



The Reliable Route from Nonmoral Evidence to Moral Conclusions

Neil Sinhababu¹

Received: 26 February 2021 / Accepted: 28 October 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

We can infer moral conclusions from nonmoral evidence using a three-step procedure. First, we distinguish the processes generating belief so that their reliability in generating true belief is statistically predictable. Second, we assess the processes' reliability, perhaps by observing how frequently they generate true nonmoral belief or logically inconsistent beliefs. Third, we adjust our credence in moral propositions in light of the truth ratios of the processes generating beliefs in them. This inferential route is noncircular, and progress along it is driven fundamentally by induction.

1 Introduction

There is a *reliable route* from nonmoral evidence to moral conclusions. Progress through its three stages relies fundamentally on inductive inference.¹ First, we divide up the psychological processes generating belief so that their reliability in generating true belief is statistically predictable. Second, we measure their reliability – the proportion of true beliefs they generate. Third, we infer probabilities of truth for moral propositions from the reliability of the processes generating belief in them. The three parts of this paper map out the three stages of the reliable route. This first section explains basic concepts that will guide us along the way.

The reliability of a process is measured by its *truth ratio* – the proportion of true beliefs it generates. Truth ratios range from 0 (none true) to 1 (all true).² Reliable

¹ Singer (2005), Greene (2008, 2014), and de Lazari-Radek and Singer's (2012) arguments for consequentialism raise the issue. Tersman (2008), Bruni et al (2014), Kahane (2016), and Rini (2016) share Selim Berker's (2009) objection: "either attempts to derive normative implications from these neuroscientific results rely on a shoddy inference, or they appeal to substantive normative intuitions (usually about what sorts of features are or are not morally relevant) that render the neuroscientific results irrelevant to the overall argument" (294).

² Truth ratios might be hypothetical frequencies, as Hájek (2011) critically discusses.

✉ Neil Sinhababu
neiladri@gmail.com

¹ NUS Philosophy, National University of Singapore, 3 Arts Link, Singapore 117570, Singapore

processes like visual perception generate high proportions of true belief and have high truth ratios; unreliable processes like wishful thinking generate low proportions of true belief and have low truth ratios. While reliabilist theories of knowledge and justification invoke these processes, the reliable route requires no commitment to reliabilism or any other distinctive account of knowledge or justification. Any epistemological theory supporting induction will support it.

Individuating processes so that their truth ratios are statistically predictable permits *truth-ratio induction*, in which the truth ratios of some beliefs inductively suggest truth ratios for other beliefs generated by the same process. The three stages of the reliable route are steps in implementing truth-ratio induction. First, we individuate processes so that their truth ratios can be inductively projected from observed cases to unobserved cases. Second, we determine the processes' truth ratios by observing them under various conditions. Third, we calculate probabilities of truth from the processes' truth ratios. In the first two stages, we follow empirical evidence to propositions of the form "Process P, which generates belief in moral proposition M, has truth ratio T". In the third, we adjust credence in M as T suggests.³

The truth-values of moral beliefs can't be assumed when determining the truth ratios with which they're generated, or the reliable route would be circular. Fortunately, there are at least two ways to assess truth ratios without such assumptions. First, the truth ratio of a process can be inductively inferred from the truth ratios of nonmoral beliefs it generates. Second, contradictory moral beliefs push processes generating them towards unreliability, regardless of which belief is true. These methods require only nonmoral assumptions. With the processes' reliability determined, truth-ratio induction provides probabilities of truth for moral beliefs they generate. Induction thus carries us across the inferential gap between nonmoral *is* and moral *ought* noted by David Hume (1740/2000).⁴

We need not assume that moral terms have definitions in nonmoral terms, or that moral properties are physical, causal, empirically observable, or otherwise natural.⁵ If true, such assumptions might provide helpful shortcuts to the moral truth. For example, if "good" can be defined in terms of empirically observable natural things, clarifying and applying the definitions might reveal what is good. GE Moore (1903) warns that such definitional shortcuts lead nowhere, as proper definitions leave substantive moral questions wide open. If definitions of moral terms contradict some widely accepted moral theories, many people misunderstand definitions in their own

³ As Walter Sinnott-Armstrong (2008b) describes, the idea is to inductively "apply information about the reliability of general classes to particular beliefs within those classes" (100). This general insight can be preserved even granting Demaree-Cotton's (2016) criticisms of the framing intuitions he invokes. This paper grounds truth-ratio induction in contemporary theories of process-individuation, giving it a general form that encompasses a wider variety of empirical discoveries, and showing how it might lead to conclusions from error theory to specific realist theories as the empirical data suggest.

⁴ Prior (1960) and Karmo (1988) present valid deductive inferences from *is* to *ought*. These raise issues about the *is-ought* gap discussed by Pigden (1989), Schurz (1997), Guevara (2008), Russell (2010), Brown (2014, 2015), and Singer (2015). Huemer (2005) notes that these inference schema don't provide substantive moral knowledge. The reliable route does.

⁵ Railton (1989), Brink (1989).

languages, about which they should be highly reliable. Definitional shortcuts therefore won't be used here, though their advocates are encouraged to follow Frank Jackson (1998) in exploring whether they offer more direct paths to the moral truth.

The reliable route isn't widely known, but its underlying philosophical assumptions are widely shared. Inferential progress from nonmoral evidence to moral conclusions is driven fundamentally by inductive inferences validated by all standard epistemological theories, as Sect. 1.3 explains. Truth-ratio induction gains power if moral and nonmoral belief are generated by unified processes with unified reliability, and as Sect. 2.2 explains, standard cognitivist metaethical theories are developed with care to preserve this unity.⁶ The reliable route proceeds from nonmoral evidence to moral truth on solid ground – the common ground of our best theories.

There are reliable routes to nonmoral truths as well. The first two stages can take us to the more general proposition "Process P, which generates belief B, has truth ratio T". B might concern nonmoral normativity, or something altogether non-normative. The nonmoral norms of law and etiquette are plausibly entailed by the structure of our social practices, making them easier to discover. It's unclear whether Moore's warning against definitional shortcuts applies to normative truths concerning prudential or epistemic value, so perhaps they too can be discovered more easily. Discovering the moral truth may be harder. This paper seeks it in the same spirit that Kennedy sent the astronauts to the moon: "not because it is easy, but because it is hard".

1.1 Identify Processes that Support Truth-Ratio Induction

How should process types be individuated? This is the generality problem bedeviling reliabilism.⁷ Instances of belief-formation fall under many process types with different truth ratios. Suppose Rene sees that Elisabeth sent him a letter. Is his belief that he has a letter generated by visual perception, a reliable process with a high truth ratio? Or is it generated by visual-perception-of-letters-or-wishful-thinking, an unreliable process with a low truth ratio?

This section offers tools for individuating processes so that their truth ratios will be statistically predictable, making it possible to use truth-ratio induction. Section 1.2 argues that differences in belief-content don't block truth-ratio induction from the nonmoral to the moral. Section 1.3 shows that internalist opponents of reliabilism can use truth-ratio induction just as reliabilists can.

Reliabilists suggest many potential ways of individuating processes.⁸ Two recent innovations are especially helpful in ensuring that processes will have the statistically predictable truth ratios needed for truth-ratio induction. James Beebe (2004)

⁶ The reliable route may presuppose that moral judgments are beliefs. Other mental states including desire and emotion might not require evidential support or be true or false. Questions about how to infer moral truths from nonmoral evidence will then presuppose some form of cognitivism, accepted by naturalistic realists like Railton (1989), non-naturalists like Enoch (2011), constructivists like Street (2010), and error theorists like Mackie (1977). Noncognitivists like Stevenson (1937) differ.

