CrossMark

Episteme (2021), **18**, 367–383
doi:10.1017/epi.2021.32

episteme

ARTICLE

Normative Principles are Synthetic A Priori

Paul Boghossian*

NYU, New York, USA
*Corresponding author. Email: pb3@nyu.edu

(Received 4 July 2021; accepted 5 July 2021; first published online 24 August 2021)

Abstract

I argue for the claim that there are instances of a priori justified belief – in particular, justified belief in moral principles – that are not analytic, i.e., that cannot be explained solely by the understanding we have of their propositions. §1–2 provides the background necessary for understanding this claim: in particular, it distinguishes between two ways a proposition can be analytic, Basis and Constitutive, and provides the general form of a moral principle. §§3–5 consider whether Hume's Law, properly interpreted, can be established by Moore's Open Question Argument, and concludes that it cannot: while Moore's argument – appropriately modified – is effective against the idea that moral judgments are either (i) reductively analyzable or (ii) Constitutive-analytic, a different argument is needed to show that they are not (iii) Basis-analytic. Such an argument is supplied in §6. §§7–8 conclude by considering how these considerations bear on recent discussions of "alternative normative concepts", on the epistemology of intuitions, and on the differences between disagreement in moral domains and in other a priori domains such as logic and mathematics.

Keywords: A priori justification; synthetic a priori; G.E. Moore; Open Question Argument; Hume's Law, alternative normative concepts; intuition; analyticity; moral principles; thick concepts

1. Analyticity and the synthetic a priori

There are many propositions that we are in a position to justifiably believe merely in virtue of understanding them: they are, in the terminology of Boghossian (1996), "epistemically analytic." A good example is given by (1).

(1) If Mary ate the apple and the pear, then Mary ate the apple.

The notion of epistemic analyticity should be sharply distinguished from the notion of "metaphysical analyticity," according to which certain propositions are true purely in virtue of their contents alone and without any contribution from the facts. This latter notion is both of dubious coherence and also of dubious explanatory value. What's important to see is that the notion of epistemic analyticity – of a proposition's being justifiably believable merely in virtue of its being understood – is distinct from the notion of 'truth in virtue of meaning' and need not involve it.

How does a priori justification in virtue of understanding work? In the case of (1), it's tempting to think that what the understanding does is put you in a position to *reason* as follows:

© The Author(s), 2021. Published by Cambridge University Press

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

a. If Mary ate *both* the apple and the pear, then one of the things she would have had to do is eat the apple.
So,
b. If Mary ate both the apple and the pear, then she ate the apple.

But this, of course, can't be right, since the premise of this reasoning leading to (1) is basically just (1) itself. So, we had better say that the understanding justifies (1) directly, without inference.

This is even clearer when we look not at the justification of belief, but at the justification of *reasoning*. A good example of a piece of reasoning that we are justified in performing just in virtue of understanding it is the following instance of Modus Ponens (MP) reasoning:

(2) If Mary ate the apple, then she didn't eat the pear.
Mary ate the apple.
So, Mary didn't eat the pear.

MP, as a rule of logic, may be represented as follows:

(MP) For any propositions P, Q, given premises P and If P, Q, you may deduce Q.[1]

Any attempt to articulate the justification provided by the understanding for this piece of reasoning in inferential terms would run into the obvious point that the inference would involve the very pattern being justified. So, we had better say that the understanding justifies the pattern in (2) directly.

How does the understanding provide direct, non-inferential justification for a belief, or a pattern of reasoning? Focus first on the case of reasoning according to MP.

The first possibility is that reasoning according to MP is *constitutive* of understanding (2), in particular, of understanding 'if.' The second possibility is that, although reasoning according to MP is not constitutive of understanding 'if,' your understanding of 'if' provides you with a strong *epistemic basis* for seeing that MP reasoning is valid.

On the first option, we would need to explain why an inference rule's being constitutive of understanding that inference is a prima facie justification for inferring with it. This is not the place to go into this question in detail (see Boghossian 2003 and Chapters 13 and 15 of Boghossian and Williamson 2020), but the basic idea is that "That's just what I mean by w" is a perfectly good prima facie justification for reasoning with one of w's meaning-constituting inferences. Of course, it's possible that the rules in question will lead to inconsistencies when combined with the rule constitutive of other concepts, and this might lead to our prima facie justification being defeated. But that doesn't negate the initial claim.[2]

However, even if we could make this first type of constitutive, understanding-based explanation work, there would be need for the non-constitutive model as well, since there are beliefs/inference patterns that are clearly epistemically analytic but that are

[1]This formulation of the rule, with the word 'deduce' as opposed to 'infer' or 'conclude,' is designed to get around Harman's (1986) objection that nothing of the form MP could be thought of as a rule of reasoning. The formulation makes it clear that the rule is not meant as a rule of belief management, but rather as a rule of deduction, which is itself a component part of the overall set of norms that govern belief management.

[2]Some philosophers, like Williamson (2007), are skeptical that any beliefs or inferences are constitutive of understanding any word or concept. I believe such skepticism is misguided, as I explain in Boghossian and Williamson (2020).

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

not constitutive of our understanding of any ingredient concept. For example, while it is probably constitutive of your understanding of simple quadrilateral that you assent to

(Quadrilateral) A simple quadrilateral is a plane four-sided figure,

it is not constitutive of your understanding of 'simple quadrilateral' that you assent to 'all simple quadrilaterals have four interior angles.' However, once you have understood what a simple quadrilateral is, you are in a position to see that all simple quadrilaterals have four interior angles on the basis of your understanding alone. Once more, I refer the reader to the previous citations for details.

To sum up, then. There are two types of understanding-based accounts of a priori justification, Constitutive and Basis. On the first, S's understanding of p justifies S's assent to p in virtue of the fact that S's understanding of p is *constituted* in part by S's disposition to assent to p.

