Volume X ◇ Issue 2

2019

Logos & Episteme

an international journal of epistemology

Romanian Academy
Iasi Branch
◆◆◆
"Gheorghe Zane" Institute for Economic and Social Research

Contact address:

Institutul de Cercetări
Economice şi Sociale „Gh.Zane”

Iaşi, str.T.Codrescu, nr.2, cod 700481

Tel/Fax: 004 0332 408922

Email: logosandepisteme@yahoo.com

http://logos-and-episteme.acadiasi.ro/

https://www.pdcnet.org/pdc/bvdb.nsf/journal?openform&journal=pdc_logos-episteme

ISSN 2069-053

TABLE OF CONTENTS

RESEARCH ARTICLES

THERE ARE ACTUAL BRAINS IN VATS NOW

Adam Michael BRICKER

ABSTRACT: There are brains in vats (BIVs) in the actual world. These "cerebral organoids" are roughly comparable to the brains of three-month-old foetuses, and conscious cerebral organoids seem only a matter of time. Philosophical interest in conscious cerebral organoids has thus far been limited to bioethics, and the purpose of this paper is to discuss cerebral organoids in an epistemological context. In doing so, I will argue that it is now clear that there are close possible worlds in which we are BIVs. Not only does this solidify our intuitive judgement that we cannot know that we are not BIVs, but it poses a fundamental problem for both the neo-Moorean (i.e. safety-based) anti-sceptical strategy, which purports to allow us to know that we aren't BIVs, and the safety condition on knowledge itself. Accordingly, this case is especially instructive in illustrating just how epistemologically relevant empirical developments can be.

KEYWORDS: BIVs, cerebral organoids, scepticism, neo-Mooreanism, safety

0. Introduction

When Hilary Putnam first introduced the "case of the brains in a vat," he categorised BIVs as a "science fiction possibility."[1] A mere three decades later, this science fiction possibility is now a rapidly advancing research programme in developmental and molecular biology. Taking only a slight liberty with the word "vat," but none whatsoever with the word "brain," we might now observe that there are quite literally brains in vats. These BIVs—cerebral organoids—are not quite as Putnam imagined. They are not harvested, but rather grown from stem cells *in vitro*. More significantly, they are not yet comparable to conscious human brains, but closer to that of a three-month-old foetus.[2] Nevertheless, the science of cerebral organoids is advancing at a remarkable rate,[3] and bioethicists are already

[1] Hilary Putnam, *Reason, Truth and History* (Cambridge: Cambridge University Press, 1981), 5.

[2] Iva Kelava and Madeline A. Lancaster, "Dishing out mini-brains: Current progress and future prospects in brain organoid research," *Developmental Biology* 420, 2 (2016): 199-209.

[3] See Kelava and Lancaster, "Dishing out mini-brains," Stefano L. Giandomenico and Madeline A. Lancaster, "Probing human brain evolution and development in organoids," *Current Opinion in Cell Biology* 44 (2017): 36-43.

looking forward to a future containing conscious cerebral organoids.[4] As I will argue below, we are unsettlingly close to the world Putnam imagined.

This paper explores an especially noteworthy epistemological consequence of the modal closeness of BIV-containing worlds: We cannot know that we aren't BIVs. Here we will see this consequence play out not only at the intuitive level, but also with regard to the neo-Moorean anti-sceptical strategy, which purports to allow us to know that we aren't BIVs. In short, not only is the safety condition unable to successfully anchor an anti-sceptical argument, but it now faces a sceptical problem of its own. In arguing along these lines, this paper will proceed in the following way: First, I will discuss the BIV scenario through the lens of epistemology, focusing specifically on its relationship with neo-Mooreanism (§1). Next, I will provide a brief, non-technical introduction to the actual-word BIVs that are cerebral organoids (§2). Finally, I will argue that the present status of cerebral organoids in the actual world indicates that there are close possible worlds in which any number of us are cerebral organoids (§3). It follows directly from this that we cannot know that we aren't cerebral organoids, that neo-Mooreanism is unable to explain how we might have this knowledge, and that the safety condition on knowledge now faces its own sceptical problem.

1. Brains in Vats

In this section I discuss the epistemological significance of brains in vats. After introducing the case of BIVs, I'll discuss neo-Mooreanism, the predominant anti-sceptical strategy that seeks to explain how we might know that we aren't BIVs.

Let's begin with Putnam's original BIV scenario:

> Imagine that a human being (you can imagine this to be yourself) has been subjected to an operation by an evil scientist. The person's brain (your brain) has been removed from the body and placed in a vat of nutrients which keeps the brain alive. The nerve endings have been connected to a super-scientific computer which causes the person whose brain it is to have the illusion that everything is perfectly normal. There seem to be people, objects, the sky, etc; but really all the person (you) is experiencing is the result of electronic impulses travelling from the computer to the nerve endings.[5]

[4] Megan Munsie, Insoo Hyun, and Jeremy Sugarman, "Ethical issues in human organoid and gastruloid research," *Development* 144, 6 (2017): 942-945, Andrea Lavazza and Marcello Massimini, "Cerebral organoids: Ethical issues and consciousness assessment," *Journal of Medical Ethics* 44, 9 (2018): 606-610, Joshua Shepherd, "Ethical (and epistemological) issues regarding consciousness in cerebral organoids," *Journal of Medical Ethics* 44, 9 (2018): 611-612.

[5] Putnam, *Reason, Truth and History,* 5-6.

When Putnam first put forth this BIV scenario, it was for the purpose of illustrating a semantic point about reference.[6] Nevertheless, brains in vats proliferated throughout philosophy in the decades following, including not only the philosophy of language, but also the philosophy of mind, metaphysics, and epistemology.[7] Given the breadth of the impact of the BIV scenario, here I will focus only on the epistemological significance of brains in vats. Within an epistemological context, the BIV scenario is something like the contemporary successor to Descartes' Evil Demon. While Descartes imagined a wholesale illusion of the external world as the product of supernatural forces, not the presumably naturalistic processes employed by the evil scientist, the upshot is the same: Were we deceived by an evil demon, just as if we were BIVs, all our experiences would be indistinguishable from what they are now.

This type of scenario forms the basis for the familiar sceptical argument from closure: It seems we intuitively judge that we cannot know that we aren't BIVs. After all, were we BIVs, all our experiences might be identical to how they are in the actual world. Moreover, intuitively, it also seems that knowledge is closed under known entailment. That is, if S knows that p, and S knows that p entails q, then S knows that q. However, if we commit to both these claims, this entails that we cannot have any knowledge of the external world that is inconsistent with being a BIV, which is to say most all knowledge of the external world. Following closure, knowing about the external world would entail knowing that we aren't BIVs.

Defeating this sceptical argument has been one of the primary aims of contemporary epistemology. One widely employed strategy involves denying or restricting closure in some way, which would allow us to have knowledge of the external world despite not knowing that we aren't BIVs.[8] However, here I will focus on one predominant anti-sceptical approach that purports to allow us to know that we aren't BIVs: neo-Mooreanism.

In his infamous argument against scepticism, G.E. Moore[9] offered a simple inversion of the above sceptical argument: If (i) we know some ordinary

---

[6] Roughly, Putnam maintained that the thoughts of a BIV cannot refer to entities in the external world. Therefore, the thought, "I am a BIV," cannot be true, even when thought by a BIV.

[7] For an overview, see *The Brain in a Vat,* ed. Sanford Goldberg (Cambridge: Cambridge University Press, 2016).

[8] For an introduction, see Steven Luper, "Epistemic Closure," *The Stanford Encyclopedia of Philosophy,* ed. Edward N. Zalta (Spring 2016 Edition), https://plato.stanford.edu/archives/spr2016/entries/closure-epistemic.

[9] George Edward Moore, "Proof of an External World," *Proceedings of the British Academy* 25 (1939): 273–300.

proposition, e.g. that we have hands, and (ii) we know that this ordinary proposition entails the negation of the sceptical hypothesis (i.e. that we aren't BIVs), then (c) we know the negation of the sceptical hypothesis. While this argument finds little love in contemporary epistemology, multiple prominent philosophers (see below) have defended a more robust version of Moore's argument, dubbed "neo-Moorean" by Pritchard.[10] The key difference between Moore and neo-Moorean arguments is that the latter approach seeks a broader theoretical motivation for our ability to know that we aren't BIVs.[11] While Moore appealed, unconvincingly, only to our common sense, pre-theoretic intuitions about knowledge of the external world,[12] neo-Moorean arguments seek to establish this via some general epistemological principle.

The preferred epistemological principle of the neo-Moorean response is *safety*: If S knows that p, then there are no close possible worlds in which S falsely believes that p. Assuming, as is standard, that there are no close possible worlds in which we are BIVs, this allows us to have ordinary knowledge of the external world, as there are ordinarily no close possible worlds in which we are mistaken about such things. Moreover, safety even allows us to know that we aren't BIVs, again assuming that there are no close worlds in which we might be BIVs and accordingly believe falsely that we aren't BIVs. In this way, neo-Mooreanism preserves both our knowledge of the external world and closure. Different versions of this of anti-sceptical approach have been advanced by Sosa,[13] Williamson,[14] and Pritchard.[15]

In section 3, I will argue that the core assumption of neo-Mooreanism—i.e. that there are no close possible worlds in which we are BIVs—seems mistaken

[10] Duncan Pritchard, "Recent Work on Radical Skepticism," *American Philosophical Quarterly* 39, 3 (2002): 215-257, 237

[11] See Duncan Pritchard, "Contemporary Neo-Mooreanism," in *Epistemological Disjunctivism* (Oxford: Oxford University Press, 2012).

[12] See Duncan Pritchard, "Mooreanism," in *Epistemological Disjunctivism* (Oxford: Oxford University Press, 2012).

[13] Ernest Sosa, "How to Defeat Opposition to Moore," *Philosophical Perspectives* 13(1999): 141-153.

[14] Timothy Williamson, *Knowledge and Its Limits* (Oxford: Oxford University Press, 2000), chapter 8.