⁷ Conee and Feldman (1998). Joyce (2006) and Kahane (2016) address moral epistemology.

⁸ Goldman (1986), Sosa (1991), Heller (1995), Becker (2008), Lepock (2009).

and Samuel Kampa's (2018) *statistical solution* to the generality problem lets processes support truth-ratio induction as natural kinds do. Jack Lyons' (2019) *algorithm & parameters theory* predicts truth ratios for beliefs formed under a range of different conditions.

Beebe and Kampa's statistical solution treats process types as the broadest unified causal structures with statistically predictable truth ratios.⁹ Their emphasis on causal structure treats process types like natural kinds.¹⁰ Because instances of a natural kind share causal structure, observing some instances inductively supports predictions about other instances.¹¹ Samples of gold share causal structure at the atomic level, explaining their melting points. So seeing some samples melt at 1064 °C inductively suggests that other samples melt at 1064 °C. Instances of visual perception similarly share causal structure, so observing a high truth ratio for some instances inductively suggests a high truth ratio for others. Causal structure is an especially useful criterion when the truth ratios of processes are unknown. This may often happen with processes generating many moral beliefs, whose truth-values can't be assumed without circularity. Empirically investigating how these processes generate beliefs allows us to classify them under a process type sharing their causal structure.

Natural kinds support induction, and correctly individuated processes support truth-ratio induction. Since process types are for predicting which beliefs are true, the best theory of processes for present purposes is the one that best predicts which beliefs are true. Trial and error in using truth-ratio induction will aid progress towards this theory, as accounts of natural kinds are confirmed by supporting successful inductive inferences and disconfirmed by supporting unsuccessful ones. This sort of second-order induction – observing whether a first-order inductive system generates correct predictions – generally helps to distinguish natural kinds (understood here as whatever inductively projectible predicates refer to). As WVO Quine (1969) writes, “We newly establish the projectibility of some predicate, to our satisfaction, by successfully trying to project it. In induction nothing succeeds like success” (129).¹²

The truth ratio of visual perception varies. It's unreliable in darkness and blind-light, and with small or distant objects. Truth-ratio induction from its successes shouldn't suggest that it's reliable in darkness, and truth-ratio induction from darkness shouldn't suggest that it always fails. Lyons' algorithm & parameters theory

⁹ “Unified causal structures” and “statistically predictable truth ratios” are informal terms for the properties the statistical solution uses to individuate processes. Causal structures are individuated by the computational problem solved and the algorithmic structure used, two levels of computational neuroscientist David Marr's (1982) influential *tri-level hypothesis*. Beebe excludes Marr's third level, physical implementation, so processes will be multiply realizable. Dutant and Olsson (2013) object that Beebe's statistical relevance criterion generates spurious processes with only true belief or only false belief. Kampa addresses this by emphasizing unified causal structure, which the spurious processes lack. He lets external factors individuate computational processes but not algorithmic processes, following Dawson (2013). I invoke the statistical solution because it metaphysically explains how truth-ratio induction works.

¹⁰ Alston (1995) suggests that process types are natural kinds.

¹¹ Quine (1969).

¹² Following Goodman (1965).

helps avoid these bad inductive inferences. Parameters such as brightness, object size, and distance are inputs to visual perception, which has a belief as its output. Algorithms describe how the parameters generate the belief, and can be realized by the causal structures invoked by the statistical solution. Visual perception has parameters including the brightness of the light and the distance of the object. Both darkness and blinding light can make viewers form beliefs that correspond poorly with reality, rendering visual perception unreliable under these conditions. Only between these extremes is visual perception reliable. Extrapolating with these parameters lets us predict processes' future reliability. Suppose visual perception of frogs at 20 m is reliable in daylight, but not in moonlight. Will it be reliable on a moonless night, or in total darkness? No, because more darkness doesn't improve reliability.

The algorithm & parameters theory suggests understanding statistical predictability in terms of a single function modeling a processes' truth ratio, with each parameter in the algorithm being a variable in the function. If the algorithm for visual perception treats brightness, distance, object size, and attentional focus as parameters explaining belief-generation, the function modeling its truth ratio will include each of these parameters.¹³ Where one function best models truth ratios, there is one process. Where many functions do, there are many processes. Visual perception can be described as typically or generally reliable because typical parameter settings make it reliable, even if it's unreliable under atypical settings.

Both the statistical solution and the algorithm & parameters theory explain why wishful-thinking-or-visual-perception-of-letters isn't a genuine process type. Visual perception and wishful thinking differ greatly in causal structure. Different functions with different parameters model their truth-ratios, with one function having a brightness parameter and the other perhaps having a desire-strength parameter. The processes are therefore different.

1.2 Truth-Ratio Induction Invokes Parameters, not Content

If a process reliably generates true nonmoral belief, that's inductive evidence that it reliably generates true moral belief.¹⁴ This section explains why these inductive inferences are valid despite differences in the contents of the beliefs. First I'll illustrate how truth-ratio induction across different contents helps in avoiding error and making scientific discoveries. Then I'll explain why parameters explain reliability rather than content, which differs between moral and nonmoral propositions. Invoking parameters generates better predictions, and helps the reliable route avoid circularity.

To see how truth-ratio induction across different contents helps us avoid error, consider Ed, who knows that wishful thinking leads him mostly to false beliefs about

¹³ Lyons treats processes as having specified parameter values rather than variable values. His approach makes visual perception a group of processes like visual-perception-in-direct-sunlight-at-19-m. Present purposes permit externalism about parameter values.

¹⁴ Berker (2009) expresses misgivings about such inferences.

topics including sports, investing, and love. Ed has paid to have his head cryogenically frozen after death, expecting an advanced future society to reanimate him and extend his life. His friends tell him that this is just wishful thinking. Ed accepts this and says, “I know wishful thinking is unreliable about other things, but it’s never led me wrong about cryogenically freezing my head! So there’s no evidence that wishful thinking is unreliable about that”. If truth-ratio induction didn’t work across different contents, Ed would be right to distinguish his belief from others with different contents. But he’s clearly wrong.

To see how truth-ratio induction across different contents leads us to scientific discovery, consider this objection to Galileo’s use of the telescope to discover Jupiter’s moons: “The telescope’s reliability in detecting distant terrestrial objects doesn’t entail or even suggest reliability in detecting distant celestial objects. The terrestrial and the celestial are very different domains, and reliability in one doesn’t suggest reliability in another”. If this objection had been generally accepted, Galileo’s breakthrough would’ve come to nothing.

As astronomy after Galileo shows, processes observed to be reliable should be trusted with new contents even before their reliability with these contents is independently verified. Spacecraft and other technologies to verify telescopes’ reliability about celestial matters wouldn’t be developed for centuries. But even in 1676, accepting Galileo’s findings and the telescope’s reliability produced the first good estimates of the speed of light. Using a telescope to observe an eclipse of Jupiter’s moon Io, Ole Rømer measured the speed of light as 214,000 km/s.¹⁵ This remained the most accurate measurement for a century, coming within 30% of the actual speed of light. Rømer’s discovery required assuming the telescope’s reliability about Jupiter’s moons before independent verification was available, even while widely trusted processes such as scriptural exegesis contradicted it. Independent verification is of course valuable, and should raise credence that a process is reliable in the new domain. But requiring independent verification before admitting evidence about new domains would nullify the evidential value of scientific breakthroughs.