(Constitutive) S's assent to p is justified by S's understanding of p in virtue of the fact that S's understanding of p is partly constituted by S's disposition to assent to p.

On the Basis type of account, S's understanding of p justifies S's assent to p in virtue of the fact that S's understanding of p serves as a good *epistemic basis* for S's assent to p.

(Basis) S's understanding of p justifies S's assent to p in virtue of the fact that S's understanding of p serves as a good *epistemic basis* for S's assent to p.

As I say, there is much more to be said about both of these types of explanation, but I won't be saying it here since my focus in this essay is the claim that there are instances of a priori justified belief that *cannot* be explained solely by the understanding that we have of their propositions. To put it in Kantian terminology (although needless to say, my usage will not completely coincide with his), there are *synthetic a priori* propositions.

I will illustrate my argument by focusing on the case of *moral* principles, but the argument is a general one that applies to all *normative* principles, including, importantly, the principles of epistemic rationality. My discussion will presuppose that there are moral truths that are not constituted by our attitudes or practices, but nothing much beyond that by way of a realist meta-ethics.

Of course, I am hardly the first philosopher to argue that normative claims are synthetic a priori; the claim has a venerable history, counting among its proponents the likes of G.E. Moore and Derek Parfit. My hope is that we can use advances in our understanding of the notions of analyticity and apriority to explain more clearly than heretofore why the thesis is true.

2. The content of normative statements

Let's assume, plausibly enough, that the following is a true moral principle:

(NoTorture) Necessarily, it is wrong to inflict extreme pain on a person just for one's own amusement.

NoTorture illustrates the general logical form of a moral principle:

(DtoM) Necessarily, If x is D, then x is M

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

where 'D' is a canonical descriptive property – in this case, 'inflicting extreme pain on a person just for one's own amusement' – and 'M' is a canonical moral property – in this case, 'wrong.' When an instance of (DtoM) is true, there will be a corresponding inference that is valid:

(DtoM-I) (1) x is D
So, (2) x is M

The inferential form of NoTorture would therefore be:

(NoTorture-I) (3) x is an inflicting of extreme pain just for one's own amusement.
So, (4) x is wrong.

Now, to say that a principle like NoTorture is synthetic a priori is to say that, while it can be justifiably believed a priori, that justification does not derive *merely* from our understanding of what it says, in either of the two ways just now distinguished. Being disposed to assent to it is not constitutive of understanding it; nor does understanding it by itself serve as a good epistemic basis for believing it.

Since the principle would appear to be amongst the most obvious of moral claims, one that is as close to commanding universal assent as one could hope to get, how could one possibly hope to establish that it is nonetheless not epistemically analytic? Here we face one of Quine's (1951) most powerful challenges to the notion of an analytic claim: how to distinguish an analytic claim from one that is so obvious that well-nigh anyone who understood it *would* assent to it? I will try to show that Quine's challenge can be answered in the case of normative claims like NoTorture.

3. Hume's Law

It is arguable that the first philosopher to insist on the synthetic character of moral claims was Hume. Hume famously remarked:

> In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs; when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*. This change is imperceptible; but is, however, of the last consequence. For as this *ought*, or *ought not*, expresses some new relation or affirmation,'tis necessary that it shou'd be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it … [I] am persuaded, that this small attention [to this point] wou'd subvert all the vulgar systems of morality, and let us see, that the distinction of vice and virtue is not founded merely on the relations of objects, nor is perceiv'd by reason. (Hume 1987: T 3.1.1.27, pp. 469–70)

This is, of course, the classical statement of 'Hume's Law,' which says that an 'ought' can never be validly derived from an 'is.' But what sense of 'validly derive' is at issue?

Hume exegesis is not my aim here. Rather, my point will be that Hume's Law is only both interesting and potentially defensible if it is understood as denying that there is an

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

epistemically analytic inference from an is to an ought (or to any other normative claim).

One could read Hume, as many have done, as denying merely that there is a *logically* valid inference from an is to an ought. This is correct, of course, but boring: even the most passionate defenders of NoTorture would not wish to claim that the inference from (3) to (4) is valid merely in virtue of its logical form.

One might also read Hume as denying that there is a *strict entailment* from (3) to (4). A strict entailment holds in all metaphysically possible worlds, whether or not the entailment is logically valid or known a priori. Thus, we are inclined to think that there is a strict entailment from (5) to (6)

(Water) (5) x is water
So, (6) x is H_2O

even though the entailment is neither logically valid nor could be known a priori.

Read this way, though, Hume's Law would be open to severe objections.

First, the existence of such strict entailments follows from the widely accepted claim that particular moral facts supervene on the descriptive facts: necessarily, no two possible acts can be alike in all relevant descriptive respects (which might include its relational properties) yet differ in moral respects. Given such supervenience, there will always be a necessary claim of the form:

(Supervenience Conditional) Necessarily: If x is D*, then x is M*.

where 'D*' is a description of the act's relevant descriptive properties and 'M*' is the specification of its moral properties.

Second, even if, as some philosophers are inclined to do, an adherent of Hume's Law were to reject the supervenience of the moral on the physical, it would be hard to see how such a philosopher could be so confident that there can be no strict entailments from is to ought, since the entailments in question would not be claimed to be a priori obvious.

Arguably, therefore, the only available reading of Hume's Law on which it is both interesting and potentially defensible is the reading on which what's being denied is that there is an *epistemically analytic entailment* from an act's descriptive properties to its moral properties. But what would be the argument for this potentially defensible claim?

4. Moore on defining normative concepts

The beginnings of the argument that's needed were given by Moore in his famous Open Question Argument, although this is not necessarily evident from the discussion as Moore develops it.

Moore concentrated on the term 'good' and invited us to consider whether there could be a reductive definitional analysis of it. Some Utilitarians had proposed that to say that an act was 'good' *meant* that it would maximize aggregate happiness. Moore wanted to show that neither this, nor any other reductive definition of 'good' could be correct.