[15] Duncan Pritchard, "Resurrecting the Moorean Response to the Sceptic," *International Journal of Philosophical Studies* 10, 3 (2002): 283-307 and "How to be a Neo-Moorean," in *Internalism and Externalism in Semantics and Epistemology,* ed. Sanford Goldberg (Oxford: Oxford University Press, 2007).

given recent advancements developmental biology. However, first I need to say a bit on what these advancements actually are.

2. Brains *in Vitro*

The purpose of this section is to introduce the basics of cerebral organoids to epistemologists. Accordingly, this discussion will be limited in scope to two epistemologically interesting features of cerebral organoids:[16] (1) Cerebral organoids are quite literally brains in vats;[17] (2) although they are not functionally or structurally analogous to conscious human brains, it seems that this is only a matter of time. I will discuss both points in turn.

Let's begin with a brief overview of these brains *in vitro*. Cerebral organoids begin as stem cells, which, under conditions that partially mimic those that facilitate early embryonic neural development *in utero*, quickly grow into recognisable cortical and sub-cortical structures.[18] Neurons begin to appear in 8-10 days, with recognisable brain structures appearing in under a month.[19] Cerebral organoids famously recapitulate *in utero* development seen between 1 to 4 months of gestation.[20] At the structural level, cerebral organoids display a number of distinct neural regions, including forebrain/hindbrain differentiation, a hippocampus, a choroid plexus, an immature retina, and a cortex with distinct dorsal (including a prefrontal cortex and occipital lobe) and ventral regions.[21] The last of these is most important, as it is the cortical areas that ultimately provide the neural basis of higher-level cognitive function in further developed brains. Beyond these fascinating structural properties, cerebral organoids also display remarkable functional characteristics. Cortical neurons in cerebral organoids form mature neurons with functioning synapses, which display spontaneous neural activity.[22] To be clear, the neurons in the cortex of cerebral organoids—*in vitro* brains that presently develop to the rough equivalent of a three-month-old foetus—actually fire! In short, it is no exaggeration to say that there are presently brains in vats.

---

[16] For a technical introduction, see Madeline A. Lancaster, Magdalena Renner, Carol-Anne Martin, Daniel Wenzel, Louise S. Bicknell, Matthew E. Hurles, Tessa Homfray, Josef M. Penninger, Andrew P. Jackson, and Juergen A. Knoblich, "Cerebral organoids model human brain development and microcephaly," *Nature* 501, 7467 (2013): 373-9.

[17] Again, this might depend on your definition of "vat."

[18]Byoung-il Bae and Christopher Walsh, "What are mini-brains?" *Science*, 342, 6155 (2013): 200-1.

[19] Bae and Walsh, "What are mini-brains?"

[20] See Kelava and Lancaster, "Dishing out mini-brains."

[21] Lancaster et al., "Cerebral organoids model human brain development and microcephaly."

[22] Lancaster et al., "Cerebral organoids model human brain development and microcephaly."

This is not a sci-fi future, as it was when Putnam first described the BIV scenario, but an actual feature of the actual world.

At this point, it's important to not get ahead of ourselves. Are there actual brains in vats? Yes. Are these Putnam-esque brains in vats? No. Cerebral organoids aren't yet grown in conditions similar enough to those found *in utero* to allow for later stages of neural development, nor are they vascularised, which is also necessary for later cortical development.[23] In short, today's cerebral organoids are almost certainly not conscious. The vat technology just isn't there yet. Nevertheless, it is not unreasonable to expect that the *conscious* cerebral organoid is only a matter of time. Indeed, there appears to be a growing sense among researchers that this is the case. One indication of this comes from the recent interest in cerebral organoids displayed by medical ethicists, who have already begun laying the groundwork for the ethics of handling conscious cerebral organoids.[24] Moreover, researchers note that cerebral organoid technology is "extremely fast-moving,"[25] and that only a decade ago, today's level of *in vitro* cerebral development was "thought to be unattainable."[26] Further still, as noted by Kelava and Lancaster, current limitations don't appear to be in any way insurmountable:

> It is easy to envisage that in 10–20 years from now (or even less) we will be able to almost fully mimic development of certain tissues *in vitro.* In addition, further improvements in the technique might allow us to model adult brain physiology and disorders of the adult and ageing brain.[27]

As I will argue in the next section, this alone poses a significant problem for both neo-Mooreanism and the safety condition generally. The key here is that while there aren't yet conscious cerebral organoids in the actual world, from the epistemological perspective this doesn't actually matter. They are widely thought to be on the horizon, not a new technology as much as an incremental improvement on existing technology.

---

[23] Kelava and Lancaster, "Dishing out mini-brains."

[24] Lavazza and Massimini, "Cerebral organoids: Ethical issues and consciousness assessment," Shepherd, "Ethical (and epistemological) issues regarding consciousness in cerebral organoids."

[25] Giandomenico and Lancaster, "Probing human brain evolution and development in organoids," 41.

[26] Kelava and Lancaster, "Dishing out mini-brains," 199.

[27] Kelava and Lancaster, "Dishing out mini-brains," 205.

Finally, before moving on, I want to highlight one crucial point that is all too easy to miss if we don't discuss cerebral organoids simultaneously from the epistemological and developmental-neuroscientific perspectives: The very conditions necessary to create neural architecture seen in the postnatal brain, if applied to cerebral organoids, would necessarily transform the cerebral organoid into a sceptical scenario. One of the most basic facts of postnatal neural development is that it is dependent upon perceptual input from external sources.[28] The neural circuits and functional connections found in the mature brain require external signals in order to form. Visual pathways won't develop without visual input; auditory pathways won't develop without auditory input, etc. This means that developing cerebral organoids that are analogous to even the infant brain will require some sort of targeted stimulation of the relevant neural pathways.[29] However, this is all it takes to transform the cerebral organoid into a Putnam-esque BIV. Even crude stimulation of the primary visual cortex, which is easily accomplished non-invasively with current technology (i.e. TMS), will result in consciously perceptible flashes of light or disturbances in the visual field. Moreover, further development of the visual pathway would require more complex inputs, which would in turn produce more complex conscious perceptions. The point here is that simply getting anything close to a mature brain would require giving the cerebral organoid false perceptual input, and the more developed we want it, the more sophisticated this false perceptual content needs to be. That is to say, it requires that we play a role eerily close to Putnam's "evil scientist."

3. Brains in Close Possible Words

The existence of cerebral organoids has immediate consequences for how we think about scepticism and sceptical hypotheses. In this section, I want to discuss these consequences in two different ways. First, at the intuitive level, the existence of cerebral organoids fundamentally changes our judgements about whether we can know that we aren't BIVs. The rapid development of cerebral organoids solidifies

[28] For a textbook introduction, see Dale Purves, *Neuroscience*, (Sunderland, Mass.: Sinauer, 2016), ch. 25.

[29] It should be noted that "artificial" neural stimulation is well within the means of current medical technology. One excellent example of this comes from cochlear implants, which entirely bypass biological transduction of sound and directly stimulate the auditory nerve with electrical impulses. In principle, there is no reason why a cochlear implant must transmit auditory signals from the external world instead of, say, artificial ones generated to correspond with a simulated reality.

what were previously fluid or ambiguous intuitions: We cannot know that we aren't BIVs. Second, these intuitive attitudes are reflected in the problems cerebral organoids pose for safety and the neo-Moorean anti-sceptical strategy. In short, the existence of cerebral organoids in the actual world means (i) that there are close possible worlds in which we are BIVs and thus (ii) neo-Mooreanism is unable to explain how we might know that we aren't BIVs, or even ordinary information about the external world. Moreover, (iii) the safety condition now appears to *entail* that we cannot have knowledge of the external world, giving us ample reason to reject it on anti-sceptical grounds.

Let's begin with the intuitive judgements engendered by the emergence of cerebral organoids. Put simply, I think we clearly and unambiguously judge that we cannot know that we are not cerebral organoids. This marks a departure from the intuitions elicited by distant-world sceptical hypotheses, which are notably unstable and varied. It sometimes might seem that we can know that we aren't deceived by an evil demon, while in other cases it may not. While advocates of denying closure concede that that we cannot know the negation of sceptical hypotheses, neo-Mooreans have thought we might know the negation of sceptical hypotheses due to a theoretical feature of knowledge, and Moore of course thought it just obvious. The point here is that this sort of ambiguity no longer seems appropriate. Our intuitive response to learning about cerebral organoids is markedly different from that elicited by Putnam's BIV counterfactual. There are already BIVs. There will soon be conscious BIVs, and there is even already talk of connecting "numerous human organoids into working complexes" in the near future.[30] It now seems chillingly obvious that we cannot know that we aren't BIVs, and that any theory of knowledge that entails that we have this knowledge is simply mistaken.

The observation that we intuitively judge we cannot know that we aren't cerebral organoids is reflected in the problems these actual-world BIVs pose for the neo-Moorean response to scepticism. First, the neo-Moorean response, insofar as it is committed to the safety condition on knowledge, only works if there are no close possible worlds in which a sceptical hypothesis maintains. If there are close possible worlds in which we are BIVs, then those are close possible worlds in which we falsely believe that we aren't BIVs, meaning that such beliefs aren't safe. The problem for neo-Mooreanism, and the safety condition on knowledge generally, is that it certainly seems like there are close possible worlds in which we are cerebral organoids. While it is of course difficult to definitely settle questions of modal ordering, I think I might illustrate the modal closeness of these sceptical-

[30] Munsie et al., "Ethical issues in human organoid and gastruloid research," 943.

hypothesis worlds in the following way: First, as discussed above, conscious cerebral organoids seem an inevitability of the near future. With improvements in *in vitro* conditions and complimentary developments in the sophistication of neural stimulation, there will be conscious cerebral organoids in the actual world. The key here is that remarkably little needs to change in the actual world for conscious cerebral organoids, which suggests that there are close possible worlds in which this is the case. Let's call these W1 worlds. Next, not too distant from W1 worlds are worlds in which any one of us is a cerebral organoid. Let's call these W2 worlds. Again, the key here is that not much needs to change from W1 worlds for any given individual to be a conscious cerebral organoid. All that is necessary is that your parents opted to have a lab grow a cerebral organoid you for the purposes of "personalised medicine," as is already being envisioned as a primary application of the technology.[31] As discussed above, developing a mature brain requires complex external stimuli, which your W2 parents opt for in the interest of ensuring your non-organoid siblings won't suffer from neurodegenerative disorders later in life.

