made true by whatever systematically causes them.³¹ As different things systematically cause the moral beliefs (and moral feelings) of different possible beings, these theories fail to uphold universality.³² Qualitative identity, or sameness, is the standard of accuracy for pictures, quotations, replicas, shape-perception, onomatopoeia, and empathy. Pictures are accurate if they look the same as what they depict. Accurate quotations are the same as the quoted sentences. Replicas are accurate to the extent that they're the same as whatever they're replicas of. One sees a shape accurately by having the same shape in one's visual field. "Vroom vroom" is accurate onomatopoeia for car sounds to the extent that it sounds the same as a car.³³ Empathic experience is accurate to the extent that the person empathized with has the same experience.³⁴ Greater resemblance counts as greater accuracy for these forms of representation.³⁵ Even when identity isn't achievable or feasible, it serves as a standard of perfect resemblance that would constitute perfect accuracy.

If qualitative identity constitutes accuracy, any possible moral perceiver's pleasure is accurate about the same kind of thing as itself, namely other instances of pleasure. While relativism violates universality by treating different moral perceivers' pleasure as accurate about different things, treating qualitative identity as accuracy upholds universality by making all

pleasure accurate about the same thing. All should be pleased by pleasure. This makes ethical hedonism the universal moral truth.

Non-naturalists reject such reductions of *should*. They might argue that an irreducible and non-natural *should* better accommodates intuition. Since reducing *should* to identity entails ethical hedonism, it will violate whatever intuitions ethical hedonism violates, perhaps as well as others. Irreducible *should* isn't constrained by the reduction to identity, and may escape any such counterintuitive commitments.

To understand why reducing *should* to identity is still better than leaving it unreduced, consider why reducing water to H₂O is better than joining alchemists in treating it as an irreducible element of Water alongside Fire, Earth, and Air. Alchemists could rightly claim that positing irreducible Water avoids the counterintuitiveness of treating water as half of one gas molecule bound to two halves of another. But reducing water to H₂O lets the nature of the underlying hydrogen and oxygen atoms explain why it's a liquid at room temperature, how it shapes Earth's ecology, and how it supports life. In the end, the deep explanations made possible by reducing water to its component atoms are better than the superficial intuitiveness of positing an irreducible element of Water.

The reduction of *should* to identity offers similarly deep explanations. It tells us why all metaphysically possible beings should be horrified by the suffering of those in poverty, and of animals in factory farms. It tells us why all should hope life on Earth continues for a long and happy future.³⁶ These explanations begin in the nature of horror and hope themselves. Both are experiences, and should be accurate about reality. Displeasure is in both horror and others' suffering; pleasure is in hope and future happiness. Sharing the same properties in this way makes replicas accurate about their objects, "vroom vroom" accurate about car noises, and empathy accurate about others' feelings. It also makes horror accurate about suffering and hope accurate about pleasure for all.

Natural reality contains everything needed to explain how all possible minds should feel. When you hope that others experience pleasure, *should* connects your hope to their pleasure. Your pleasure in hoping is like a picture of their pleasure, accurately representing its object by sameness. That is how the moral experience of all should be.

Conclusion: Naturalistic Foundations for Utilitarianism

The Reliability Argument provides empirical evidence that pleasure is the only good. The Universality Argument provides a naturalistic explanation of why all should hope for more pleasure in the world. I'll conclude by showing how these arguments support each other, and how they suggest hedonic utilitarianism.

Both arguments suggest that all pleasure is good, whenever it occurs and whoever has it. The goodness detected by phenomenal introspection is present in anyone's pleasure at any time, just as the volume it detects in sound-experience will be present in anyone's sound-experience at any time. Relations of qualitative identity similarly hold between all equally strong pleasures at all times, making them equally to be hoped for.

Though the arguments stand independently, they also defend each other from important objections. One objection to the Universality Argument is that pleasure isn't the only good thing, so identity can't be accuracy for all moral feeling. But according to the Reliability Argument, belief in other good things is unreliably generated. So we should abandon such belief rather than modifying our conception of accuracy.

Likewise, one objection to the Reliability Argument is that phenomenal introspection can't detect moral value. But according to the Universality Argument, moral value is identical to pleasure. Phenomenal introspection can determine whether an experience is identical to pleasure. The two arguments come together into a unified naturalistic metaethics supporting hedonism.