A propos of the Utilitarian definition, Moore argued as follows: there could be someone who understood 'good' perfectly well but who coherently wondered whether whatever maximized aggregate happiness was really good. Indeed, isn't that precisely what anti-Utilitarian ethicists did; and, surely, they grasp the concept expressed by 'good' if anyone does. However, if 'good' really did mean 'that which maximizes aggregate happiness' then someone questioning whether what maximized aggregate happiness

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

was itself good would be questioning whether what maximized aggregate happiness maximized aggregate happiness, which would be nonsensical. Therefore, Moore concluded, the Utilitarian definition of 'good' could not be correct.

Moore believed that this point generalized. For any reductive definition of 'good', of the form

(Good) x is good = (def) x is F

there can be someone who understands 'good' perfectly well, but who nonetheless competently doubts whether F is good. Therefore, no such reductive definition of 'good' could be correct. Mutatis mutandis for the other basic moral concepts.

Moore's argument can be summarized as follows:

A. For any reductive definition of 'good,' there can be someone who has *mastered* the meaning of 'good' but who *doubts* whether whatever satisfies the definiendum is good.

Therefore,

B. No such definition is correct.[3]

As so presented, the argument can seem unsound: there can seem to be a problem both with substantiating its first premise, and with the claim that that premise entails the conclusion.

First, assuming that we grant the first premise, how does the conclusion follow? Couldn't the master of a concept rationally doubt a *correct* definition of it? Certainly, the history of philosophy is littered with examples of experts accepting a definition as correct when it wasn't. For example, epistemologists are experts on the concept expressed by 'know'; yet, for many years they accepted as correct the analysis of knowledge as justified true belief which we now believe to be incorrect. If it's possible for an expert to accept an incorrect definition as correct, it's presumably also possible for them to reject a correct definition as incorrect. In general, experts in particular domains are not guaranteed to be *infallible* about the definitions of the concepts of that domain.

Second, how does Moore know that, for *any* reductive definition of 'good,' there *can* be someone who understands all the ingredient concepts perfectly well and yet can doubt the equivalence? There are certainly concepts for which Moore's claim would *not* be true. For example, anyone who rejected or doubted the definition Quadrilateral we gave above would disqualify themselves from understanding that concept. How does Moore know that 'good' isn't similar to 'quadrilateral' in this respect – that there isn't a reductive definition of 'good' which is such that, once we are presented with it and have sufficient time to reflect on it, we would all agree on the equivalence it postulates, and so agree that it cannot be competently doubted?

Stated thus abstractly, Moore's official argument looks problematic. However, there is a simple observation at its heart which goes a long way towards establishing its

[3]I have strengthened Moore's argument by using the notion of the mastery of a concept in premise 1 rather than the notion of the mere *possession* of a concept. As illustrated by Burge-style thought experiments, there is a weak notion of concept possession on which one can possess a concept while misunderstanding it or having only a partial understanding of it. However, if we employed such a notion in the first premise, Moore's argument would obviously fail. Since, on such a weak notion, you could possess a concept while misunderstanding it, or having only a partial understanding of it, there would be no entailment from A to B: you could possess a concept while being unable to recognize a perfectly correct definition of it.

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

conclusion. It may be put like this: To say of an act that it is good analytically entails *endorsing* it, *approving* of it, or taking some other kind of positive attitude towards it. That, after all, is part and parcel of saying that 'good' is a positive normative or evaluative term. Anyone saying of x that it's good would have to be expressing some sort of positive attitude towards it; that much is clear. Mutatis mutandis for anyone judging that something is bad.

What, though, is it to take a positive evaluative attitude towards an act, to approve of it in this way? On one influential view of the matter, anyone, S, who judges that x is good will be herself *motivated* to do x and will be disposed to *praise* anyone who has done x and to feel positive *emotions* towards them for having done it. By contrast, if S judges that x is bad, she will be disposed to avoid doing x, and will be critical of someone who does perform x and may feel negative emotions towards them. Importantly, these dispositions will be there as a mere result of S's having made the relevant evaluative judgment and independently of S's other desires. (These are obviously mere sketches of accounts and they can be filled in, or altered, in various ways – the details won't matter for present purposes).

Such a view is often referred to as an *internalist* view of moral judgments: 'internalist' because it posits internal connections between the making of a moral judgment and facts about the motivation and moral emotions of the judger. According to such a view, then, moral concepts, and the judgments into which they are embedded, have a constitutive *normative role*. Their role is given by their links to dispositions to behave in certain ways and to feel certain emotions. When someone makes a moral judgment, it is necessarily true that they are disposed to behave in certain ways and to feel certain things. Of course, the link here is merely dispositional, so it's consistent with the truth of judgment internalism that those dispositions may, on occasion, fail to manifest themselves.

If we accept such an internalist view of moral judgments, then we have the means by which to substantiate Moore's conclusion, for there seems to be no similar constitutive tie between making a *descriptive judgment* and any positive attitude or emotion. Saying of an act that it maximizes aggregate happiness does *not* constitutively imply any such dispositions to action or emotion: looking just at the nature of such a speech act, and at the meaning of what is said, it leaves the question of the evaluative attitude one bears to it an *open question*. It might not be nice of someone to *disapprove* of x precisely *because* it would maximize aggregate happiness, but it would not be metaphysically impossible for them to do so. If that is correct, then 'good' could not mean the same as 'what maximizes aggregate happiness,' since these two expressions have distinct constitutive properties.

Now, since the attribution of any descriptive property would leave the question of the judger's evaluative attitude open in exactly the same way, no naturalistic definition of 'good' could be correct, just as Moore claimed. Hence, Moore's claim that no reductive analysis of 'good' could be correct seems to hold up.