The question at this point is whether the changes between the actual world, the closest W1 world, and then the closest W2 world to that are so great as to preclude W2 worlds from being close to the actual world. When viewed in this manner, we can understand just how close W2 worlds are. The change from the actual world to W1 is simply a modest, easily foreseeable advancement in technological capabilities. Moreover, the change from W1 to W2 could be as little as your parents deciding not to grow a cerebral organoid before having you (W1) vs growing a cerebral organoid that is you (W2). Put this way, the modal closeness of W2 worlds to the actual world is striking, and more than a little unsettling. As discussed above, this poses a fundamental challenge for neo-Mooreanism, which requires that there are no such close W2 worlds. Not only is the neo-Moorean argument unable to explain how we might have knowledge that we aren't BIVs, it additionally fails to explain how we might have ordinary knowledge of the external world.

Further still, if we accept that there are (1) close possible worlds in which we are cerebral organoids and (2) that safety is a necessary condition on knowledge, this entails that we have little, if any, knowledge of the external world. Given (1), for (at least) most any belief p by S about the actual external world, there will be a close possible world in which cerebral organoid S falsely believes that p. However, the very definition of the safety condition on knowledge is that such beliefs cannot be known. In a fascinating interplay between epistemology and

[31]Kelava and Lancaster, "Dishing out mini-brains," 206.

developmental biology, we can understand that the emergence of cerebral organoids has transformed safety from a plausible inoculation against scepticism into a premise in its own sceptical argument. Accordingly, it seems we now have good reason to reject safety on anti-sceptical grounds.

There is a worry here that all this might be premature. Perhaps the worlds in which we are BIVs won't be close enough to pose a problem for another 20 or so years, when there are conscious cerebral organoids in the actual world. While I understand this sentiment, anti-sceptical strategies and necessary conditions on knowledge shouldn't come with expiration dates. If we think that safety might not be a necessary condition on knowledge 20 years from now, then there is something profoundly mistaken about maintaining that it is now. Beyond this, there is an additional worry that perhaps the emergence of cerebral organoids simply means that we cannot have knowledge about the external world. Unfortunately, I cannot rule this out. However, as such a concession would of course represent a new paradigm in epistemology, I will not address it here.

Additionally, one could object that there isn't actually anything new here, and that any epistemologically interesting work done by modally close BIVs was already possible via modally close dreamers (i.e. close possible worlds in which we could have been dreaming whatever it is we perceive in the actual world).While an understandable impulse, this objection I think exaggerates the extent to which the phenomenal content of dreams might recapitulate that of conscious perception. Yes, there is a sense in which, when we dream, we might form something like false beliefs. However, there is a major concern here that these beliefs will never share the same content as those formed during waking consciousness. When I dream that I'm writing a philosophy paper, I don't form beliefs with comparable content to those I am forming now. Any belief content in dreams, if we might call even it that, is lo-fi, confused, and incoherent. The key to BIVs is that, unlike dreams, their mental content might exactly mimic that of the (waking) actual world. Perceptual input might even follow the same visual and auditory pathways, via the retina and auditory nerve, respectively. Given some perceptual belief that p, with content derived from actual experience, this is what is needed to underwrite a false belief that p. Accordingly, it is clear BIVs can do far more epistemological work here than dreams.

Finally, I want to say a bit more about the significance of this paper. Admittedly, support for safety appears to be waning, not the least among those cited above as initial proponents, and neo-Mooreanism is rather niche. However, it's important not to miss that there is a broader lesson to be learned here about the relationship between empirical science and technological advancement on the one

hand, and epistemological theory on the other. Despite the increasing incorporation of empirical findings and methods into epistemology, this tends to be rather shallow and one-dimensional, focusing largely on the scientific description of patterns of knowledge attribution.[32] Even you aren't especially interested in safety or neo-Mooreanism, it is crucial to understand that even fields like developmental and molecular biology can have non-trivial epistemological implications. While in this case it happens to be for accounts that were already a bit past their prime, the takeaway is still clear. We as epistemologists need to be sensitive to the fact that empirical developments from far outside our orbit might be directly relevant to the accounts we develop. The onus is on us to seek out such findings.

In conclusion, recent advances in developmental biology have precipitated a shift in how we might think about sceptical hypotheses: There are now close possible worlds that contain sceptical hypotheses in which any number of us are BIVs. This poses a foundational problem for the neo-Moorean anti-sceptical strategy, as well as the safety condition on knowledge, and it seems quite clear at this point that we cannot know that we aren't cerebral organoids. Moreover, it serves to underscore the unexpected ways in which empirical developments might be epistemologically relevant.

[32] For example, see Jennifer Nado, *Advances in Experimental Philosophy and Philosophical Methodology* (New York: Bloomsbury, 2016).

POSSIBILITY VERSUS POSSIBLE WORLDS

James CARGILE

ABSTRACT: It is a common idea in philosophy that some false propositions such as (C) that Charlottesville is the largest city in Virginia, have the property of being possibly true. It is not a clear idea but an important one which has inspired considerable effort at clarification. One suggestion is that there exist (really, not just possibly) "possible worlds" in which C or some suitable facsimile is true. One further attempt at clarification on offer is that there exists (again, really) a maximal consistent set of propositions containing C. It is argued here that these attempts at clarification are profoundly erroneous. There exist actual powers of imaginative construction which would yield a scenario sufficiently detailed to be recognized by competent reviewers as one in which C is true. (The depiction might be in film or narrative and would avoid analytic falsehoods.) This is a frail clarification, vulnerable to questions, but is the best possible direction for a clear idea of the possibility of the proposition. The notion of possible worlds is associated with very valuable work in mathematical logic. It can only improve our appreciation of this excellent work to separate it from cloudy metaphysics.

KEYWORDS: possible, maximal consistent, proposition

1. It can be difficult to explain to students that the proposition (C): that Charlottesville is the largest city in Virginia, is contingent, neither necessarily true nor necessarily false. They will agree that it could turn out that Charlottesville is the largest, meaning that it could become the largest, in a variety of possible ways. This brings out a primary problem – about the proposition itself, before explaining its possibility or contingency. By C, I mean the proposition that Charlottesville is now the largest city in Virginia, a proposition we know is false. It is strictly impossible that this, or any other proposition, should become true. No proposition ever changes in truth value.

Unfortunately, this is not accepted by all parties to this discussion. It is common to say that the proposition that I am sitting is true at some times and false at others. When I stand up, the proposition becomes false. Is it that, when I said I was sitting, I was right, but now I am wrong? Propositions should be understood as clear about time references important to truth. If I say truly the sentence "I am sitting," the sentence is poorly described as "what I said," "the thing I asserted," because the sentence (in reference to me) will soon be false and that does not affect what I said.With such propositions as that 2+2=4, time reference is not important.

It often is important. Any proposition, if true, is true at all times, if false, false at all times. Its truth value may not be settled until later, but that does not mean that it will not have one until then.

The proposition C would have been difficult for an ancient Greek to grasp. Abstruse propositions of modern science would be even more difficult. Understanding might require lengthy explanation, of the sort given to modern students in the course of their education. To identify the time referred to in C would have been difficult in ancient Greece. The lengthy explanation would not be part of the content of the proposition, just part of the requirement for grasping it.

The possibility or even the nonzero probability of the sentence used to express C coming to express a true proposition is not what is meant in my saying that the proposition C is possible. C is about the position of Charlottesville in the present ranking. Some philosophical skeptics will agree that it could at present be the largest, appealing to a metaphysical skepticism that has us somehow out of touch with reality. This would be explaining the possibility of C in terms of another possibility – our being radically deceived about C. That is a logical possibility, the possibility of the proposition that we are at present mistaken in believing that C is false. Citing that possibility does not offer any progress on the project of explaining what it means to say that C is possible. In fact, it would be a regression. Unfortunately "It is possible that P" is commonly used in this way, to claim that P may be true, unbeknownst to us.

The right approach is that there actually exist imaginative exercises which qualify as imagining that C is true. We would want to avoid the natural "It is possible to imagine that C," if we aim for a noncircular explanation. You can assure someone that they can imagine it in the sense that they can do twenty push-ups. Even if presently not successful, it is a lawlike truth that someone with their makeup will succeed with determined effort over a short period of time. This use of "imagine" could be quite literal for such a possibility as my turning bright red. You would simply call up an image of me, colored bright red. True, there are philosophical objectors to the existence of mental images, but the project will make sense to most people, without any actual visual images. Visualizing does not require having visual images, though images may help and reinforce confidence about imagining. "Imagining" C is very different. Actual visual images do not seem helpful. But in a looser use of "imagine," the most common use, many of us have the ability to tell a story in which C is true. We actually possess a working method of storytelling which will produce a story people will recognize as one in which Charlottesville is presently largest.

It is fair, and an understatement, to complain that this general explanation of possibility is unclear. Possibility in general is unclear. Clarity for the possibility that P increases about particular cases the more naturally everyone recognizes a story as one which successfully depicts P as true. "Here's a story: C is true" is a poor showing. "Once upon a time, lots of bachelors were married" is another bad attempt on a different proposition. The C story could be improved by suitable imaginary details. Suitability is somewhat vague and subjective, and being stuck with "It is possible to provide suitable details" is circular. With the bachelor story, rich details could be provided so that many would describe the result as a story in which bachelors are married. That does not establish the possibility of married bachelors. It is not a successful depiction. How can we prove such a depiction cannot be successful? We can deduce a contradiction from the premise that X is a married bachelor.