Individuating processes independently of content gets the right answers about Ed and Galileo. Wishful thinking causes beliefs about many topics by the same desire-driven causal mechanism, with dismal reliability. So Ed’s wishful thinking about cryogenic freezing belongs to an unreliable process type. Telescopes employ the same optical causal mechanism in both terrestrial and celestial applications, with high reliability. So celestial telescope use belongs to a process type with unified causal structure and a high truth ratio. Truth-ratio induction tells us that Ed’s belief is probably false, while Galileo’s belief is probably true.

One might try to address these cases by permitting truth-ratio induction only between appropriate content (or subject matter) types. The aim would be to permit truth-ratio induction from the unreliability of wishful thinking and the reliability of the telescope, while blocking other inferences. But the appropriate content-types are hopelessly disunified. Addressing Ed’s case requires grouping together sports, investing, love, and cryogenic freezing. Addressing Galileo’s case requires grouping

¹⁵ Rømer (1677).

together terrestrial things like ships, lighthouses, and clouds with celestial things like moons and planets. These diverse collections of things don't fit into well-unified kinds. As Bob Beddor (2014) notes, individuating these content-kinds throws us back into the generality problem.

While telescopes aren't reliable about nearby things, this isn't best explained by listing all the nearby things that could figure in the content, and treating telescopes as unreliable in generating beliefs with content about the floor, bees, Galileo's beard, and so on. This explanation lacks unity and generality, and violates simplicity with its endless list of contents. The better explanation is that telescopes are reliable about faraway things and not nearby things. This explains why telescopes are reliable about Jupiter's moons, as well as farther things like Neptune and nebulae which weren't contents of belief in Galileo's time.

Invoking parameters rather than content generates better predictions about reliability. Visual perception is usually unreliable about the presence of bacteria, as they are very small. So one might take visual perception to be unreliable in generating beliefs with content involving bacteria. But a better explanation of unreliability invokes the low value of the size parameter, providing a unified explanation of our unreliability in seeing bacteria, viruses, and molecules. Invoking the low value of the size parameter supports extrapolation, explaining why visual perception is unreliable about smaller things like atoms. The size parameter also explains why visual perception is more reliable about the unusually large bacterium *thiomargarita namibiensis*, which can be 0.75 mm if its internal gas bubble is fully inflated. Content doesn't explain this, since visual perception isn't usually reliable about bacteria.

Leading reliabilist approaches to the generality problem don't give content any fundamental role in distinguishing processes. Lyons notes that cognitive science generally explains belief-formation using broad parameters and not specific subject matters or contents, so that "differences of mere content don't ipso facto make for differences of process" (487).¹⁶ When a process has different reliability with different contents, "it's general parameters, rather than particular contents, that are doing the work" (489, also 472–473).

To avoid circularity, the reliable route must begin from nonmoral premises. Starting from empirical observations of how parameter values affect reliability avoids circularity. This is true even if parameters and content happen to be identical, and are therefore perfectly correlated. "That's water, and water is water, so that's water" is circular. "That's H₂O, and H₂O is identical to water, so that's water" is noncircular.

1.3 Everyone Can Use Truth-Ratio Induction, not Just Reliabilists

While reliabilists did foundational work in developing the reliable route, it requires no commitment to reliabilism. Opponents of reliabilism can use it too.

¹⁶ Goldman (2000) notes that "our ordinary thought about process types slices them broadly" (612–613). Becker (2007) agrees.

Despite rejecting reliabilism, Earl Conee (2013) makes room for truth-ratio induction within his own internalist framework. He explains how:

each justified reliability proposition asserts that the corresponding believed proposition has a certain contingent property that is shared mostly by truths. Having justification for those propositions that place the target beliefs in contingent classes that contain mostly truths gives the believers routine inductive reasons to think that the target beliefs are true. (761)

These “routine inductive reasons” justify truth-ratio induction. Conee and other epistemic internalists can happily employ truth-ratio induction, once given evidence about reliability. Knowing that a process generates a sufficiently high proportion of false belief allows inductive inference that other beliefs it generates are unlikely to be true. Truth-ratio induction is just induction applied to processes. All plausible epistemological theories permit induction. Processes individuated for statistical predictability therefore support truth-ratio induction.

Conee and other epistemic internalists are right to make room for truth-ratio induction. Doing so lets them draw the right conclusions about Ed and Galileo. Truth-ratio induction extends the general unreliability of wishful thinking to Ed’s belief about freezing his head, and extends the general reliability of the telescope to Galileo’s astronomical discoveries. Internalism would face devastating objections if it couldn’t support induction, and reliability is just another thing to which induction applies. As Juan Comesaña (2006), Michael Bishop (2010), and Jeffrey Tolly (2017) explain, the generality problem is a practical problem that must be solved for us to make the inductive inferences that we must be permitted on any epistemological view.

The reliable route bypasses issues about the fundamental nature of justification and knowledge where epistemic internalists and externalists disagree. Epistemologists dramatize these issues with cases of people mistaken about the reliability of their belief-forming processes, including Lawrence Bonjour’s clairvoyant Norman (1980), and the victims of Stewart Cohen’s New Evil Demon (1984). We need not judge these epistemic unfortunates, or accept any theory about whether their false beliefs are justified. Instead, we should avoid their unfortunate condition by scientifically discovering which of our belief-forming processes are reliable.¹⁷ This will require knowing more about the causal structure of the mind and the truth-values of our nonmoral beliefs. Scientific progress in many disciplines will generate this knowledge. By harnessing this scientific progress, we can make moral progress.

¹⁷ Invoking reliability allows more direct connections to truth than invoking justification as Sinnott-Armstrong does (2008a; 2008b). Truth ratios are probabilistic evidence about the truth of the beliefs a process generates. Examining them additionally confers justification, as changing one’s beliefs on the basis of such probabilistic evidence is justified.

2 Assess the Truth Ratios of the Processes Generating Moral Belief

The second step is to assess the truth ratios of the processes generating moral belief. We must do so without normative ethical assumptions.¹⁸ Otherwise we'd merely travel in a circle from ought to ought. You were promised a more exciting journey.

I'll suggest two noncircular methods for assessing truth ratios. First, if a process generates both moral and nonmoral beliefs, truth-ratio induction lets us inductively project the truth ratio of the nonmoral beliefs onto the process as a whole. Section 2.1 discusses this method, and Sect. 2.2 explains why standard cognitivist theories suggest that moral and nonmoral beliefs are generated with the psychological and epistemological unity it requires. Second, if a process generates inconsistent beliefs, at least one belief in each inconsistent set must be false, pushing the process towards unreliability.¹⁹ Section 2.3 discusses this method.

My examples of how to implement these two methods include psychological claims about the processes generating our beliefs. Those who reject these psychological claims can still accept the in-principle possibility of using the methods. The main purpose of the psychological claims is to illustrate the form of the methods, rather than to defend distinctive psychological or ethical theses.

There may be additional methods for assessing reliability without normative ethical assumptions. For example, evolutionary debunking arguments might establish the unreliability of some processes, as evolutionary considerations favor fitness-enhancing processes over reliable ones.²⁰ The more ways there are to assess reliability without moral assumptions, the more we discover at the conclusion of the reliable route.

2.1 Use Truth-Ratio Induction from Nonmoral Beliefs

If a process generates both nonmoral and moral beliefs, and our observations provide a truth ratio only for nonmoral beliefs, truth-ratio induction suggests that this is the truth ratio of the process as a whole.