The argument should not be misunderstood. For all that it shows, it might well be true, indeed it might well be necessarily true, that an act is good if and only if it maximizes aggregate happiness. But if such a claim is true it would not be knowable to be true merely in virtue of being understood. It would not be an epistemically analytic truth. As Moore (1903: 6) himself said, when a claim equating the good with a descriptive condition is true, it is at best a *synthetic a priori* truth and not an *analytic* one.

This argument can be summarized as follows:

1. It is necessary that, if S judges that x is good, S has certain positive dispositions towards x.

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

2. It is not necessary that if S judges that x is D, where D is any descriptive property, S has any positive dispositions towards x.
   Therefore,
3. For any D, 'good' cannot mean the same as D.

5. Normative principles

Now, it might be thought that Moore's result, while interesting and important, is too weak to show us anything about the analytic status of normative principles, which are conditionals of the form

(DtoM) If x is D, then x is M

not biconditionals, of the form

(Def) x is M iff (def) x is D

which is what Moore's argument concerns. Formally speaking, it is compatible with every instance of (Def) being false that there are true analytic conditionals of the form (DtoM), as well as true analytic conditionals of the form

(MtoD) If x is M, then x is D.

Indeed, we *know* that there are true analytic conditionals of the latter form; for, if an act is virtuous, say, then we know, by understanding alone, that it was performed by a creature with a mind, since inanimate or unconscious creatures cannot perform virtuous acts.

If that's so, might there not also be true analytic conditionals of the (DtoM) variety, too, compatibly with Moore's result? And if that's the case, the understanding alone may suffice for justifiably believing that certain normative principles are true, despite its being the case that no moral term has a reductive naturalistic analysis.

If Moore's argument works, we know at least this much: No instance of (DtoM) could be *constitutively* analytic: it can't be a condition on someone's mastering the concepts M and D that they be prepared to assent to any particular instance of

If x is D, then x is M

since Moore has already established that asserting x is D is compatible with any evaluative attitude that one may wish to combine it with.

What, though, if the tie between D and M is claimed to be not constitutively analytic, but basis-analytic, in the way outlined at the start of this paper? On a basis view, the justifiability of 'if x is D, then x is M' is extractable from the meanings of 'D' and 'M' not because assenting to the conditional is a condition on possessing M, but because the understanding of the conditional provides a sound epistemic basis for deriving the conditional, in much the way in which the understanding of 'simple quadrilateral' provides a sound epistemic basis for concluding that all simple quadrilaterals have four interior angles.

Well, wouldn't Moore's argument still apply? On a basis view, shouldn't a master of M be able, by reflection on the relevant concepts alone, to conclude that if x is D, then x is M. However, hasn't Moore established that one could be a master of D and M and yet not conclude that if x is D, x is M?

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

The problem is that the reason *why* a thinker may be unable to extract the DtoM conditional from their understanding of D and M may be due not to the conditional's not *being* basis-analytic, but because the thinker in question has a distorting background theory that prevents her from seeing the justification for the conditional that is in fact contained in her understanding. In other words, and obviously so given the notions involved, you can't refute the claim that some DtoM conditional is basis-analytic just by showing that it's not constitutively-analytic. I will illustrate this point with a neutral example.

Let's suppose that it is basis-analytic of the concept 'knowledge' that if S knows p then S believes p. Now, consider a theorist, T, who has come to be convinced through various clever, if misguided philosophical arguments, that in fact no one has beliefs, that the folk-psychological concept of belief has an empty extension. (Needless to say, T is not wholly fictional: a number of actual philosophers have held such a view.) T does not also wish to assert that no one knows anything. He may wish to avoid that conclusion at all costs, worried that it might be self-undermining. He certainly wishes to claim that we know that there are no beliefs. Such a thinker would be highly motivated to assert that knowing does not entail believing, despite its being true that it is epistemically analytic that whosoever knows believes. A distorting strongly held background theory prevents him from appreciating the analytic truth.

So, however intuitive it may be that, for any D and M, someone may competently doubt that

if x is D then x is M,

that can at most establish that the conditional is not constitutively-analytic, not that it is not basis-analytic. If we are to establish that normative principles are not epistemically analytic in either of the available senses, we need something more than Moore's argument.

In what follows, I will supply an argument for the claim that substantive moral principles are not basis-analytic. If the argument is cogent, then we will have shown that substantive moral principles are not epistemically analytic, since they are neither constitutively-analytic nor basis-analytic. On the assumption that substantive moral principles are not a posteriori, we will have shown that they are synthetic a priori.

6. Moral concept correctness

Consider again the principle

(NoTorture) Necessarily, it is wrong to inflict extreme pain on a person just for one's own amusement.

and imagine a subject who is as free of distorting background philosophical views and prejudices as possible. She considers this principle and, after adequate reflection, including the consideration of potential counterexamples, she concludes that it is true, indeed, perhaps obviously so. Should we say that she arrived at her conclusion *solely* on the basis of her understanding of NoTorture? I will develop an argument for thinking otherwise.

As we saw previously, the minimal core of a normative term is its normative role, which we fleshed out in terms of its constitutive connections to positive (or negative) attitudes when it features in an actual normative judgment. That much is essential to anything properly regarded as a normative term.

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

If a proposition like NoTorture is to be epistemically analytic in general, and basis-analytic in particular, its central normative term, 'wrong,' must consist not merely of this normative role, which in and of itself leaves it wide open which particular act-types are to be disapproved of; it must also contain a *substantive* conception of which acts are wrong. In particular, it must either directly, or indirectly, include the information that inflicting pain on a person just for one's own amusement is wrong, so that someone who understood NoTorture would be able to see that it is true simply by reflecting on its ingredient concepts and the way they are put together, assuming she were not precluded from doing so by some distorting background theory. Let's call any moral concept that incorporates a view about which acts or traits fall under its extension a *substantive* moral concept. By way of contrast, I will designate any moral concept that consists merely in its normative role a *minimal* moral concept.