Deduction in some formal systems, such as standard propositional or predicate calculus, is defined with mathematical precision. We must not let this foster illusions about the precision of the general idea of deduction in natural language. Some philosophers hold that there are contradictions which are true. If "A proves P to B" entails that B accepts P, then we may not be able to prove the impossibility of married bachelors to everyone.

Whether a story "to the effect that P" establishes P as possible is vague and subjective. To argue the impossibility of P by trying to come up with an acceptable story and failing is even weaker. It is important to recognize the weaknesses of the popular conception of necessity, possibility and contingency. That does not mean there are no such properties. Even eminent thinkers may be unable to reach agreement about the possibility of certain propositions. Their imaginative powers and logical skills may not be adequate for the particular task. They are still adequate for many cases and provide reason to believe that our imaginative power, along with deductive skill, shows us the presence of a feature, possibility, possessed by some propositions and lacked by others, even though these powers vary widely among people and are fallible for everyone.

The possibility of the proposition that I am standing right now can easily be proved by imagining. Imagining that C is more difficult. Imagining that the Earth is round has given many people trouble. Such variability about whether a certain fabric sample is red, is good reason to suspect that "red" does not generally express any objective property. Some say the same thing about goodness. That is spiritually bad. Possibility is not as important as goodness, but it is objective in spite of similar difficulties.

For any true proposition, we can establish possibility without depending on imagination, by verifying it. That proves that it is possible. Nothing impossible is true. We could sail around the Earth and prove its roundness without being able to visualize the whole Earth, or needing to. For the possibility of false propositions, imagination is crucial. This is bad news for verifying possibility claims, because imagination is so variable. For false P, telling a story which qualifies as successful imagining that P is a frail criterion for possibility. We should recognize that it is the best we have.

2. Successful story telling might be said to describe a "world" in which C is true. This could be a harmless figure of speech. However, there are those who hold that saying that there is a possible world in which P is true makes explicit something important. One leading authority writes as follows: "It is uncontroversially true that... things could have been different in countless ways... I therefore believe in the existence of entities that might be called 'ways things might have been.' I prefer to call them 'possible worlds'."[1]

It is not disputed here that things might have been different in countless ways, one way being C's being true. But C's being true, though a way things might have been, is not a possible world in the meaning of the above author. This is not based on a position as to what "possible world" must mean, but rather, on the assumption that the author would not consider C's being true a possible world. Insofar as C is a way things might have been but not a possible world, being a way things might have been is not sufficient for being a possible world.

We should all agree that, in order for things to be otherwise as with C, lots of other things would also have to have been otherwise. C being true would require a change in the truth value of indefinitely many propositions. The size of Charlottesville could stay the same and other cities be smaller, or other cities stay the same and Charlottesville be larger, etc. It should be clear how to arrive at very many ways (subways) for C to be true. And it should be clear that none of those ways are possible worlds. For each one of those ways C might have been true might itself be true in very many ways, varying names of inhabitants and adding details, etc.

3. A main attempt to make the idea of a possible world palatable is based on appeal to the idea of a maximal consistent set of propositions, where a set of propositions is maximal consistent if it is consistent (no contradiction can be derived from premises drawn from the set) and is such that adding to the set any proposition

[1] David Lewis *Counterfactuals* (Cambridge: Harvard University Press, 1973), 84.

which is not a member results in a set which is inconsistent. The contingency of C would then be explained in terms of there being maximal consistent sets of propositions containing C and also such sets containing not-C. One purpose here is to argue that this position is unsound.

In the lower predicate calculus (LPC) there are two main meanings for saying that a set of sentences is consistent. One (syntactic) is that there does not exist a proof using premises from the set, of a formal contradiction, say P&~P. This existence, though actual, would be highly idealized, not restricted to physical proof lists on paper. The other (semantic) is that there does exist an interpretation which makes every sentence in the set true. These can be proved to be coextensive. To say that a maximal consistent set of interpreted LPC sentences represents a "possible world" could only be a stipulative introduction of a technical term. Such a set of sentences is not essentially connected with any interpretation. So it is not a story of any "world." However, a set of LPC sentences does constrain interpretations in a way which could be called "describing a world structure." Thus, even if the terminology is not helpful, it is a reasonably clear way of speaking.

For a given consistent LPC sentence S, a maximal consistent extension of S is determined, albeit non-constructively, by an enumeration (the enumerations are constructive) of all LPC sentences. Then a series runs through the enumeration and includes or excludes a sentence depending on whether or not it is formally consistent with the set so far formed. There will be infinitely many different maximal consistent extensions for S, depending on which of infinitely many enumerations of the other sentences is chosen. We can call the idea that a proposition (such as C) is possible iff it has a good translation into LPC which has a maximal consistent extension, "LPC possibility."

We concede (solely) for purpose of discussion that this is a sensible way of speaking of something, which could be called "possible worlds." It could be a non-circular way of defining "possible world" in terms of syntactic consistency, and such consistency can be proven to be equivalent (in LPC) to semantic consistency. However, such a "possible world" would not be any specific interpretation and would not represent anything plausibly called a "story." "Once upon a time, Fa…" is a fair caricature.

It is large understatement to say this LPC possibility is no better as an account of the possibility of C than the vague "imaginative exercise" account. Translating C into LPC is a vague project, starting with a sentence letter translation (say 'C' as an SC sentence). Translation into a formal system is a wayof bringing out logical connections relative to the system. There are different translations and different formal systems. The idea of bringing out "all" connections is implausible.

I concede, again for purpose of argument, that there is a set of all propositions P such that C strictly implies P. Our access to that set is just as feeble as our imagination for testing the possibility of C.

The inadequacy of LPC possibility is underscored by what is often counted as a virtue of LPC – its compactness. If every finite subset of a set of LPC sentences is consistent (syntactically or semantically) then the set is consistent. This compactness is not a feature of sets of propositions. The infinite set (S-minus) : {1. There is at least 1 F, 2. There are at least 2 Fs,…,n. There are at least n Fs,…}is consistent for some appropriate choice of being F, such as being a rabbit. The set S: {S-minus, + 0. There are only finitely many Fs} is inconsistent (called "omega-inconsistent") even though every finite subset is consistent. That LPC does not allow this to happen is not a virtue as far as representing the possibility of propositions is the project. If an LPC possible set can include S, that is too bad for LPC possibility. It is better to say LPC cannot represent such a set, but that is too bad for LPC representation of possibility.

Second Order Logic (SOL) is not compact, but this is not improvement. It just counts as syntactically consistent some sets which do not have models, which is no virtue. The right verdict is that a set with no model is not "really consistent." This is an important notion which is not well captured by formal systems. This counts against trying to define possibility in terms of a maximal consistent set of propositions. For the consistency cannot be adequately based on the nonexistence of a formal deduction of a contradiction. A formal deduction can only use a finite number of premises. The absence of a formal deduction of contradiction is thus no guarantee of genuine consistency for sets of propositions. We would have to rely instead on the semantic idea of it being possible for all the propositions in the set to be true. That is obviously circular as an account of possibility.

4. This by itself is not sufficient to refute the view that there are a vast number of "possible worlds." C, as a way things might have been, might itself have been the case in many ways, and each of these many ways might have been the case in further ways. This might seem to lead to an infinite branching series with some hope for a possible world as a path through the lattice. Someone might even call such a path a "way" in another sense. The path, or world, or maximal consistent set of propositions, would presuppose the idea of possibility but would work out the idea that C being true would require infinite changes as to the totality of truths. Infinitely many infinite sets of such changes would then be the infinity of distinct "possible worlds" in which it is the case that C. This ontological view might be

advocated even after the hope of noncircular analysis is surrendered. Whether members of this set are really "worlds" would be an idle matter of terminology.

I will not attempt to refute the idea that C, as a way things might have been, heads a lattice of ways such that a path through the lattice could be called a "complete way" or "possible world" in which C is true. My complaint is just that, even so, saying truly that there are such things is not only no account of possibility – it would merely reflect a fact about possibility.

5. For all its inadequacy, the LPC notion of a maximal consistent set of sentences may have influenced the idea of a possible world as a maximal consistent set of propositions. Consider the set of propositions. As a totality, it has its detractors, who see it involved in paradoxes and deny the existence of such a set. Here we will set those objections aside. The paradoxes about the set of all propositions are answerable. There really is such a set. We will not burden our present argument with establishing that point. Rather, we observe that, if the set of propositions were really an inconsistent idea (the set itself is of course, highly inconsistent, but that is another matter)then so is the idea of the set of all true propositions. If there is no set of all propositions then there is no set of all true propositions. (The primary trouble for the all-set are self reference and excessive size. Those same troubles arise for the true-set.) And if there is no set of all true propositions then it is not reasonable to hold that there are nonetheless many distinct maximal consistent sets of propositions. We are assuming the existence of a set of all (and only) true propositions.

The set T of true propositions is a subset of the set of propositions, separated by the property of being true. This is a paradigm of a maximal consistent set of propositions. There are other large size sets of propositions, such as the set of false propositions or the set of propositions which are consistent with C, but those are not consistent. There is the set of necessarily true propositions, but it is a proper subset of T even if it is of the same cardinality. But, granting T, why not lots of other maximal consistent sets? The answer depends on a point which is, unfortunately, not unanimously accepted: any set is nothing but the extension of one or more properties. For a set to exist is for there to exist a property of which that set is the extension. Russell's Paradox has inspired the idea that some properties, such as the property of being a non-self-membered set, do not determine sets, on pain of contradiction. I agree with Bertrand Russell---there is no such property.

T is the extension of the property of being a true proposition. What property could determine a distinct maximal consistent set containing C? For any

contingently true proposition P, the property Pv, of being either a true proposition or the proposition P, determines the set T of all true propositions. Pv is a distinct property from the property of being true. It is possible for a proposition to have the property Pv without having the property of being true, since P has that property whether or not it is true. Thus there are infinitely many properties distinct from being true that will separate a maximal consistent set from the set of all propositions.