Truth-ratio induction from the nonmoral to the moral won't work if different processes generate moral and nonmoral beliefs.²¹ For example, sensory processes generating beliefs about concreta might generate none about abstracta. Non-naturalists who regard moral properties as abstract may then argue that the reliability of sensation in generating nonmoral belief doesn't apply to moral belief. This non-naturalist

¹⁸ O'Neill (2015) argues that processes are unreliable if they generate moral belief in ways that "are inconsistent with plausible general features of moral facts", arguing for example that the associative tendencies explaining what disgusts us don't track moral disvalue (1074). Rini (2016) argues that all extant attempts at selectively debunking moral judgments depend on such assumptions. The reliable route avoids such assumptions.

¹⁹ The first and second stages of the reliable route may in practice overlap, as empirical evidence about belief-formation is integrated into theories of processes and their reliability.

²⁰ Street (2006), Joyce (2001). Joyce (2006) addresses the issue in terms of reliability.

²¹ This would make truth-ratio induction an illegitimate conflation of distinct natural kinds. Kahane (2016) describes confluations.

argument against the reliability of sensation in generating true moral belief is valid, though it requires the controversial premise that moral properties are abstract.

Truth-ratio induction still applies to non-naturalist epistemology, as demonstrated by non-naturalists who argue that since rational intuition is reliable about mathematical and modal abstracta, it's reliable about moral abstracta. These “companions in guilt” arguments have the structure of truth-ratio induction from reliably caused mathematical and modal beliefs to moral beliefs.²² As this example shows, and as the next section explains in detail, applying truth-ratio induction to moral belief requires no special ontological assumptions about moral properties.

Which processes might generate both moral and nonmoral belief, and allow us to use truth-ratio induction from the nonmoral to the moral? Here are two examples, one unreliable and the other reliable. While the psychological hypotheses invoked in both examples require further empirical defense, they at least serve to illustrate how processes generating moral belief can be assessed on the basis of the nonmoral beliefs they generate.

Intuitive reasoning with large magnitudes is unreliable, and seems to affect moral beliefs about large amounts of harm. Cognitive scientists call this error “compression”. Resnick et al (2017) write, “people possess compressed representations of the relatively larger, unfamiliar magnitudes. Compression refers to the observed pattern of overestimation of relatively smaller magnitudes and underestimation of relatively larger magnitudes” (1022). Experimental subjects often put one million halfway between one thousand and one billion on number lines, rather than very close to one thousand. They make accordingly erroneous estimates of proportional relations involving large numbers.²³ It's a wry adage of political budgeting that all numbers ending in “illion” sound the same to voters.²⁴ Human quantitative reasoning inaccurately compresses millions, billions, and trillions into poorly distinguished “illions” rather than recognizing vast differences between them.

Moral judgments about large amounts of harm display similar features. Desvoves et al. (2010) asked three different groups of participants how much they'd pay to save large numbers of seabirds from dying in an oil spill. Those asked about saving 2000 seabirds offered a median sum of \$80. Those asked about saving 20,000 seabirds offered \$78. Those asked about saving 200,000 seabirds offered \$88. Just as people often don't distinguish large magnitudes in nonmoral cases, they often don't distinguish the moral weights of large magnitudes of bad events. Cognitive scientists call this phenomenon “scope insensitivity”.²⁵ Since compression and scope insensitivity both involve nonresponsiveness to increasing magnitudes, a single process with a size-of-number parameter explains their unreliability with unity and simplicity. Large numbers make this process unreliable in the nonmoral case of compression. Truth-ratio induction then suggests the same in the moral case of scope insensitivity.

²² Mackie (1977) anticipates such arguments, which Huemer (2005) delivers.

²³ Dehaene (1997), Siegler and Opfer (2003).

²⁴ Miller (1978).

²⁵ Chang and Pham (2018), Fetherstonhaugh (1997), Slovic (2010).

For a case where high reliability in generating nonmoral belief suggests high reliability in generating moral belief, consider phenomenal introspection, which generates beliefs about what one's conscious experience is like. As I write my paper on my computer while listening to music, phenomenal introspection generates belief that I'm having color-experiences of varying brightness, and sound-experiences of varying volumes. While other forms of introspection are less reliable, phenomenal introspection is widely agreed to be highly reliable.²⁶

Some ethicists hold that phenomenal introspection generates belief that pleasure is good. Shelly Kagan (1998) claims that "The value of pleasure and the disvalue of pain seem virtually self-evident to anyone experiencing them" (30). Then the same process telling us that color-experiences have brightness and sound-experiences have volume will also tell us that pleasure-experiences have goodness and displeasure-experiences have badness.²⁷ Phenomenal introspection can access all these experiences and their intrinsic features. If belief in pleasure's goodness indeed comes from phenomenal introspection, induction suggests that it's reliably generated.

The reliability of rational intuition, assessing large magnitudes, and phenomenal introspection can be inductively projected from observed cases to unobserved cases. Their reliability with moral belief should be equal to their reliability with similarly generated nonmoral belief. It's conceivable that induction fails here because hidden parameters or process distinctions line up with the moral/nonmoral content distinction.²⁸ But if you bet against induction because its failures are conceivable, you'll lose more than you win.

2.2 The Unity of Psychology and Epistemology Across the Moral and Nonmoral

Truth-ratio induction from the nonmoral to the moral requires two assumptions. The first is psychological: the processes generating moral belief also generate nonmoral belief. The second is epistemological: the processes' reliability in moral cases is statistically predictable from nonmoral cases. As I'll explain, cognitivist metaethical theories are typically built to support these assumptions. Non-naturalists, constructivists, error theorists, and naturalistic realists all accept that the same processes generate beliefs with both moral and nonmoral content, with nonmoral reliability predicting moral reliability. This helps them avoid the implausibility of giving moral belief-formation an unprecedented and *sui generis* psychological and epistemological status.

²⁶ Brentano (1969). Even Schwitzgebel (2008), a rare skeptic about phenomenal introspection, concedes its reliability in cases like severe pain. Feest (2014) further supports its reliability. As Wu (2018) discusses, introspection is essential to neural research on consciousness.

²⁷ Shaver (2004), Rawlette (2014), Broi (2022).

²⁸ Gray and Wegner (2010) argue that a process of agent-detection correctly identifies human agents and creates untestable (presumably false) belief in divine agents. If processes differ similarly with nonmoral and moral belief, truth-ratio induction leads us astray. But getting Ed and Galileo's cases right requires getting the agent-detection case wrong. Epistemically limited beings like ourselves must live and die by induction.

Non-naturalists regard moral properties as abstract and non-natural, and invoke processes also generating nonmoral belief to explain beliefs about them. Robert Audi (2013) invokes perception, suggesting that its reliability in generating true beliefs about natural properties generalizes to the non-natural moral properties supervening on them. Michael Huemer (2005) invokes rational intuition, suggesting that its reliability in generating true beliefs about non-natural mathematical and modal properties generalizes to similarly non-natural moral properties. Audi and Huemer identify processes reliably generating true nonmoral beliefs, and take this to suggest reliability in producing true moral beliefs.

Constructivists regard moral properties as metaphysically constituted by our evaluative attitudes and their objects, much as deliciousness is constructed from our tastes and favorite foods, and as beauty is in the proverbial eye of the beholder. David Hume (1740/2000) claims that we can identify moral properties this way: “virtue is distinguished by the pleasure, and vice by the pain, that any action, sentiment or character gives us by the mere view and contemplation”. He suggests that in tracking our attitudes, these feelings are reliable about morality, as they are about other attitude-dependent value: “The case is the same as in our judgments concerning all kinds of beauty, and tastes, and sensations. Our approbation is implied in the immediate pleasure they convey to us” (3.1.2).