Now, it is obvious, of course, that we have *some* substantive moral concepts, namely, the so-called "thick" moral concepts – concepts such as 'selfish,' 'brave,' 'cruel,' and so on. Each of these thick moral concepts singles out a particular act- or trait-type for moral approbation or disapprobation. The concept of a selfish person, for example, is the concept of a person who puts his own interests ahead of those of others in a morally objectionable way. There continues to be a great deal of controversy about exactly how thick moral concepts incorporate both descriptive and normative elements; but that is clearly what they do, and any putative master of such a concept would have to understand that.

By contrast, the 'thin' moral concepts, the ones that we've been working with so far – 'good,' 'bad,' 'right,' 'wrong,' 'ought,' and 'ought not,' – do not, in any overt way, incorporate views about particular act- or trait-types. As far as their surface appearance is concerned, they seem to be minimal moral concepts. However, the proponent of the epistemic analyticity of NoTorture would have to insist that that surface appearance is misleading, and that one can after all recover from them that inflicting extreme pain on another just for one's own amusement is wrong.

To test this, let's suppose that our concept of 'wrong' is substantive in this way and so is able to provide a sound epistemic basis for our concluding that NoTorture is true. About such a substantive concept, embodying a substantive view of which act-types count as wrong, it seems that we can *always* competently ask:

> (Moral Concept Correctness Open Question, MCCOQ) We can all agree that act-type A is ruled wrong by *your* substantive concept of 'wrong.' But is A *genuinely* wrong? Is your concept of 'wrong' the *correct* concept here, the one that we *ought* to be using, if we are to capture those act-types that are genuinely wrong?

Of course, the questions here will be felt more vividly if they were directed against a view that seemed obviously false, rather than, as in the present case, directed against one that seems obviously true! But that shouldn't matter to the point at issue, which is that no matter which view is analytically embedded in the concept, even one that seems obviously true, it still makes sense to ask the concept correctness open question (CCOQ).

To get a feel for this, consider the contrast with the concept 'simple quadrilateral'. If someone said: "According to my concept of 'simple quadrilateral,' a quadrilateral always has four sides," it wouldn't make any sense to ask: "Yes, but is that the *correct* concept of 'quadrilateral'? Is that the concept that delivers the genuinely correct verdicts about how many sides quadrilaterals have?" So, the concept correctness question that I have highlighted for substantive moral concepts does not always sensibly arise for other concepts. In the case of a putatively substantive moral concept, however, it does always seem to

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

make sense to ask whether that concept is correct, whether it correctly reflects whatever moral facts there are about the rightness and wrongness of acts.

Notice, too, that a CCOQ could *not* sensibly be raised for a minimal moral concept, one that consists *wholly* in its normative role. If 'wrong' were said to consist only in enabling a certain kind of criticism, but without analytically committing its user to any particular set of things as worthy of being criticized, it would make no sense to ask whether this is the correct concept of 'wrong.' You are presumably entitled to have available to you a concept whose job it is to enable that type of criticism, even if it turns out that nothing out there *is* worthy of that type of criticism (as an Error theorist of Mackie's stripe may hold).[4]

It makes sense to ask of a concept of 'wrong' that incorporates a substantive view of which act-types are worthy of a certain kind of criticism, whether it is correct; but it does not make sense to ask of a concept whose sole job it is to be able to raise the question whether a given act-type is worthy of a certain type of criticism, whether it is correct. Since such a minimal concept is making no claims, there is nothing for it to get wrong or right, even if we end up concluding that its extension is empty.[5]

What is the significance, for our purposes, of the alleged fact that a CCOQ can always be raised for a *substantive* moral concept? Well, if showing that some substantive moral proposition is conceptually extractable from a given moral concept always leaves the question of the correctness of that concept open, then the correctness of the substantive moral proposition is also left open. As a result, we are not able to claim that the truth of the substantive moral proposition has been settled by conceptual considerations alone. In other words, we are not able to claim that the proposition in question is epistemically analytic. And that is the conclusion I was after.

It would be natural to be suspicious of this argument. Can it really be that easy to establish such a strong and interesting conclusion, one that has potentially large and potentially problematic ramifications in the epistemology of the a priori? And is it really true that one can't raise similar 'concept correctness' open questions about a host of other *non-normative* concepts?

Let me start with the second question. We have already seen that the CCOQ does not make sense as applied to the concept of a simple quadrilateral. Obviously, there is nothing special about quadrilaterals in this respect: a similar limitation applies to 'triangle,' 'table,' 'church,' – the list goes on.

It might be thought, though, that what unites all these concepts is not that they are non-normative but, rather, that they are, in a sense to be specified, *stipulative* concepts: they are concepts in which we exercise our freedom to specify which set of things we want to be grouping together under one label. However, an objector might continue,

[4]What if there's a disagreement between A and B both of whom hold that 'wrong' is a minimal moral concept but who propose different accounts of that normative role? Wouldn't that show that it can make sense to question the correctness of a normative concept that consists wholly in its normative role? This would not be an instance of a CCOQ in my sense, which involves a substantive disagreement over which moral claims are true and not the kind of meta-dispute about which account correctly captures what normative roles our basic moral concepts have. For our purposes, we may simply assume that we have to hand correct accounts of those normative roles.

[5]Or, to put the point more carefully, perhaps we could use the thin 'ought' of practical deliberation, the one that we use to express our all-things-considered judgment about what we ought to do, to raise a question about whether we ought to be using the minimal 'ought' of moral deliberation; still, we won't be able to iterate that question any further. At some point we will simply have to assume the integrity of some very thin, purely role-like, normative vocabulary about whose integrity no further normative question is possible. For the purposes of this paper, I will largely ignore this complication, although I will return to it towards the end.