However, all of these trivially distinct properties determine the same set: T. There does not exist any property which will separate, from the set of propositions, any maximal consistent set containing C. There is only one maximal consistent set: T, and thus (if we must use the expression) only one possible world. There are nonetheless countless alternative possibilities. For every contingent truth, its negation represents a way things might have been. And "possible world" could still be used here, just in a properly loose way, according to which novels and plays may depict possible worlds. These will not be the infinitely detailed "possible worlds" of the modal metaphysicians.

The clear notion of a maximal consistent set of LPC sentences may have encouraged the idea of a maximal consistent set of propositions. The former is based on the clear construction of an enumeration of LPC sentences. This enumeration can be run through, not constructively, but plausibly, from the standpoint of classical logic, adding at each step in the enumeration if and only if doing so is consistent with what is formed so far. This is not possible for the set of propositions, since it is not denumerable, having a cardinality as great as the Continuum. However, if we accept Zermelo's Well Ordering Axiom, we might think of proceeding through this set on a transfinite selection of a maximal consistent extension of C, to obtain a maximal consistent set containing C, to be called a possible world in which C is true.

Zermelo's Well Ordering Axiom, and its equivalent, the Axiom of Choice, are propositions there is good reason to reject. And the background assumptions of Zermelo's Axiom include the presumption that there are no sets as big as the set of propositions. That set is a good candidate for a counterexample to the axiom – a set that cannot be well-ordered. However, this will not be argued here. Those who choose to believe there are such well-orderings can go on to believe in transfinite runs through to separate a maximal consistent subset other than T. We should add possible worlds metaphysics to the Banach-Tarski Paradox as just another wondrous result of these axioms. We should note at least that these run-throughs are not based on consistency defined as the nonexistence of a formal deduction of contradiction. They would appeal to the unanalysed idea of possible truth. This

appeal to well-ordering is no improvement on our idea of possibility as revealed to imaginative survey. Imagination is indeed a highly fallible guide, especially for some imaginers. It is an illusion to think these questionable postulates constitute any improvement.

6. In debates about whether the alethic modalities iterate, a relation, assumed to be reflexive, "accessibility" between possible worlds is introduced. "Possible" then becomes "true in some accessible possible world" and "necessary" is "true in all accessible possible worlds." The question whether what is necessarily true is necessarily necessarily true (Np-->NNp) then turns on the question whether the accessibility relation is transitive. The question whether what is possible is necessarily possible (Lp-->NLp) turns on the further question whether accessibility is symmetric. This is a very clever and creative way of recasting the questions about iterated modalities. As a tool for constructing meta-theoretical results about formal systems of modal logic it is unimpeachable. Various relations of accessibility could be constructed, one modeling one set of axioms, another modeling some other.

However, this modeling of different axioms would have little to do with whether one system represented the truth of the matter. The constructed "worlds" with their relations of accessibility could be adequately represented by sets of numbers. The question as to which axioms represent the facts of modality would depend entirely on the separate question what those facts are. On the other hand, if possibility really consisted in the existence of "worlds" in some more substantial sense, that were really accessible, then the question about iterated modality would be a question about these entities and their relations. It might then be tempting to think that the objective existence of possible worlds is required for the objectivity of the questions about iteration. This does not follow from the fact that the objective existence of possible worlds and an accessibility relation would make the questions objective. But the hope for objectivity about questions of iteration does create a pressure in favor of objective existence for possible worlds.

The pressure is even greater in the case of analyses of subjunctive conditionals in terms of possible worlds. Consider the sentence (S) "If it were the case that Virginia Beach is not (at present) the largest city in Virginia, then it would be the case that Norfolk is the largest." It is common to pretend that a sentence such as the antecedent of S expresses a proposition, which could be named 'not-V' and that the consequent also expresses a proposition (N). Many philosophers would want to conclude that then S expresses a definite proposition and would enquire as to whether it is true, false, or neither. This is quite

implausible. We will grant for purpose of discussion that the antecedent and consequent sentences do indeed, to parties to this discussion, express definite propositions. Various truth functional relations between these propositions can be claimed without need for further explanation. Thus ~V materially implies N, and conversely, their disjunction is false, etc. But the frame of S, though it is grammatically just as regular as the grammatical forms for truth functional connections, (in fact, is even more natural in English than they are), nonetheless does not represent any relation between ~V and N.

It can represent various relations if it is supplemented with sufficient further explanation. Thus someone might say that they assert S because, if we hold all the facts about the size of Virginia cities other than Virginia Beach constant, and add to a precise statement of those facts, the premise ~V, then N can be straightforwardly deduced. This could be formulated precisely so as to yield a rather trivial truth. That ~V, accompanied by this other material, entails N, is a relation between ~V and N which can be asserted to hold. That proposition is not expressed by S in the way that the antecedent and consequent express propositions.

Another person might assert "Not-V would not have implied N" on the grounds that Virginia Beach was almost not the largest city in Virginia because an industry was started in Fairfax which was very likely to succeed and would have attracted half the population of Virginia Beach, etc. The line would be that the best explanation for ~V would be the success of that venture and the result that Fairfax, not Virginia Beach, was largest. The contribution of the frame is not the expression of one standard relation.

This verdict will be unsatisfactory to many philosophers because they will feel that the two claimants just described would be in disagreement. The denier might say that the trivial deduction, though valid, is an objectively wrong way to interpret S. Here possible worlds have seemed to have promise. The idea is that for propositions A and C, "If A were the case then C would be the case" means, roughly, that with respect to the class of possible worlds in which A is true there is at least one in which C is true, which is "closer" to the actual world than is any world in which C is false. There can be considerable further discussion aimed at clarifying this relation of closeness. Here again we see the hope of objectivity.

The truth is that S does not objectively express a proposition at the level where its antecedent and consequent do. We have to go beyond that. The S affirmer above provides supplementary assumptions allowing a deduction of the consequent. The S denier asserts an incompatible counterfactual, not explicitly listing the needed premises. The denier claims there are laws of social science which, along with assumptions about probable developments and the assumption

that not-V, entail F. These are not the only ways of supplementing, but supplementation is needed. The "closest world" analysis does not duly recognize this.

Of course, the objectivity view about S does not depend on the possible worlds analysis. There could be an account according to which asserting S is claiming that there exists support of the "right kind" such that added to the antecedent it entails the consequent. A possible worlds analyst of subjunction can reject the view that there is in general a right kind of support and can appeal to the idea that support has to be a finite set of propositions to add to the antecedent, while the "really right kind" of support would require a specification of details that would go on *ad indefinitum*. The objective truth would be determined by the metaphysical closeness relation between possible worlds. Thus it will seem that, just as reference to the real world is required to give our claims about what did cause what the hope of objective truth, reference to possible worlds is needed to give such objectivity to our claims about what would have caused what. This value of a possible worlds analysis for objectivity depends on the objective existence of possible worlds.

7. The fact that there are many logically distinct properties for determining T does not mean that there is any way of determining some maximally consistent set of propositions other than T. We might feel inclined to say "If the possibility that ~V were realized, then T would not be the set of true propositions, because V would still belong to T and would be false." This is not correct. "The set of true propositions would *in that case* not be the set of true propositions." What case would that be? "The case in which V is false"? There is no such case. Its being possible that V is false does not consist in there being a "case" or "state of affairs" or "world" "in which" V is false.

This is not to deny that the term "proposition" has sometimes been used with a different meaning from ours, according to which "propositions" so called could change truth values. They might be sentences of some kind, for example, or "tensed propositions." But properly chronologized propositions do not and cannot change truth value. If this is recognized by a philosopher who thinks of being necessary as being true at all times, the result will be the view that every true proposition is necessarily true. This is Spinoza's view. (However, it is worth noting that Spinoza holds that some true propositions are such that their being false is possible. That is, he has a use of "possible" which is compatible with being necessarily false. It is a kind of "epistemic" possibility.) Leibniz may have been led to the idea of possible worlds to avoid this necessitarianism. This attitude, that

possible worlds of some sort are needed to avoid Spinozism, is reflected in the saying that "those philosophers who go about saying loudly and defiantly, 'There's only one possible world, the actual one' are either Spinozists or fools."[2] The right answer to the threat of Spinozism is that the possibility of a proposition does not consist in its being possible for it to become true. So there is no need, having noted that this could not happen in this world, to find some other world for it to happen in, or to have already happened in.

8. Of course, if it were false that V, then lots of things would be different. Every proposition that entails V (and there are infinitely many distinct such propositions) would be false. This specification of what would be true could be extended further by adding all the propositions deducible from ~V in conjunction with the laws of a genuine science, except for the vagueness of "genuine science" and the paucity of such science concerning such topics as the relative size of cities. We should not put up with just any true universal generalizations to determine what would be true if ~V. We would need true "laws of nature." But the distinction between laws of nature and mere accidental generalizations is vague and subjective. Even when it is objectively true that a candidate has been advanced with good reason and defended reasonably this does not mean that it has an objective property of being truly "natural."

When Q follows from P either as a logical consequence or an analytic consequence or by a law of nature, the conditional "If P then Q" may be called "independent." By contrast, "If Smith is elected, I will be pleased" claims a connection which depends on facts additional to what is stated in the antecedent or available just from general scientific knowledge. There is no objective division between "independent" and "dependent" conditionals. While ever so many things which are not true would be true if ~V, there is no such thing as "what would be true if ~V." Every P for which there is an "independent" conditional "If ~V then P" would be true if ~V were true. But this does not determine a set. A "dependent" conditional is aptly named. To the question whether such a conditional is true or false, it is appropriate to reply, "It depends – on what, if anything, you mean."

[2] Peter van Inwagen, "Two Concepts of Possible Worlds" reprinted in *Identity, Ontology, and Modality: Essays in Metaphysics* (Cambridge: Cambridge University Press, 2001), 206-42. Page 232 note 19. Note that, while I reject Spinozism, I only say loudly and defiantly that there is just one maximal consistent set of propositions, refraining from calling it the "actual world." However, even if I do happen to instantiate the above generalization, this is at best an accidental truth, not a lawlike one.