Error theorists claim that there are no moral properties, rendering all moral belief false. They often support their view by invoking unreliable processes generating all moral belief. John Mackie (1977) invokes a socially inculcated form of “the tendency to read our feelings into their objects. If a fungus, say, fills us with disgust, we may be inclined to ascribe to the fungus itself a non-natural quality of foulness” (98). Mackie can use truth-ratio induction to argue that the same process generating this nonmoral error will also generate moral error. Friedrich Nietzsche (1889/1968) claims that “there are altogether no moral facts. Moral judgments agree with religious ones in believing in realities which are no realities”. He describes unreliable social and psychological processes explaining moral and religious errors at once.

Naturalistic realists regard moral properties as natural properties. They think reliable empirical methods can reveal which natural properties they are. The Cornell Realists invoke forms of inference to the best explanation used by natural and social scientists. David Brink (1989) argues that the injustice of apartheid can be discovered by inference to the best explanation of social phenomena like protest and regime instability. Many processes are reliable about natural properties, and naturalistic realists can search among them for processes that generate moral beliefs. Phenomenal introspection may be such a process, as the final section of this paper considers.

Advocates of all these metaethical theories have good reasons to psychologically and epistemologically unify moral and nonmoral belief-formation. Explaining moral belief-formation in terms of processes not attested elsewhere invites psychological objections that the human mind simply doesn't have those processes. Anyone claiming that ESP, radar, or clairvoyance generates some of our beliefs will face objections that the human mind doesn't do these things. Invoking processes known to generate nonmoral belief avoids these objections.

Similarly, extending processes' truth ratios from nonmoral to moral cases avoids conflict with observations of the processes' reliability. Ed fails to achieve such consistency in defending wishful thinking as reliable about cryogenic freezing, as do those who deny the reliability of the telescope regarding space. Plausible meta-ethical theories suggest that the same processes generate both moral and nonmoral belief with broadly unified reliability, supporting truth-ratio induction from the nonmoral to the moral.

2.3 Find Jointly Inconsistent Moral Beliefs, Which Reduce Reliability

If a process generates jointly inconsistent combinations of moral beliefs, it must generate some false beliefs, pushing its truth ratio towards unreliability.²⁹

A process generating a belief that genocide is impermissible and a belief that genocide is permissible generates at least one false belief, as these beliefs are inconsistent. This reduces the upper bound on its truth ratio. It raises the lower bound on its truth ratio as well, since at least one belief must be true. But the net effect is to push the process towards unreliability. In the extreme case where a process generates equal numbers of beliefs in two inconsistent propositions, one true and one false, it's no better than the paradigmatically unreliable process of having one's beliefs determined by the flip of a coin.

Many processes generate inconsistent beliefs interpersonally. Wishful thinking often leads fans of rival sports teams to inconsistent beliefs about who will win. Narcissistic self-regard may lead many people to each believe that they are the sole deserving recipients of an award. Any process causing belief in different religions generates true belief in at most one religion, and false beliefs in all others. This suggests the unreliability of wishful thinking, narcissistic self-regard, and any process responsible for most religious belief. Since at least one party to any disagreement must be in error, processes generating widespread disagreement must generate widespread error.

There seems to be a great deal of interpersonally inconsistent moral belief, historically and cross-culturally. The example of moral disagreement about genocide wasn't idly chosen. Belief that genocide is permissible or even obligatory is historically prevalent from the Old Testament to colonial empires to recent ethnic nationalists.³⁰ The anti-genocide consensus of today's educated people starkly disagrees with many previous cultures. There is great trans-historical moral disagreement on issues including slavery, gender norms, punishment, sexual behavior, the treatment of animals, and obligations to aid those who suffer.³¹ When investigating the general

²⁹ Sinnott-Armstrong's (2008a) example of someone whose judgments about whether a man is short or tall differ depending on context illustrates the principle. Inconsistencies between framing intuitions suggest arguments with this structure. Doris and Plakias (2008a; 2008b) and Leiter (2008) provide support for similarly structured arguments from disagreement.

³⁰ Jones (2011) provides a comprehensive historical survey. Van Wees (2010) is also helpful.

³¹ While arguments from disagreement have received considerable discussion – see Mackie (1977), Huemer (2005), and Enoch (2009) – few except Stewart (2005) and Horn (2017) note this chain of reasoning from disagreement to error to unreliability.

reliability of human belief-forming processes, beliefs across all human cultures for all of human history must be assessed.

Some moral disagreements result merely from fundamental moral agreement combined with nonmoral disagreement. For example, those who share a moral theory might disagree about whether a medical treatment is morally required, if they disagree about whether it is safe. Errors within such moral disagreements reduce the reliability of processes generating nonmoral beliefs, not moral beliefs.

In other cases, false nonmoral beliefs may result from false moral beliefs. Nonmoral beliefs about the mental incapacity of women, ethnic minorities, and lower social classes often arise not as innocent scientific mistakes, but to rationalize a sense that these groups are morally inferior and don't deserve equal treatment.³² Nonmoral beliefs favoring genocide may similarly rationalize ethnic hatred. These false nonmoral beliefs provide no defense for the reliability of the processes generating the moral beliefs.

If widespread inconsistency renders a process unreliable, truth-ratio induction suggests abandoning beliefs it generates where inconsistency is absent, since they're also likely to be mistaken. If a process responsible for inconsistent religious belief has caused me to believe something so distinctive that only my religion explicitly addresses it, I should abandon the belief. Even if Hinduism never explicitly denies the virgin birth of Christ, truth-ratio induction tells Christians to abandon this doctrine if it's generated by the same process as many Hindu beliefs that they reject. If a unified process generates all religious belief, and great disagreement on a number of topics renders it unreliable, agreed-upon religious beliefs such as belief in an after-life are undermined as well. Christians might therefore deny that their beliefs issue from the same process as Hindus. Doing so would reflect their realization that if inconsistency demonstrates that a unified process is unreliable, this unreliability will carry over to narrow areas where the process doesn't generate inconsistency.³³

Better understanding the psychological processes generating inconsistent beliefs will help in drawing conclusions about reliability. The causal mechanisms generating moral belief are vigorously debated by philosophers and psychologists. We should expect similarly vigorous empirical debate about how these causal mechanisms should be grouped into processes assessable for reliability. If widespread disagreement about sexual morality results from differences in which sexual practices disgust people, is the unreliable process disgust about sexual practices, disgust in general, or emotion as a whole?³⁴ Careful empirical research and reasoning is needed to individuate processes and identify their parameters.

³² Hoffman and Hurst (1990), Greenwald and Banaji (1995).

³³ These arguments are independent of Elga (2007) and Christensen's (2007) conciliationism about peer disagreement, and compatible with Kelly's (2005) steadfast view. Disagreement with one peer about one proposition doesn't significantly reduce process reliability because it entails only one false belief, and those who disagree can be demoted from peerhood as Vavova (2014) describes. As Decker and Groll (2013) argue, peer disagreement doesn't motivate moral skepticism. Stewart (2005) helpfully describes how widespread disagreement about many propositions can entail enough error to undermine reliability.

³⁴ Kelly (2011) argues that disgust causes such belief. May (2014) replies.

While more nonmoral information may be needed, using inconsistency to assess the reliability of processes generating moral belief doesn't require moral assumptions. It requires only the psychological assumption that people have inconsistent moral beliefs caused through the same process. Knowing that one of two inconsistent beliefs is false doesn't require knowing which is false. Either way, inconsistency entails error, pushing processes towards unreliability.