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

not all concepts are like that – in particular, natural kind concepts are not like that. With one of those, we aim to group together things that, in some hard to specify sense, belong together naturally. Given that, won't natural-kind concepts provide a fertile source of concepts for which CCOQs might easily arise?

I don't believe so: the type of open question in the two cases seems to me to be quite distinct. In the case of a normative concept, we are asking whether any of the propositions yielded by the substantive normative concept is actually *true*. In the case of a natural kind concept, we are not asking whether the propositions yielded by the concept are true, but rather whether they are law-like, or counterfactual-supporting, or something along those lines. Even if a concept that purports to name a natural kind – 'grue' for example – does not actually carve at nature's joints, that doesn't entail that the propositions that it yields analytically are not true: things that are green before 3000 (emeralds, for example) are just as much grue as they are green.[6]

I'm not averse to there being types of concepts other than the normative for which a 'conceptual correctness' open question can arise, but I don't believe natural kind concepts provide an example of such a type. Moreover, there is a reason why normative concepts are susceptible to CCOQs, one which may not apply to other concept types, a reason which I will now describe.

We previously saw that the non-negotiable minimal core of a moral concept is its normative role – in the case of a positive concept, its constitutive ties to praise and motivation; in the case of a negative concept, its constitutive ties to criticism, blame and resentment. Call the view that moral concepts consist wholly of their normative roles, with no substantive conception of what falls under them, a Minimalist View of those concepts; and call the view according to which they harbor substantive views about which acts fall in their extensions, a Substantive View.

Consider a negative concept, such as 'wrong,' and let's simplify our account of its normative role so that we say that, at its core, 'wrong' is used to mount a certain distinctive type of criticism of an act, a distinctively *moral* criticism of that act. When we imagine a Substantive version of 'wrong,' what we are imagining is a concept with this critical normative role at its core to which has been superadded a substantive conception of which acts are deserving of that type of criticism.

It should now be clear why a CCOQ will sensibly apply to such a substantive concept of 'wrong' in a way that may not generalize. For when we ask the CCOQ of a given substantive concept of 'wrong,' *we are falling back on its core normative role* to ask whether the substantive list that has been superimposed on that normative role is the one that *ought* to be there. And this latter 'ought' need only be a minimalist 'ought,' since what we are asking in the CCOQ is whether the act- or trait-types contained in the superimposed list are ones that are genuinely deserving of moral criticism. Even if the answer in each case is 'yes,' the question is a coherent one.

What the applicability of a CCOQ to a putatively substantive concept of 'wrong' shows is that our concept of 'wrong' is actually minimalist in nature: it consists merely in its characteristic normative role. That is why it can be used to ask of any act whether it is genuinely wrong.

Of course, we could, by stipulation, introduce a concept – 'wrong*' – that, by definition, contains a substantive conception of the acts deserving of moral criticism; but this concept

[6]If someone insisted that, on their concept of 'water,' it followed analytically that XYZ is also a kind of water, all that would show is that they were not using 'water' as a natural kind concept after all, but rather as a type of 'functional' concept, like 'furniture,' where what matters to something's being 'water' is what it does, and how it looks, etc., rather than what it is composed of.

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

could not be used to settle moral disputes by appeal to concepts, since it would always be vulnerable to a CCOQ, which cannot itself be settled by merely conceptual considerations.

Since our basic moral concepts are just bare normative roles, they engender no contradiction when applied to any act, no matter how clearly that act may fall outside their extensions. And, since they are just bare normative roles, that means that no substantive verdict about right and wrong can be epistemically analytic. And the fact that we cannot treat other concepts – 'simple quadrilateral,' for example – as consisting purely in some core normative role helps explain why these other, non-normative concepts, are *not* susceptible to a CCOQ.

7. The minimalist view and alternative moral concepts

A preoccupation of some recent influential literature within metaethics has been the idea of 'alternative normative concepts' (see Eklund 2017; Dasgupta Ms). The coherence of this idea has been taken to have far-reaching ramifications, including for the very idea of moral or normative realism. Our discussion bears directly on these issues in ways that, I hope, will illuminate them both.

Eklund's (2017) book is framed by a desire to understand whether what he calls 'Ardent Realism' (AR) – the view that "reality itself favors certain ways of valuing or acting," – is a viable position. This is obviously metaphorical and vague, but its core import is supposed to be that disagreements about value have an objective adjudication and aren't just a matter of diverging tastes.

Historically, many philosophers have thought that the correct way to pursue the question about objective adjudication is by asking whether what Eklund labels "Standard Realism" (SR) is true, where SR is the familiar view that:

(I) Fundamental moral principles such as (NoTorture) are capable of truth and falsity.
(II) Some such principles are true.
(III) And those that are true are made true by mind-independent facts.

Eklund denies that the question of the objective adjudication of value can be correctly prosecuted simply by asking whether SR is true. His reason for this striking claim has to do with his taking seriously the idea that the following scenario is metaphysically possible:

> *Alternative*. There is a linguistic community speaking a language much like English, except for the following differences (and whatever differences are directly entailed). While their words "good," "right," and "ought" are associated with the same normative roles as our words "good," "right," and "ought," their words aren't coextensive with our "good," "right," and "ought." So even if they are exactly right about what is "good" and "right" and what "ought" to be done, in their sense, and they seek to promote and to do what is "good" and "right" and what "ought" to be done in their sense, they do not seek to promote what is good and right and what ought to be done. (Eklund 2017: 18)

There are features central to Alternative (to differentiate between our terms and theirs, let's call our term 'wrong,' and theirs 'wrong*').

> (Same) The normative roles of our moral terms and that of the alternative community's are the same. For example, 'wrong' and 'wrong*' are used for exactly the same type of critical role (of the type previously discussed in this paper).