It is true of the proposition (A) that 2+2=4, that it would be true if ~V. This is simply because (A) is necessarily true and so would be true no matter what. If we say that the proposition (L) that Lynchburg is the largest city in Virginia is such that it would be true if ~V, this saying lacks a clear meaning. It is clear to say that (L) would be true if ~V because (L) is necessarily true. Clear, because clearly necessarily false, since (L) is contingently false. "P would be true if Q were" is understandable if it is obvious that P is independent of Q. It is understandable along a different line if there is a law that Q would guarantee P that is known to the hearers. But "what would be true if Q were true" is not clear for general Q, just by the grammatical construction.

It can be good to hope that our dreams will come true and in the same spirit we can speak of realizing our possibilities. Such possibilities may be represented, that is, described, by contingent propositions, as when we say that it is possible that Mary will become president. But the proposition that Mary will become president is either true now or false now and if true was true at the time of the Pharaohs. (This does not represent any restriction on Mary's possibilities, such as undercutting her power to decide whether or not to run. That it is true now that I will eat dinner tonight, God willing, does not affect my power to decide to skip dinner.) Propositions in the proper sense of the term do not and cannot become true and the idea that a contingent proposition represents a person's power of acting is a bad mistake. The possibilities which represent a person's powers of action are not the possibilities in which the contingency of propositions consists. It is a necessary condition of its being in my power whether to skip dinner tonight that the proposition that I do so should be contingent. But that contingency is no significant basis for the power. It is just as contingent that I run a sub-four minute mile tonight, but that is quite definitely not in my power.

Alternate descriptions for T, such as TV, the set of propositions which are either true or identical to V,work by specifying part of the set. The number of specifications could be infinite, for example, using "1 is a number" "2 is a number" etc., as V was used in TV. We would only get the rest of T with something lame, such as "the rest of the true propositions." Such a description as "propositions about Virginia Beach" would not do, since every proposition P is equivalent to "Everything is such that P." Every proposition is about everything, so aboutness is not a good basis for individuating propositions.

9. Some hold that the notions of propositions and properties and absolute truth are risky and deserve suspicion. Others of us disagree; but we should all agree that they are incomparably clearer than "possible worlds." There is nothing inherently

overextended in thinking about the proposition V. It is reasonable to think we can understand it perfectly well. This is not to ignore the point that the topic of city size can be problematic. But it is not wildly overconfident to expect that our social science is capable of clarity on such a topic. To reject the idea that we can get clear about some propositions and properties would be extremely pessimistic. The infinity of propositions suggested to us by the limitless applicability of logical operations does indeed warrant more modesty. But this trust in the power of logic only leads (those willing to go that far) to the one maximal consistent set T.

Having gone that far, the notion of possible worlds is still far short of the properly required clarification. It has been suggested that propositions can be defined in terms of possible worlds. The proposition that Virginia Beach is the largest city in Virginia would be the class of possible worlds in which it is the largest city in Virginia (or perhaps its counterpart is the largest city in the counterpart of Virginia). This definition avoids the blatant circularity of identifying the proposition with the class of worlds in which the proposition is true, only by economical wording. We can presume that those who offer such a definition would not also endorse the view of a possible world as a maximal consistent set of propositions.

It has been said that the idea of necessary truth can be defined in terms of counterfactuals. The claim that p is necessarily true would mean "If ~p were the case then A&~A would be the case" for any arbitrarily chosen A. On the contrary, such a "saying" as "If 7 were not a prime number, then V&~V would be the case" does not make good sense. It can be subjected to arbitrary interpretation designed to fit it into a general theory about possible worlds and counterfactuals. In such a theory a convention establishing such an inter-theoretical connection between necessity and counterfactuals could be understandable. That should not be confused with establishing the alleged connection between necessity and counterfactuals.

This is not to deny that from the hypothesis that 7 is not prime we can construct a valid proof that 7 both is and is not divisible only by itself and 1. And from that explicit contradiction we can derive any arbitrary proposition Q, including any explicit contradiction. It is not obvious that such a proof is available starting from the denial of just any necessary proposition. But in any case, the existence of such a derivation is not the same thing as a proof of the alleged counterfactual. A *reductio* proof does not in general show what would have been the case had the hypothesis been true. It merely shows that the hypothesis could not have been true. We might naturally say that the *reductio* shows what would have had to have been the case in order for the hypothesis to be true. This is still

not to establish a counterfactual. Showing what would have had to have been the case is not showing what would have been the case. (You may be tempted to reply that if it would have had to be the case that P then it would have been the case that P had to be the case. Please reconsider. That alleged consequent is nonsense.) Natural tolerance and charity will lead us to understand "If ~p were the case then A&~A would be the case" as a summary or a promise of a *reductio*, but it would be an abuse of charity to take that as establishing the acceptability of the above definition.

10. I have always objected to possible worlds with a simple argument. If it is true that some proposition p is true in some possible world w, then it is necessarily true that p is true in w, since having p true is part of what it is to be w as opposed to some other world. [No doubt a formal treatment can be worked out which allows that T_wp("p is true in world w") does not entail LT_wp (using 'L' for necessity). We might even have T_wp&~$T_w T_w$p and other dizzying exotica. Still, it is reasonable to reject that and accept T_wp-->LT_wp.] Similarly, innocent reasoners naturally agree that T_aV ("It is true in the actual world that V") is equivalent to V. That is, to say that it is true in the actual world that V is equivalent to saying that V. Conjoining these reasonable concessions leads to the conclusion that it is necessarily true that V. Thus it seems that our natural understanding of possibility is not well served by the notion of possible worlds.

However, when I cited this problem to a distinguished philosopher, it was dismissed. In this exchange, we were, unfortunately laboring under the impression that the truth was not V but N ("Norfolk is the largest city in Virginia"). Hopefully we can work around this factual error here. We only need be clear what is being assumed as to the actual facts about city sizes, rather than worrying about what exactly are those facts. Anyway, the reply was as follows.

"There is no problem. The proposition 'Norfolk is the largest city in Virginia' is necessarily an element of any set of which it is a member. If X is the set of all true propositions (formulable from a given set of atomic constituents), then 'Norfolk is the largest city in Virginia' is an element of X. It is also necessarily an element of X. But it is not necessary that X is the set of true propositions. Some of the sets other than X that might have been the set of true propositions contain the proposition 'Norfolk is not the largest city in Virginia.' Hence it is not necessary that the proposition 'Norfolk is the largest city in Virginia' is an element of the set of true propositions. That proposition about the size of Norfolk retains its

contingency, consistently with the membership relation between elements and sets being necessary in any case in which it holds."[3]

There remains this problem. There is no such thing as "the sets other than X that might have been the set of true propositions." (Here I am assuming that X=T and that it does not in fact contain N, but rather, V. This modification makes no difference to the point at issue.) It is true that it might have been the case that some set other than X was the set of true propositions. But that does not mean that there is some set other than X, say X', such that X' might have been the set of true propositions. That there might have been a unicorn does not mean that some unicorn might have been.

If we were to focus on sets of propositions "formulable from a given set of atomic constituents" we might be working within LPC, where the notion of atomic constituents is meaningful. Then there would indeed be lots of maximal consistent sets of such propositions. But then it would be a misleading understatement to say that "it is not necessary that X is the set of all true propositions." It would be necessary that such an X is not the set of all true propositions, since it would be denumerable. It is not being questioned here that V could be false or that if V were false then some set other than T would qualify as the set of all true propositions. It is just that there does not exist any set other than T which could qualify as the set of all true propositions. That a maximal consistent set would come into existence with the falsity of V does not mean there exists a maximal consistent set which would come into existence. Thus the possibility of ~V cannot possibly consist in the existence of a maximal consistent set of propositions to which it belongs.

Furthermore, it is either false or badly misleading to say that V is necessarily a member of T(=X), the set of all true propositions. As was noted earlier, V is necessarily a member of the set TV (propositions such that they are either true or =V) and TV is contingently identical to T. V is necessarily identical to V but not necessarily true. Belonging to the set of true propositions is the same property as being true and so belonging necessarily would be being necessary, which V is not. Belonging to TV is not the same property as being true. It is a distinct property which happens as a matter of fact to determine the same set.

It must be conceded that the complaint that the notion of truth in a possible world leads to bogus necessities does not apply clearly to possible worlds viewed as sets of propositions. This is because saying that a proposition p belongs to a set S is not saying that p is true in S or even offering any means of making sense of "p is true in S." No doubt the proposition that V belongs to T (that is, the proposition

[3] Christopher Peacocke "The Past, Necessity, Externalism and Entitlement," *Philosophical Books* 42, 2 (2001): 106-17, 109.

that V belongs to the set of all true propositions) belongs to T, and the proposition that the proposition that V belongs to T, belongs to T, and so on and on; except that those are all the same proposition, namely, V. So there is no good sense in which "they" all belong to T or to any set of propositions, since there would not be separate members. Talk of worlds "in which" it is true that there are unicorns is thus not naturally translatable into the terminology of sets of propositions. Thus my old favorite possible worlds problem is somewhat inapt for the discussion in terms of sets of propositions. What is relevant here is the question as to whether my problem can justifiably be dismissed for the reasons just considered. I have argued those are not correct reasons.

Imagination, though its frailty must be duly noted, is the proper way to understand possibility. Properly understood, it shows us that possibility is not the possibility of becoming true or of being true contrary to what appears to us. The right way to explain to the students that it is logically possible that the University was never built does not require the point that for all they know, it really never was built, the whole appearance being managed by an evil genius. Rather, they should be led through imagining the area of the University during the relevant historical period, with no University being constructed during that time, right up to and including the present. That is how to show that the proposition that the world has never had the University has the property of possible truth, if we are to have any hope of showing such a thing. To say there is a world in which that proposition is true, or a maximal consistent set of propositions including it just encumbers the picture with false presumptions.

Spinozism is based on distrust of imagination as a guide to any genuine feature of propositions. It is ironical that Spinoza's low estimate of imagination is well supported in Descartes' writing and yet Descartes bases one of his principal arguments on what is clearly an appeal to imagination. He says that he can rationally feign that his body does not exist, but cannot do that with respect to his own existence. This comes to being able to imagine that he exists without a body but unable to imagine his body exists without a body. Spinoza's response is to dismiss such flights of imagination as worthless as indications of what is true. The imaginer has no good sense of what would be required to bring about the conditions he thinks of himself as "imagining."