The hedonic analyses alongside U1 explain conceptual relations in terms of relations between feelings, supporting inferences from pleasure's goodness to utilitarian accounts of rightness and virtue. If you strive to make the world more like all should hope for it to be, all should be proud of you and admire your motivation. So making the world a better place is right and striving to do so is virtuous. Considering the intensity of feeling suggests a utilitarian structure for quantitative relations between goodness, rightness, and virtue. The more intensely all should hope for the outcome you strive for, the prouder they should feel and the more they should admire you. Better outcomes are therefore more right to bring about and more virtuous to desire.³⁷ Skeptical arguments supporting R3 preserve the simplicity of this utilitarian structure. Relations between goodness and rightness might otherwise be complicated by deontological principles [distinguishing harm done from harm allowed](#), or intended harm from foreseen harm. Other complications might arise from the [person-affecting principle](#) that creating happy people doesn't make actions right. Unless evidence arises that belief in these principles is reliably generated, they should be rejected.³⁸ Arguments from disagreement and evolutionary debunking arguments indicate the unreliability of processes generating moral belief other than phenomenal introspection. Humans must beware of unreliably generated moral beliefs, as members of our species all too often believe that genocide is right.

Empirical observation reveals objective and universal moral value existing within reality. The reliability of phenomenal introspection, the gruesome history of moral error, and the evolutionary origins of moral judgment together provide empirical evidence for hedonism. The nature of accuracy in depiction, onomatopoeia, and empathy, together with the way moral feelings represent moral truths, together explain the universality of pleasure's moral value.

About the Author



[Neil Sinhababu](#) is Associate Professor of Philosophy at the National University of Singapore. His main interests are in metaethics, alongside various projects in Nietzsche scholarship and modal metaphysics. He hopes to write a book presenting these arguments in more detail, perhaps titled *Pleasure Fundamentalism*.

How to Cite This Page

Sinhababu, N. (2023). Naturalistic Arguments for Ethical Hedonism. In R.Y. Chappell, D. Meissner, and W. MacAskill (eds.), *An Introduction to Utilitarianism*, <<https://www.utilitarianism.net/guest-essays/naturalistic-arguments-for-ethical-hedonism>>, accessed 8/12/2025.

Further Reading

- Railton, Peter (1989). Naturalism and prescriptivity. *Social Philosophy and Policy*, 7(1): 151–174.
- Rawlette, Sharon (2016). *The Feeling of Value*. CreateSpace Independent Publishing Platform, California.
- Sinhababu, Neil (forthcoming). Pleasure is goodness; morality is universal. *Ethical Theory and Moral Practice*.
- Sinhababu, Neil (forthcoming). The epistemic argument for hedonism. In Chakraborty (ed.), *Human Minds and Cultures*. Routledge.

-
1. Bentham, Jeremy (1780). *An Introduction to the Principles of Morals and Legislation*. Dover Publications. Rawlette, Sharon (2016). *The Feeling of Value*. CreateSpace Independent Publishing Platform, California. ↩
 2. Sinhababu, Neil (forthcoming). The epistemic argument for hedonism. In Chakraborty (ed.), *Human Minds and Cultures*. Routledge. Sinhababu, Neil (draft). The reliable route from nonmoral evidence to moral conclusions. ↩
 3. Sinhababu, Neil (forthcoming). Pleasure is goodness; morality is universal. *Ethical Theory and Moral Practice*. ↩
 4. Bramble, Ben (2013). The distinctive feeling theory of pleasure. *Philosophical Studies*, 162(2): 201–217; Crisp, Roger (2006). *Reasons and the Good*. New York: Oxford University Press. ↩
 5. Railton, Peter (1989). Naturalism and prescriptivity. *Social Philosophy and Policy*, 7(1): 151–174. ↩
 6. Nietzsche, Friedrich (1954). *The Portable Nietzsche*. Penguin Books. Mackie, John Leslie (1977). *Ethics: Inventing Right and Wrong*. Penguin Books. ↩
 7. Huemer, Michael (2005). *Ethical Intuitionism*. Palgrave Macmillan. Parfit, Derek (2011). *On What Matters*. Oxford University Press. ↩
 8. Goldman, Alvin (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press. ↩
 9. Beebe, James (2004). The generality problem, statistical relevance and the Tri-level hypothesis. *Noûs*, 38(1): 177–195. Lyons, Jack (2019). Algorithm and parameters: solving the generality problem for reliabilism. *Philosophical Review*, 128(4): 463–509. ↩