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

(Extensions) However, our terms are not co-extensive. For example, given our term 'wrong,' inflicting pain on a person merely for fun is 'wrong,' while given their term 'wrong*', inflicting pain on a person merely for fun need not be wrong*. (Note: this is not just a matter of our having divergent views about what falls under a uniform concept of 'wrong'; it's that they have a different concept with a genuinely different extension.)

(Bad Guys) From our point of view, their judgments about what they call 'right' and 'wrong' are repugnant to us. They are bad guys.

Assuming Alternative to be a possible scenario, why does this raise a problem for using SR to prosecute the issue of objective adjudication? The argument looks to be something like this:

a. Suppose Alternative is possible.
b. Then we could have two communities, each with its own set of normative concepts, but with different extensions, so that the terms are respectively true of different acts or objects.
c. So, it looks as though it will be possible for the two communities to each be judging correctly *relative* to their *own* sets of basic normative concepts.
d. But, on SR, there will be nothing in the world that can adjudicate between these competing sets of basic normative concepts. (*No Privilege*)
e. I.e., there will be no facts which will tell you which 'ought' is the *better* 'ought' to have, which 'ought' you ought to be operating with.
f. But we wanted fundamental disagreements about value to be objectively adjudicated.
g. So, in the presence of the claim that Alternative is possible, SR seems not to be realistic enough.

Does this argument work? There are two questions here: First, is Alternative really possible? Second, even assuming Alternative to be possible, can we use that fact to raise a problem for SR? Let me take these questions in turn.

There is a prima facie difficulty in supposing Alternative to be genuinely metaphysically possible. We are to imagine that there are two conflicting types of 'ought' concept, 'ought' and 'ought*'which share a core normative role: for example, in both languages, the all-things-considered 'ought' has the role in practical deliberation of settling the question what is to be done.

But if that's so, then how could they end up referring to different properties, and so having different extensions? Wouldn't this normative role be sufficient, along with facts about the world, to determine what lies in the extension of these normative terms? Let's follow Eklund in dubbing the thesis that the normative role is sufficient to determine the extension of the thesis of "Referential Normativity" (RN).

To suppose that RN is false is to assume that the correct theory of reference for such terms must appeal to some further factor, in addition to their normative role, to fix their extensions. For example, it might appeal either to causal connections to the environment, as in the case of natural kind terms, or, as on a Canberra-style view, it might appeal to a set of platitudes about what what's right and wrong, with the reference determined by fit with these associated descriptions.

The Canberra view should be immediately recognizable as amounting to what I earlier called a Substantive view of moral concepts, regarding them as consisting in a body

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

of doctrine over and above their core normative roles. I have already argued that such a Substantive view has been refuted by the susceptibility of moral concepts to CCOQs.

As for the idea that the reference of moral terms might depend on their contingent causal relations to aspects of their environments, although some philosophers have advocated for such a view, construing them as similar to natural kind terms, I believe the view has very little plausibility (see, for example, Horgan and Timmons 1991).

So, there is a very real question, it seems to me, whether Alternative really is metaphysically possible. However, let's set that aside. Let's assume that Alternative is possible, and ask whether we can use it to raise a problem for the adequacy of SR to capture the issue of objective adjudication.

A crucial component of the relevant argument is the thesis of No Privilege: if we confronted two sets of normative concepts which shared their conceptual roles but had divergent extensions, there would be no facts out there about which set of concepts it would be *correct* to use, which ones we *ought* to use.

This, of course, is highly reminiscent of what I was calling the Concept Correctness Open Question. But the moral of my CCOQ is that our basic moral concepts consist purely in their normative roles and so are concepts about which no such CCOQ can be raised. And this is borne out by the very argument being put forward. In order to ask the question to which the No Privilege thesis is an answer, we have to assume the integrity of *some* normative vocabulary: to ask which set of normative concepts we *ought* to use, we have to assume that the notion of 'ought' that we're using is a privileged normative concept with which we are able to raise the question of the correctness of these putative normative alternatives.

The point I am making shows up starkly in Dasgupta's (Ms) paper on the meta-ethics of artificial intelligences, or AIs. Dasgupta wants to argue that there are no facts to discover about the ethics of AIs: it is more a matter of deciding on an ethical code. Dasgupta makes a distinction between parochial ethics and cosmopolitan ethics.

> Ethics involves asking what to do, how to live, and who to be. It therefore requires evaluating actions, projects, and characters. *Parochial* ethics uses our received normative concepts as the evaluative standards. It evaluates the action of destroying Hal according to whether Hal has moral status, whether there are reasons to treat him well, and so on. By contrast, *cosmopolitan* ethics involves asking which concepts to use as our standard in the first place, e.g. reasons or shmeasons. I claim that the ethics of AI must be cosmopolitan. We must recognize that the standards that have suited our limited lives so far might not be suited to the wider world. (Dasgupta Ms: 5)

His argument then goes something like this:

> Suppose that our concept 'should,' delivers a verdict about Hal, say that we should not treat him as we treat other humans. Well, it seems we can always ask: I see that that follows from our concept 'should'. But is that the right concept to be using in doing cosmopolitan ethics? There is another concept, should*, which delivers the opposite verdict. Since there can be no facts in virtue of which one or the other concept 'should' is the right one to use, there can be no facts to discover about the ethics of AI.

Dasgupta is aware that there is a problem here in so much as formulating the question:

> The main objection … I believe, is not that the cosmopolitan question is inappropriate but that it is incoherent. It asks which standard to use, yet this presupposes

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

> that it's *possible* to use an alternative to our parochial standard in the first place. (Dasgupta Ms: 13)

His eventual solution is to substitute for the question

> Which concepts should we use?

the question:

> Are there concepts which limn reality at its normative joints?

But we have yet to be convinced that there is a real alternative question here, since talk of 'normative joints' is just an unhelpful metaphor: it's just another way of talking about which normative concepts we should be using, which of course presupposes the integrity of our own minimalist normative concepts. We haven't got as far as setting up a putative dispute between parochial and cosmopolitan ethics.