This suggests the requirement that one should consider how the actual facts would have to be changed in order to bring about the alleged possible condition. This process of making adjustments to the facts might seem easy with respect to merely chance happenings such as coin tosses or quantum phenomena. But for many possibility candidates the changes seem to ramify endlessly, suggesting the

need for a new world. This leads into the idea of alternate possible worlds. It is a false lead. We do not have to accept Spinoza's low estimate of imagination. We should recognize it is the best guide we have to possibility. We can duly recognize its limitations without giving up claim to grasp (somewhat) a genuine property or sliding into claims to grasp a merely somewhat genuine property.

The target here has been the essentially meinongian realism of the ontology of possible worlds. Those who work in terms of possible worlds are likely to resist description as meinongian. It has been thought especially that this is mitigated by the switch to maximally consistent sets of propositions. That is not so. But this is nothing at all against the platonic reality of propositions and the eternal forms. Confusing platonism with meinongianism has been a frequent source of misunderstanding. We have been arguing against the latter, not for or against the former, but hopefully, combating the confusion can be a help to the larger cause.

PRAGMATIC ENCROACHMENT AND CONTEXT EXTERNALISM

David COSS

ABSTRACT: Pragmatic Encroachment (PE hereafter), sometimes called 'anti-intellectualism,' is a denial of epistemic purism. Purism is the view that only traditional, truth-relevant, epistemic factors determine whether a true belief is an instance of knowledge. According to anti-intellectualists, two subjects S and S*, could be in the same epistemic position with regards to puristic epistemic factors, but S might know that p while S* doesn't if less is at stake for S than for S*. Motivations for rejecting purism take two forms: case-based and principle-based arguments. In considering both approaches, I argue that PE is best viewed as externalist about epistemic contexts. That is to say, I claim that what determines a subject's epistemic context is external to her mind.

KEYWORDS: pragmatic encroachment, anti-intellectualism, purism, contextualism

1. Pragmatic Encroachment

Pragmatic Encroachment (PE) is a rejection of epistemic purism. According to purism, only truth-related factors determine whether a belief is an instance of knowledge. According to PE,[1] when the stakes rise, the evidential threshold for knowledge rises proportionally, thus making it more difficult for a true belief to be an instance of knowledge.

A consequence of PE is that two subjects, S and S*, could be in the exact same position regarding puristic epistemic constraints (such that both have a non-gettierized justified true belief), but S might know that p, while S* might not, if more is at stake for S* than S. Put more formally, anti-intellectualists typically endorse variants of the following principle.

Practical Interest Condition

The strength of S's epistemic position with regards to p matches the practical costs S is expected to incur if p were false.[2]

[1] The view is often also called 'anti-intellectualism.' I will use both terms 'pragmatic encroachment' and 'anti-intellectualism' interchangeably throughout the paper.

[2] There are many ways of formulating and describing the practical interest condition. Similar formulations can be found in Jason Stanley, *Knowledge and Practical Interests* (New York: Oxford University Press, 2007) as well as Ram Neta, "Anti-Intellectualism and the Knowledge-

Both S and S* can be in different knowledge states, even if both subjects satisfy all the classical knowledge constraints. Anti-intellectualists explain the discrepancy in knowledge states by linking the presence of stakes to an elevation in the evidential threshold for knowledge. In other words, even though all the classical conditions of knowledge may remain fixed for both S and S*, S could know that p, even if S* doesn't, due to the expected cost to S* if she is mistaken.

2. Contextualism

Contextualism is the view that the meaning of the word 'knows' is context sensitive. More specifically, contextualists argue that the truth of knowledge attributions shift with the relevant contextual standardsin play. For example, contextualists maintain that when one entertains skeptical hypotheses—or even alternate possibilities—the epistemic threshold for knowledge shifts upward, making it more difficult for attributors to know the proposition in question. However, in ordinary contexts—those that obtain outside of philosophical study, discussion and reflection—the standards of knowledge are usually lower. In other words, S's utterance "I know that I have hands" could be true at time t (assuming, of course, S isn't a handless BIV), while false at t+1 depending on S's context. At time t, S might be unaware of skeptical hypotheses, but perhaps at t+1, S is absconded by an epistemological skeptic. In such a situation, S's utterance of "I know that I have hands" is true at t, but false at t+1.[3]

In this way, contextualists deny knowledge invariantism, the view that there's only one standard of knowledge. Rather, there are multiple standards of

---

Action Principle," *Philosophy and Phenomenological Research* LXXV, 1 (2007): 171-195. On the other hand, Fantl and McGrath prefer to understand the condition in terms of what's rational to act on. For their various formulations, see Jeremy Fantl and Matthew McGrath, "Evidence, Pragmatics, and Justification," *Philosophical Review* 111, 1 (2002): 67-94; Jeremy Fantl and Matthew McGrath, "On Pragmatic Encroachment in Epistemology," *Philosophy and Phenomenological Research* 75, 3 (2007): 558-589; Jeremy Fantl and Matthew McGrath, *Knowledge in an Uncertain World* (New York: Oxford University Press, 2009); and Jeremy Fantl and Matthew McGrath, "Pragmatic Encroachment: It's Not Just About Knowledge," *Episteme* 9, 1 (2012): 27-42.

[3] While this is a plausible, and arguably a good first approximation, for describing the basic idea behind Contextualism's response to skepticism, the view is more nuanced. DeRose outlines ways non-skeptics can block the standards of knowledge from rising using various conversational maneuvers. See Keith DeRose, *The Case for Contextualism* (New York: Oxford University Press, 2009). While it's common to think that contextualists capitulate to the skeptic when such hypotheses are entertained, Contextualism has resources for salvaging ordinary knowledge attributions even when skeptical scenarios are made salient. See Ch. 4 of DeRose's *The Case for Contextualism* for more on such maneuvers.

knowledge determined by the conversational context. Contextualists typically adhere to the following thesis about knowledge.

The Contextualist Thesis

Whether a knowledge attribution, 'S knows that p,' made by an attributor A, is true or false, depends upon whether A's evidence (or, if one prefers, strength of epistemic position) is strong enough for knowledge relative to standards of knowledge in A's context.

3. Case-Based and Principle-Based Arguments for PE

Anti-intellectualists defend the practical interest constraint in several ways, such as the reliance on knowledge-action (or knowledge-rationality) principles and contrast cases (such as DeRose's bank cases).I would first like to draw attention to the kind of cases anti-intellectualists employ in making the case that anti-intellectualism is best viewed in light of context externalism before briefly discussing principle-based approaches.[4]

Anti-intellectualists employ contrast cases to illicit the intuition that absent high stakes, subjects would know the proposition in question.[5] Traditionally, the first contrast case is low stakes, while the second is high. Consider the following example.

Lazy Fred

It's Saturday evening and Fred realizes he's running low on vitamins and that if he doesn't get them over the weekend, he won't have enough to sustain him until

[4] While case-based arguments and principle-based arguments are the two dominant strategies used to make the case for PE, it's unclear which is the better approach, and as some have argued, neither succeed. Roeber for example, argues that both fail to sufficiently motivate PE. I remain neutral on which approach is better (or successful). See Blake Roeber, "The Pragmatic Encroachment Debate," *Nous* 2 (2018): 171-195. Even anti-intellectualists who prefer the principle-based approach employ contrast cases, and explain the intuition behind them in virtue of various principles of rational action. For simplicity and ease of exposition, I will primarily discuss contrast cases and how the PE interpretation of them supports a context externalist position.

[5] It's worth pointing out other options for purists to accommodate the anti-intellectualist intuition. For example, Foley would account for the intuition in these cases by appealing to the notion of an epistemically responsible belief. In other words, contrary to what anti-intellectualists suggest, subjects in both high and low know the respective proposition, but in the high case variant, the subject is acting in an epistemically irresponsible way. See Richard Foley "Epistemically Rational Belief and Responsible Belief," in *Proceedings of the Twentieth World Congress of Philosophy*, ed. Richard Cobb-Stevens (Bowling Green: Philosophy Documentation Center, 2000), 181-188.

> Monday. He remembers the pharmacy being open on Sunday, having been there just two weeks ago, and decides he can hold off until Sunday morning. Fred thinks to himself: *I was there just two weeks ago, and pharmacies rarely change their hours. I know the pharmacy will be open tomorrow. I'll get my vitamins then*.
>
> **Diabetic Sally**
>
> Sally has diabetes and needs insulin to survive. It's Saturday evening and she realizes at 7pm that her insulin supplies are too low to maintain her through Sunday. If she runs out of insulin, she runs the risk of serious injury, and possibly death. She remembers from two weeks ago that her pharmacy was open on Sunday. She doesn't feel like braving the weather to get an insulin refill and thinks to herself: *I know the pharmacy will be open tomorrow. I was there just two weeks ago. I'll get my supplies in the morning*.

In the first case, the intuition is supposed to be that since little is at stake for Fred, memorial evidence suffices for him to know the pharmacy will be open. In the second case, however, since more is at stake if Sally is mistaken, the evidential threshold rises, and memorial evidence alone is insufficient for her to meet the more demanding standards of knowledge. Thus, while Sally and Fred are in identical epistemic positions regarding evidence, Fred's is sufficient to meet or surpass the lower epistemic threshold for knowledge, while Sally's isn't.

Now consider DeRose's classic bank cases in which the evidential threshold is elevated in virtue of potential financial loss.