3 Revising Beliefs to Discover Moral Truths

This section explains why evidence of the reliability with which moral beliefs were generated might suggest revising them. Section 3.1 addresses how evidence of reliability interacts with prior probabilities of truth for moral propositions. Section 3.2 illustrates how the reliable route might lead to striking moral conclusions.

Individuating the processes generating moral belief and assessing their reliability, as discussed in previous sections, results in propositions of this form:

Process P, which generates belief in moral proposition M, has truth ratio T.

As T is a function, it will be computed with the parameter values filled in. If T is higher than expected, credence in M should rise. If T is lower than expected, credence in M should fall. Even in the absence of explicit expectations regarding process reliability, propositions of the above form can motivate revisions. Suppose one believes that M, discovers that the truth ratio of the process generating belief that M is 0.1, and has no other evidence bearing on the truth of M. Conjoining M with the probability suggested by its truth ratio entails:

M, and there's an 0.1 probability that M.

One shouldn't believe such a proposition or anything entailing it. It's irrational, perhaps because it's Moore-paradoxical or probabilistically inconsistent. If one's evidence supports the probability assignment, one should give up belief that M.

David Lewis' (1986) Principal Principle suggests the same conclusion. According to Lewis, one's subjective probability in a proposition should equal its objective probability, if one has no other evidence bearing on its truth.³⁵ After learning that a coin has a 0.5 objective probability of heads, one's subjective probability of heads should be 0.5. The odds of getting particular results from coin-flips and other chancy events are naturally understood as objective probabilities. Processes of belief-formation can be seen this way, with their truth ratios being objective probabilities of truth for beliefs they generate. Lewis' Principal Principle then suggests matching our subjective probabilities to their objective probabilities – that is, our credences to their truth ratios.³⁶

It's hard to justify maintaining beliefs generated by unreliable processes. Even the epistemic conservative Roderick Chisholm (1989) allows evidence of unreliability

³⁵ Hájek (2011) formalizes the principle as " $C(A | ch(A)=x)=x$." C represents credence. $ch(A)$ represents the objective probability of A, with ch signifying "chance".

³⁶ I see sufficiently significant revisions in credence as constituting belief-revisions, following the Lockean thesis of Locke (2014). See also Jackson (2018).

to reduce credence. Otherwise he'd have to defend Ed, who maintains his belief in the efficacy of cryogenically freezing his head despite acknowledging that it's wishful thinking. Evidence about processes' truth ratios is evidence of the truth of the beliefs generated. Moral beliefs should be revised in light of evidence of their truth, just like any other beliefs.

Changing credence in moral propositions on the basis of evidence concerning the reliability of the processes generating them is reasoning from nonmoral evidence to moral conclusions. This reasoning is fundamentally inductive. Hume's gap between *is* and *ought* can thus be crossed on the power of induction alone.

3.1 Process Reliability and Prior Probabilities

How should discoveries about process reliability affect prior probabilities in moral propositions? In general, lower-than-expected reliability suggests reducing credence, while higher-than-expected reliability suggests raising credence. In the simplest cases, the scarcity of evidence for moral propositions suggests shifting prior probabilities of truth for moral propositions towards the truth ratios of processes generating them. When a single process generates beliefs of different logical strength, credences in these propositions should differ as logical reasoning or complexity-related parameters suggest. When prior evidence includes a belief generated by a reliable process, and another process generates the same belief less reliably, the truth ratios of both processes should be aggregated to determine the correct credence.

Suppose one discovers that a process with truth ratio T generated one's belief in moral proposition M . The less additional evidence one has, the closer one's credence in M should go towards T . Ed's high credence in the efficacy of cryogenically freezing his head should decline towards the low truth ratio of wishful thinking, when he realizes that it's the sole process generating his belief and that he has no additional evidence. If Galileo gives his daughter Livia the telescope to see Jupiter's moons and she asks whether it's reliable about celestial bodies, her credence that Jupiter has moons should rise towards whatever truth ratio Galileo's answer suggests. These examples stipulate that Ed and Livia's only putative evidence comes from a single process – wishful thinking or telescope use – which generated their prior probabilities. When their evidence suggests a truth ratio for that process, these prior probabilities should move towards it.

In similarly simple cases, evidence about how reliably moral beliefs were generated should likewise move prior probabilities towards truth ratios. It's a widespread assumption in contemporary metaethics that few if any processes reliably generate moral belief, keeping things similarly simple. Moral realist Michael Huemer describes Hume's "is-ought gap" (72) as preventing the many processes reliably generating nonmoral belief from evidentially supporting fundamental moral beliefs. Huemer therefore vigorously defends the reliability of moral intuition, seeing it as the sole source of good evidence for fundamental moral principles. Error theorists including Mackie correspondingly attack the reliability of moral intuition, aiming to undermine the entire evidential foundation for moral belief. Both sides seek to

demonstrate how reliably prior probabilities of truth were generated, with realists arguing that our case is like Livia's and anti-realists arguing that our case is like Ed's. The scarcity of evidence for moral propositions makes such evidence about reliability more important than in nonmoral domains where evidence is abundant. This paper aims to alleviate this scarcity by bringing new evidence about reliability across the is-ought gap, so that prior probabilities of moral propositions can be adjusted accordingly.

When the same process generates beliefs in multiple logically related propositions, the same truth ratio can't be applied to all of them. Suppose a 90% reliable process of intuition generates belief in two independent atomic propositions and their conjunction. Giving all three propositions credence 0.9 would be probabilistically incoherent.

There are various ways to avoid this incoherence. If all three propositions are separately intuited with reliability 0.9, an additional reliable process of logical reasoning might set things right, raising credences in the atomic propositions or reducing credence in their conjunction. The credences following logical reasoning will be the correct ones to have. Belief in logically stronger propositions may also be less reliably generated, with different parameter values reflecting greater risk of error. Visual perception is more reliable about the American flag having stripes than about its having exactly 13 stripes, as one might miscount the stripes. A parameter representing logical strength or complexity would express this difference in reliability. Different parameter values would allow logically related beliefs to be generated with different truth ratios, allowing the general rule of moving credences towards truth ratios to apply. Evidence from various instances of visual perception can vary in reliability and occasionally be inconsistent, for example when different lighting conditions make something seem to have different colors. Our understanding of how to aggregate complicated visual evidence can guide us in aggregating similarly complicated evidence from any single process generating moral belief, especially after discovering the parameters explaining its reliability.

If prior probabilities and new evidence come from different processes, their truth ratios must be aggregated to find the right credence. Here the simple rule of moving credences towards truth ratios no longer serves us, as there are two different truth ratios to move towards. Suppose I'm certain that an urn contains 99 red marbles and one green marble, and also that my color vision is only 90% reliable in the dimly lit room. If the marble I pick from the urn looks red, my credence that it's actually red should be above 0.99 rather than matching the truth ratio of color perception at 0.9. If it looks green, my credence that it's actually green should be below 0.9. Either way, it's a mistake to shift my credence to the 0.9 truth ratio of color perception, as prior evidence about the marbles suggests other credences.

Discovering the truth ratio of any process generating prior evidence allows its aggregation with our other evidence about how reliably beliefs are generated. For present purposes, prior certainty of 99 marbles being red is analogous to prior certainty of the 99% reliability of the process generating prior belief in moral proposition P. If a 90% reliable process also generates belief that P, one's credence in P shouldn't be 0.9 but above 0.99, as the 99% reliable process strongly confirms the 90% reliable process. If the 90% reliable process generates belief that not-P,

credence in not-P should be below 0.9, as the 99% reliable process contradicts the 90% reliable process.