8. Consequences

If these considerations are correct, then at least three important consequences follow.

First, and as I previously emphasized, we cannot think of the a priori justified moral principles that we are tempted to endorse as simply revealing to us information that is already encoded within our moral concepts. They are, rather, synthetic a priori. (Although I continue to talk primarily of morality, my claims apply to normative principles quite generally, since any fundamental normative concept can be shown, via the arguments I've developed, to consist merely in its characteristic normative role.)

Second, if these a priori justified moral judgments are supported by intuitions, as they seem to be, then those intuitions cannot be thought of as simply revealing to us the nature of our moral concepts, but must be thought of more ambitiously, as revealing to us the true facts about moral right and wrong. If intuition-backed moral knowledge is to be possible, then there must be a type of intellectual seeming whose justificatory power is not sourced in the understanding alone.

The acknowledgement of such non-understanding-sourced intuitions will sit ill with the theories of many philosophers, including many of those who are friendly to intuitions. Philosophers such as Sosa, Goldman and Pust, for example, all insist that intuitions ultimately derive whatever probative power they have from being sourced in the understanding (see Sosa 1998, 2007; Pust 2000; Goldman 2007). However, if my argument works it shows the need to recognize the existence of intuitions whose probative power derives not merely from our understanding of concepts. To put the point another way, we must recognize a distinction between a proposition that is *self-evident* – that is justifiably believable simply on the basis of being understood – and a proposition that is *intuitive* – that is justifiably believable not merely on the basis of being understood but also because, when one considers it, one has the intellectual impression that it is true.

Finally, my argument implies that considerations of meaning constitution do not constrain *disagreement* in normative domains in the way in which they do in non-normative domains, such as those of mathematics or logic. Even Clarke-Doane (2020), who is eager to claim that there may be as much disagreement within mathematics and set theory as there is in ethics, concedes that anyone denying the axiom of Extensionality or the axiom of Pairing would betray a lack of conceptual competence:

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press

> Consider the Axiom of Extensionality, $\forall x \forall y \forall z[(z \in x \leftrightarrow z \in y) \rightarrow (x = y)]$. This says that if "two" sets have the same members, then they are really one and the same (the converse is a logical truth in first-order logic with identity). Perhaps this is some kind of truism about sets. Similarly, the Axiom of Pairing, $\forall x \forall y \exists z \forall w [w \in z \leftrightarrow (w = x \vee w = y)]$, says that, whenever you have "two" (perhaps not distinct) sets, there is another containing just those "two". This also seems hard to deny … (Clarke-Doane (2020: 40–1)

And even Quine (1974), the most implacable enemy of epistemic analyticity, conceded that in certain cases the deviant logician would simply be changing the subject.

But, as we have seen, in the case of moral propositions, you can deny *any* substantive moral claim without necessarily being guilty of conceptual incoherence. This makes it possible to have a wider range of disagreement about moral propositions than is possible for non-normative domains, like mathematics, a possibility that has clearly been exploited by people across cultures and history.[7]

References

Boghossian P. (1996). 'Analyticity Reconsidered.' *Noûs* **30**(3), 360–91.
Boghossian P. (2003). 'Blind Reasoning.' *Aristotelian Society Supplementary Volume* **77**(1), 225–48.
Boghossian P. and Williamson T. (2020). *Debating the a Priori*. Oxford: Oxford University Press.
Clarke-Doane J. (2020). *Morality and Mathematics*. Oxford: Oxford University Press.
Dasgupta S. (Ms). *The Meta-Ethics of Artificial Intelligence: Are We Beholden to Normative Joints?*
Eklund M. (2017). *Choosing Normative Concepts*. Oxford: Oxford University Press.
Goldman A.I. (2007). 'Philosophical Intuitions: Their Target, Their Source, and Their Epistemic Status.' *Grazer Philosophische Studien* **74**(1), 1–26.
Harman G. (1986). *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press.
Horgan T. and Timmons M. (1991). 'New Wave Moral Realism Meets Moral Twin Earth.' *Journal of Philosophical Research* **16**, 447–65.
Hume D. (1978). *A Treatise of Human Nature* (L.A. Selby-Bigge and P.H. Nidditch, eds). Oxford: Clarendon Press.
Moore G.E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.
Pust J. (2000). *Intuitions as Evidence*. London: Taylor & Francis.
Quine W.V. (1951). 'Two Dogmas of Empiricism.' *Philosophical Review* **60**, 20–43; reprinted in W.V. Quine (1953). *From a Logical Point of View*, pp. 20–46. Cambridge, MA: Harvard University Press.
Quine W.V. (1974). *The Roots of Reference: The Paul Carus Lectures*. Chicago, IL: Open Court Publishing.
Sosa E. (1998). 'Minimal Intuition.' In M. DePaul and W. Ramsey (eds), *Rethinking Intuition*, pp. 257–69. Lanham, MD: Rowman and Littlefield.
Sosa E. (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I*. Oxford: Clarendon Press.
Williamson T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell.

Paul Boghossian is Silver Professor of Philosophy at New York University and the director of the New York Institute of Philosophy. His research interests are primarily in epistemology and the philosophy of mind, although he has written on a wide range of topics, including: color, rule-following, naturalism, self-knowledge, a priori knowledge, analytic truth, realism, relativism, the aesthetics of music and the concept of genocide. His most recent book is *Debating the A Priori*, co-authored with Timothy Williamson.

[7]For helpful comments on an earlier version of this essay I am grateful to Yu Guo, Antonella Mallozzi, Declan Smithies and Crispin Wright.

Cite this article: Boghossian P (2021). Normative Principles are Synthetic A Priori. *Episteme* **18**, 367–383. https://doi.org/10.1017/epi.2021.32

https://doi.org/10.1017/epi.2021.32 Published online by Cambridge University Press