10. Stewart, Todd (2005). The competing practices argument and self-defeat. *Episteme*, 2(1): 13–24. ↩
11. Gertler, Brie (2001). Introspecting phenomenal states. *Philosophy and Phenomenological Research* 63(2): 305–28. Feest, Uljana (2014). Phenomenal experiences, first-person methods, and the artificiality of experimental data. *Philosophy of Science*, 81(5): 927–939. ↩
12. Shaver, Robert (2004). The appeal of utilitarianism. *Utilitas*, 16(3): 235–250. ↩
13. Kagan, Shelly (1992). The limits of well-being. *Social Philosophy and Policy*, 9(2): 169–189. ↩
14. Grundmann, Thomas (2021). Dependent reliability. *Philosophy and Phenomenological Research* (forthcoming). ↩
15. Horn, Justin (2017). Moral realism, fundamental moral disagreement, and moral reliability. *Journal of Value Inquiry*, 51(3): 363–381. ↩
16. Jones, Adam (2011). *Genocide: A Comprehensive Introduction*. New York: Routledge. ↩
17. *Deuteronomy* 20.16–17 ↩
18. Street, Sharon (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies*, 127:1, 109–166. ↩
19. Korman, Daniel Z. & Locke, Dustin (2020). Against Minimalist Responses to Moral Debunking Arguments. *Oxford Studies in Metaethics*, 15: 309–332. ↩
20. Kauppinen, Antti (2013). A Humean theory of moral intuition. *Canadian Journal of Philosophy*, 43(3): 360–381. Tappolet, Christine (2016). *Emotions, Values, and Agency*. Oxford: Oxford University Press; Werner, Preston (2016). Moral perception and the contents of experience. *Journal of Moral Philosophy*, 13(3): 294–317. ↩

21. Moore, George Edward (1903). *Principia Ethica*. Dover Publications. ↩
22. Kant, Immanuel (1785). *Groundwork of the Metaphysic of Morals*. Foot, Philippa (1972). Morality as a system of hypothetical imperatives. *Philosophical Review*, 81(3): 305–316. ↩
23. Siegel, Susanna (2010). The contents of visual experience. *Oxford University Press*, New York. ↩
24. Schnall, Simone; Jonathan Haidt, Gerald Clore, and Alexander Jordan (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34(8): 1096–1109. Laham, Simon, Adam Alter, & Geoffrey Goodwin (2009). Easy on the mind, easy on the wrongdoer. *Cognition*, 112(3): 462–466. ↩
25. Hume, David (1739/2000). *A Treatise of Human Nature*. Oxford: Oxford University Press. ↩
26. Borowich, A. (2008). Failed reparative therapy of orthodox Jewish homosexuals. *Journal of Gay & Lesbian Mental Health*, 12(3): 167–177. Fjelstrom, J. (2013). Sexual orientation change efforts and the search for authenticity. *Journal of Homosexuality*, 60(6): 801–827. ↩
27. Díaz-León, Esa (2017). Sexual orientation as interpretation? Sexual desires, concepts, and choice. *Journal of Social Ontology*, 3(2): 231–248. ↩
28. Cf. Smith, Michael (1994). *The Moral Problem*. Oxford: Blackwell. Korsgaard, Christine (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press. ↩
29. Sinhababu, Neil (2017). *Humean Nature*. Oxford University Press, Oxford. ↩
30. Wong, David B. (2006). *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press. ↩

31. Boyd, Richard (1988). How to be a moral realist. In Sayre–McCord (ed.), *Essays on Moral Realism*. Cornell University Press. pp. 181–228. Dowell, Janice (2016). The metaethical insignificance of Moral Twin Earth. In Shafer–Landau (ed.) *Oxford Studies in Metaethics*. Oxford: Oxford University Press. 1–27. ↩
32. Sinhababu, Neil (2019). One–person Moral Twin Earth cases. *Thought*, 8(1): 16–22. ↩
33. Johnson, Michael (2018). Pure quotation and natural naming. *Journal of Philosophy*, 115(10): 550–566. ↩
34. Marshall, Colin (2016). Lockean empathy. *Southern Journal of Philosophy*, 54(1): 87–106. ↩
35. Blumson, Ben (2014). *Resemblance and Representation*. Open Book Publishers. ↩
36. MacAskill, Will (2002). *What We Owe the Future*. New York: Basic Books. ↩
37. Mill, J. S. (1861). *Utilitarianism*. Oxford: Oxford University Press. Driver, Julia (2001). *Uneasy Virtue*. Cambridge: Cambridge University Press. ↩
38. Sinhababu, Neil (2013). Unequal Vividness and Double Effect. *Utilitas*, 25(3): 291–315. ↩