However things go, discovering that a process is unexpectedly reliable should raise credences it generates. One's credence in P should rise if one discovers that the initial process is 99% reliable rather than 50% reliable, just as one's belief that the marble is red should rise after discovering that 99 rather than 50 of the 100 marbles are red. Discovering that a process is unexpectedly unreliable should likewise reduce credences it generates. Theories of evidence aggregation help determine the right credence when multiple reliable processes interact.³⁷

I'll conclude by describing a scenario in which evidence aggregation would be straightforward, and indicating the moral theory it would suggest.

3.2 Concluding Illustrative Hypothesis: from Introspection to Hedonism

If the only reliable process generating moral belief is phenomenal introspection, which reveals pleasure's goodness and displeasure's badness, the reliable route leads to ethical hedonism. Those unsympathetic to ethical hedonism may treat this hypothesis as a mere illustration of how nonmoral evidence might lead to moral conclusions; those more sympathetic are invited to explore whether it can be empirically confirmed.

Scientific investigation might discover phenomenal introspection to be a unified process reliably generating beliefs about intrinsic features of conscious experience. It's reliable when the introspector isn't impaired or distracted, and when the experiences being introspected aren't too complex. The same mathematical function would describe the truth ratio of phenomenal introspection in generating belief that sound-experience has volume and belief that pleasure is good. Both would be generated with high reliability.

Scientific investigation might also discover a second process generating all other fundamental moral beliefs, in which emotional responses to things lead to moral beliefs about them. Anthropological investigations might reveal that this process systematically generates jointly inconsistent moral beliefs in individuals with different emotional dispositions, leading to widespread disagreement on moral issues including the permissibility of genocide. As every inconsistent set includes a false belief, enough inconsistency will render this second process unreliable.

Truth-ratio induction from the nonmoral to the moral would then suggest high probability that pleasure is good. With inconsistency rendering the emotional process unreliable, truth-ratio induction would suggest low probability that any moral beliefs it generates are true. If no other processes generate fundamental moral beliefs, the reliable route leads to ethical hedonism. We cross Hume's gap between *is* and *ought*, and Bentham comes out to welcome us.

The reliable route may lead to a surprising destination. Discovering which beliefs are generated by unreliable processes and dismissing them may lead us away from

³⁷ Williamson (2009), Pettigrew (2019), Bright, Dang, and Heesen (Unpublished).

intuitive but erroneous moral theories. Then by extending demonstrably reliable processes into new domains, as Galileo did in pointing his telescope to the starry sky above, we can discover the moral law within.

References

- Alston, W. (1995). How to think about reliability. *Philosophical Topics*, 23(1), 1–29.
- Audi, R. (2013). *Moral Perception*. Princeton University Press.
- Becker, K. (2007). *Epistemology Modalized*. Routledge.
- Becker, K. (2008). Epistemic luck and the generality problem. *Philosophical Studies*, 139, 353–366.
- Beddor, B. (2014). Process reliabilism's troubles with defeat. *Philosophical Quarterly*, 65(259), 145–159.
- Beebe, J. (2004). The generality problem, statistical relevance and the Tri-Level Hypothesis. *Noûs*, 38(1), 177–195.
- Berker, S. (2009). The normative insignificance of neuroscience. *Philosophy and Public Affairs*, 37(4), 293–329.
- Bishop, M. (2010). Why the generality problem is everybody's problem. *Philosophical Studies*, 151(2), 285–298.
- BonJour, L. (1980). Externalist theories of empirical knowledge. *Midwest Studies in Philosophy*, 553–573. <https://doi.org/10.1111/j.1475-4975.1980.tb00396.x>
- Brentano, F. (1969). The origin of our knowledge of right and wrong. Trans. R. Chisholm and E.H. Schneewind. Routledge and Kegan Paul
- Bright, L., Haixin, D., & Remco, H. (Unpublished). Judgment aggregation in science.
- Brink, D. (1989). *Moral realism and the foundations of ethics*. Cambridge University Press.
- Broi, A. (2022). Une défense de l'hédonisme axiologique. *Dialogue*, 61(2), 325–346.
- Brown, C. (2014). Minding the is-ought gap. *Journal of Philosophical Logic*, 43, 53–69.
- Brown, C. (2015). Two versions of Hume's law. *Journal of Ethics and Social Philosophy*, 9, 1–8.
- Bruni, T., Mario, M., & Regina, R. (2014). The science of morality and its normative implications. *Neuroethics*, 7, 159–172.
- Chang, H., & Michel, P. (2018). Affective boundaries of scope insensitivity. *Journal of Consumer Research*, 45(2), 403–428.
- Chisholm, R. (1989). *Theory of knowledge*. Prentice-Hall.
- Christensen, D. (2007). Epistemology of disagreement: The good news. *Philosophical Review*, 116(2), 187–217.
- Cohen, S. (1984). Justification and truth. *Philosophical Studies*, 46, 279–295.
- Comesaña, J. (2006). A well-founded solution to the generality problem. *Philosophical Studies*, 129(1), 27–47.
- Conee, E. (2013). The specificity of the generality problem. *Philosophical Studies*, 163, 751–762.
- Conee, E., & Feldman, R. (1998). The generality problem for reliabilism. *Philosophical Studies*, 89(1), 1–29.
- Dawson, M. (2013). *Mind, body, world: foundations of cognitive science*. Athabasca University Press.
- De Lazari-Radek, K., & Peter, S. (2012). The objectivity of ethics and the unity of practical reason. *Ethics*, 123(1), 9–31.
- Decker, J., & Daniel, G. (2013). On the insignificance of moral disagreement for moral knowledge. In E. Shafer-Landau (Ed.), *Oxford studies in metaethics* 8. Oxford University Press.
- Dehaene, S. (1997). *The number sense*. Oxford University Press.
- Demaree-Cotton, J. (2016). Do framing effects make moral intuitions unreliable? *Philosophical Psychology*, 29(1), 1–22.
- Desvousges, W. (2010). *Measuring nonuse damages using contingent valuation*. RTI Press.
- Doris, J., & Alexandra P. (2008b). How to locate a moral disagreement. In: Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 2, pp. 345–354). MIT Press
- Doris, J., & Alexandra P. (2008a). How to argue about disagreement. In Sinnott-Armstrong Ed., *Moral Psychology* (Vol. 2, pp. 303–332). MIT Press
- Dutant, J., & Eric, O. (2013). Is there a statistical solution to the generality problem? *Erkenntnis*, 78(6), 1347–1365.
- Elga, A. (2007). Reflection and disagreement. *Noûs*, 41(3), 478–502.

- Enoch, D. (2009). How is moral disagreement a problem for realism? *Journal of Ethics*, 13, 15–50.
- Enoch, D. (2011). *Taking morality seriously*. Oxford University Press.
- Feest, U. (2014). Phenomenal experiences, first-person methods, and the artificiality of experimental data. *Philosophy of Science*, 81(5), 927–939.
- Fetherstonhaugh, D., Paul, S., Stephen, J., & James, F. (1997). Insensitivity to the value of human life: A study of psychophysical numbing. *Journal of Risk and Uncertainty*, 14, 283–300.
- Goldman, A. (2000). What is justified belief? In Sosa & Kim (Eds.), *Epistemology*, Blackwell.
- Goldman, A. (1986). *Epistemology and cognition*. Harvard University Press.
- Goodman, N. (1965). *Fact, fiction, and forecast*. Bobbs-Merrill.
- Gray, K., & Daniel, W. (2010). Blaming God for our pain: Human suffering and the divine mind. *Personality and Social Psychology Review*, 14(1), 9–10.
- Greene, J. (2008). The secret joke of Kant's soul. In Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 3, pp. 35–79). MIT Press
- Greene, J. (2014). Beyond point-and-shoot morality: Why cognitive (neuro)science matters for ethics. *Ethics*, 124(4), 695–726.
- Greenwald, A., & Mahzarin, B. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4–27.
- Guevara, D. (2008). Rebutting formally valid counterexamples to the Humean “is-ought” dictum. *Synthese*, 164(4), 45–60.
- Hájek, A. (2011). *Interpretations of probability*. Stanford Encyclopedia of Philosophy.
- Heller, M. (1995). The simple solution to the problem of generality. *Noûs*, 29(4), 501–515.
- Hoffman, C., & Nancy, H. (1990). Gender stereotypes: Perception or rationalization? *Journal of Personality and Social Psychology*, 58(2), 197–208.
- Horn, J. (2017). Moral realism, fundamental moral disagreement, and moral reliability. *Journal of Value Inquiry*, 51(3), 363–381.
- Huemer, M. (2005). *Ethical intuitionism*. Palgrave Macmillan.
- Hume, D. 1740/2000. *A Treatise of Human Nature*. Oxford University Press.
- Jackson, E. (2018). *Belief and credence: why the attitude-type matters*. Philosophical Studies.
- Jackson, F. (1998). *From metaphysics to ethics*. Oxford University Press.
- Jones, A. (2011). *Genocide: A comprehensive introduction*. Routledge.
- Joyce, R. (2001). *The Myth of morality*. Cambridge University Press.
- Joyce, R. (2006). *The Evolution of morality*. MIT Press.
- Kagan, S. (1998). *Normative ethics*. Westview Press.
- Kahane, G. (2016). Is, ought, and the brain. In Moral Brains, L. (Ed.), Oxford University Press.
- Kampa, S. (2018). A new statistical solution to the generality problem. *Episteme*, 15(2), 228–244.
- Karmo, T. (1988). Some valid (but no sound) arguments trivially span the is-ought gap. *Mind*, 97, 252–257.
- Kelly, T. (2005). The epistemic significance of disagreement. In Hawthorne & Szabo (Eds.), *Oxford Studies in Epistemology* (Vol. 1, pp. 167–196). Oxford University Press.
- Kelly, D. (2011). *Yuck!*. MIT Press.
- Leiter, B. (2008). Against convergent moral realism: the respective roles of philosophical argument and empirical evidence. In Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 2, pp. 333–338). MIT Press.
- Lepock, C. (2009). How to make the generality problem work for you. *Acta Analytica*, 24, 275–286.
- Lewis, D. (1986). A subjectivist's guide to objective chance. *Philosophical Papers: Volume II* (pp. 263–293). Oxford University Press.
- Locke, D. (2014). The decision theoretic Lockean thesis. *Inquiry*, 57, 28–54.
- Lyons, J. (2019). Algorithm and parameters: Solving the generality problem for reliabilism. *Philosophical Review*, 128(4), 463–509.
- Mackie, J. (1977). *Ethics: Inventing right and wrong*. Penguin.
- Marr, D. (1982). *Vision*. Freeman.
- May, J. (2014). Does disgust influence moral judgment? *Australasian Journal of Philosophy*, 92(1), 125–141.
- Miller, W. (1978). *Fishbait: The memoirs of the Congressional doorkeeper*. Prentice-Hall.
- Moore, G. (1903). *Principia ethica*. Cambridge University Press.
- Nietzsche, F. 1889/1968. *The Portable Nietzsche*. In Kaufman (Ed.) and trans. Penguin.
- O'Neill, E. (2015). Which causes of moral beliefs matter? *Philosophy of Science*, 82, 1070–1080.

- Pettigrew, R. (2019). On the accuracy of group credences. In Hawthorne and Gendler (Eds.), *Oxford Studies in Epistemology 6*. Oxford University Press.
- Pigden, C. (1989). Logic and the autonomy of ethics. *Australasian Journal of Philosophy*, 67, 127–151.
- Prior, A. N. (1960). The autonomy of ethics. *Australasian Journal of Philosophy*, 38(3), 199–206. <https://doi.org/10.1080/00048406085200221>
- Quine, W. (1969). *Ontological relativity and other essays*. Columbia University Press.
- Railton, P. (1989). Naturalism and prescriptivity. *Social Philosophy and Policy*, 7(1), 151–174.
- Rawlette, S. (2014). *The feeling of value*. CreateSpace Independent Publishing Platform.
- Resnick, I., Nora, N., & Thomas, S. (2017). Dealing with big numbers: Representation and understanding of magnitudes outside of human experience. *Cognitive Science*, 41, 1020–1041.
- Rini, R. (2016). Debunking debunking: A regress challenge for psychological threats to moral judgment. *Philosophical Studies*, 173(3), 675–697.
- Rømer, O. (1677). On the motion of light. *Philosophical Transactions of the Royal Society of London*, 12, 893–894.
- Russell, G. (2010). *In defence of Hume's law* (Pigden, pp. 151–161). Palgrave Macmillan.
- Schurz, G. (1997). *The is-ought problem: An investigation in philosophical logic*. Kluwer.
- Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review*, 117(2), 245–274.
- Shaver, R. (2004). The appeal of utilitarianism. *Utilitas*, 16(3), 235–250.
- Siegler, R., & John, O. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, 14, 237–243.
- Singer, D. (2015). Mind the is-ought gap. *Journal of Philosophy*, 112, 193–210.
- Singer, P. (2005). Ethics and intuitions. *Journal of Ethics*, 9, 331–352.
- Sinnott-Armstrong, W. (2008b). How to apply generalities: reply to Tolhurst and Shafer-Landau. In Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 2, pp. 97–105). MIT Press.
- Sinnott-Armstrong, W. (2008a). Framing moral intuitions. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 2, pp. 48–75). MIT Press.
- Slovic, P. (2010). The more who die, the less we care. In Michel-Kerjan & Slovic (Eds.), *The Irrational Economist* (pp. 30–40). Public Affairs.
- Sosa, E. (1991). *Knowledge in perspective*. Cambridge University Press.
- Stevenson, C. L. (1937). II—The emotive meaning of ethical terms. *Mind XLVI*(181), 14–31. <https://doi.org/10.1093/mind/XLVI.181.14>
- Stewart, T. (2005). The competing practices argument and self-defeat. *Episteme*, 2, 13–24.
- Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies*, 127(1), 109–166.
- Street, S. (2010). What is constructivism in ethics and metaethics? *Philosophy Compass*, 5(5), 363–384.
- Tersman, F. (2008). The reliability of moral intuitions: A challenge from neuroscience. *Australasian Journal of Philosophy*, 86(3), 389–405.
- Tolly, J. (2017). A defense of parrying responses to the generality problem. *Philosophical Studies*, 174, 1935–1957.
- Vavova, K. (2014). Moral disagreement and moral skepticism. *Philosophical Perspectives*, 28, 302–333.
- Van Wees, H. 2010. Genocide in the ancient world. In Bloxham & Moses (Eds.), *Oxford Handbook of Genocide* (pp. 239–258). Oxford University Press.
- Williamson, J. (2009). Aggregating judgments by merging evidence. *Journal of Logic and Computation*, 19(3), 461–473.
- Wu, W. 2018. The neuroscience of consciousness. In Zalta (Ed.), *Stanford Encyclopedia of Philosophy*.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.