> **Bank Case A.** My wife and I are driving home on a Friday afternoon. We plan to stop at the bank on the way home to deposit our paychecks. But as we drive past the bank, we notice that the lines inside are very long, as they often are on Friday afternoon. Although we generally like to deposit our paychecks as soon as possible it is not especially important in this case that they be deposited right away, so I suggest we drive straight home and deposit our paychecks on Saturday morning. My wife says 'Maybe the bank won't be open tomorrow. Lots of banks are closed on Saturdays.' I reply, 'No, I know it will be open. I was just there two weeks ago on Saturday. It's open until noon.'
>
> **Bank Case B.** My wife and I are driving home on a Friday afternoon, as in Case A, and notice the long lines. I again suggest we deposit our paychecks on Saturday morning, explaining that I was at the bank on Saturday morning only two weeks ago and discovered that it was open until noon. But in this case, we have just written a very large and very important check. If our paychecks are not deposited into our checking account before Monday morning, the important check we wrote will bounce, leaving us in a *very* bad situation. And, of course, the bank will not be open on Sunday. My wife reminds me of these facts. Then she says, 'Banks do change their hours. Do you know the bank will be open tomorrow?' Remaining as confident as I was before that the bank will be open then, still, I

> reply, 'well, no, I don't know. I'd better go in and make sure.'[6]

As we will see in the next section, anti-intellectualists give a divergent account of why in the second two cases subjects fail to know the relevant proposition.

Besides contrast cases, anti-intellectualists also employ knowledge-action or knowledge-rationality principles to motivate PE. For the sake of simplicity and ease of exposition, I will only present two common formulations. The purpose of such principles is to provide solutions to a myriad of worries purist epistemologies face, such as the epistemic threshold problem as well as explaining the intuitive knowledge discrepancy between low and high stakes cases.In dealing with the epistemic threshold problem, Fantl and McGrath write, "How probable must p be for you to know it? It must be probable enough to properly put to work as a basis for belief and action."[7] Further clarifying, they formulate the following principle:

> Your probability for p is knowledge-level iff the probability that non-p doesn't stand in the way of p's being put to work[8] as the basis for belief and action.[9]

As stated before, this principle, and ones similar to it, are used not only to explain the intuition behind contrast cases, but also as an independent reason for endorsing anti-intellectualism over purism. In Bank Case B, Keith doesn't know that the bank will be open on Saturday in virtue of the probability of its negation standing in the way of his acting on that proposition.[10]

---

[6] DeRose, *The Case for Contextualism*, 1-2.

[7] Fantl & McGrath, *Knowledge*, 65.

[8] Fantl & McGrath leave the notion of p's being put to work undefined. One could cash this out in many ways, and in Ch. 3 of *Knowledge* they present several principles which further employs the notion of 'standing in the way.' Arguably the easiest way to explicate this phrase, within the PE framework, is to see how it accounts for the bank-style cases employed. For example, consider DeRose's Bank Case B. While the probability that the bank will be open on Saturday is moderately high, the stakes stand in the way of him knowing this, thus preventing Keith from putting to work his belief that the bank will be open on Saturday.

[9] Fantl & McGrath, *Knowledge*, 65

[10] There are several formulations of knowledge-action principles. For the sake of simplicity and ease of exposition, I only present only two. For a discussion of these principles, and alternate formulations, see John Hawthorne and Jason Stanley, "Knowledge and Action," *Journal of Philosophy* 105, 10 (2008): 571-590; C.S. Sripada and Jason Stanley, "Empirical tests of interest-relative invariantism," *Episteme* 9, 1 (2012): 3-26; Fantl & McGrath "Evidence," and "Pragmatic Encroachment;" Jessica Brown, "Subject-Sensitive Invariantism and the Knowledge Norm for Practical Reasoning," *Nous* 42, 2 (2008): 167-189; Jessica Brown, "Knowledge and Assertion," *Philosophy and Phenomenological Research* LXXXI, 3 (2010): 549-566; Jessica Brown, "Practical Reasoning, Decision Theory and Anti-Intellectualism," *Episteme* 9, 1 (2012): 1-20; Mark

Arguably a simpler formulation such principles takes is that knowledge is a necessary—although not a sufficient condition—for acting on a proposition:

> For any subject S and any proposition P, S can only rationally act on P if S knows that P.[11]

This principle is able to explain the intuition that in high stakes cases, subjects either don't know the relevant proposition in question or else need stronger epistemic justification for their belief to count as an instance of knowledge. In Bank Case B, Keith can't put off going to the bank since he can't rationally act on the proposition "the bank will be open on Saturday," since he doesn't know the bank will be open.

The above formulation cast in terms of knowledge being a necessary condition on rational action is the one endorsed by anti-intellectualists, and therefore, other variants will not be considered at length.[12]

Schroeder, "Stakes, Withholding, and Pramatic Encroachment on Knowledge," *Philosophical Studies* 160, 2 (2012): 265-285; and Roeber, "The Pragmatic Encroachment Debate."

[11] While I will not entertain objections to this principle at length, it's minimally worth pointing out (i) its counterintuitiveness and (ii) an objection when employed in anti-intellectualist reasoning. Consider a case where S finds herself face to face with a tiger. Presumably the stakes rise to a point where even visual perception is insufficient for her evidence to meet the more demanding epistemic threshold. In such a case, S doesn't know the proposition: 'there is a tiger in front of me' and consequently cannot act on that proposition (such as fleeing, or if she has the proper weapons, fighting back). Another problem is that it seems that the more that is at stake, the more rational it is to act on the high stakes proposition. Considering the previous case, not only does it seem that S ought to act on the proposition, but it strikes me as obvious that she is more justified in doing so given such high stakes. In other words, one could plausibly invert the knowledge-action principle so that as the stakes rise, one is proportionally more justified in acting on it. Suppose that it wasn't well established that humans are a significant contributor to climate change. It seems that given the gravity of the situation, even given flimsy evidence that, all other things being equal, one has a good reason to act on that proposition. While these are important considerations, the purpose of outlining standard variants of knowledge-action principles is exegetical, not evaluative.

[12] Logical space affords at least two other formulations, a sufficient and a bi-conditional principle. However, such principles run into serious obstacles. One could easily imagine, for example, a case where knowing that p doesn't justify one in acting on p. Perhaps S knows that God doesn't exist, but that hardly justifies S is proclaiming such in a crowded church. For this reason, among others, interpreting knowledge-action principles in terms of sufficient conditions is unpopular among anti-intellectualists. However, Fantl and McGrath seem to suggest a pro tanto variant of the sufficient condition formulation. Given that it's unclear (at least to me) how this formulation works in service of PE ends, I will not explore it here.

4. Pragmatic Encroachment and Context Externalism

Jason Stanley argues that, for contextualists, salience of error, rather than an elevation in stakes, is what causes Keith's context to shift from a low to a high standards one. If, in Bank Case B, Keith's wife hadn't made him aware that banks sometimes change their hours, Keith would arguably remain in an epistemically less demanding position. In making the case that PE (or, as he calls it, 'Interest-Relative Invariantism') provides a superior explanation of contrast cases, Stanley argues that PE is able to explain the intuition behind traditional bank cases as well as others that contextualists struggle with. In the following case, Sarah and Hannah occupy a high stakes context, even though both are unaware of the expected consequences of being mistaken.

> **Ignorant High Stakes**
>
> Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. But neither Hannah nor Sarah is aware of the impending bill, nor the paucity of available funds. Looking at the lines, Hannah says to Sarah, 'I know the bank will be open tomorrow, since I was there just two weeks ago on Saturday morning. So we can deposit out checks tomorrow morning.[13]

Stanley claims that while PE can explain the upward shift in the standards of knowledge for Hannah and Sarah, contextualists struggle providing a satisfactory explanation since salience of error is absent.Stanley accounts for the difficulty contextualists face by ascribing to their position an intention-based notion of contextual shifts. As he writes, "On this standard account of context-sensitive expressions, their semantic contents, relative to a context, are determined by facts about the intentions of the speaker using that expression."[14] Since intentions play no, or at best a marginal role in Ignorant High Stakes, contextualistsstruggle explaining cases where subjects are ignorant of the practical cost of being wrong. It *seems* that contextualists are committed to saying that Hannah and Sarah are in a low standards context, which appears to be a mistaken account of Ignorant High Stakes.[15]

---

[13] Stanley, *Knowledge*, 5

[14] Stanley, *Knowledge*, 25

[15] One might object that Stanley has unfairly characterized contextualism as being too committed to salience of error raising the epistemic threshold for knowledge. For example, one might point out that what partially fixes a context, even for contextualists, are mind-independent factors. Consider the following attribution made by subject S, "Jack knows carrots are orange." Part of

My employment of Ignorant High Stakes is not meant to provide a reason for preferring PE over contextualism. Rather, the aim is to highlight a case in which contextual shifts move independently of a subject's awareness.

A more fundamental reason to think IRI-theorists are externalists about epistemic contexts is that epistemic shifts occur for subjects, rather than attributors. If an epistemic context shifts for a subject—whether she be the subject of a knowledge attribution or not—then it's plausible that standards of knowledge shift independently of awareness.

The claim that PE is best seen in light of context externalism might strike some as either obvious or else trivial and therefore not worthy of exposition. However, if anti-intellectualists are externalists about epistemic contexts, there are a range of cases they struggle explaining.

Problematic cases for anti-intellectualists are those in which subjects or attributors mistakenly think theyoccupy a high stakes context. The intended intuition is supposed to be that the epistemic threshold for knowledge rises for S, even though nothing is at stake if he's mistaken.While such cases are problematic for anti-intellectualists, arguably contextualists and purists alike are able to do so with ease.

Mistaken Bank Case

It's Friday evening and Keith is driving home from work. He passes the bank with his check in hand and notices long lines. Feeling tired he'd rather deposit his check on Saturday. However, he believes he has an impending bill due Monday, and if late, could lead to foreclosure on his house. Unbeknownst to him, his wife

---

what fixes the context for S is the fact that carrots are orange (and this fact is, plausibly, independent of Jack's awareness). I concede that many factors, both mental and non-mental, determine what context a subject or attributor is in. However, given what I have said about contextualism, external facts would fail to either raise or lower the epistemic threshold for knowledge independent of awareness. Perhaps one is comfortable developing an externalist account of contextualism, but this faces at least two problems. First, it robs the cases contextualists use to support their arguments for contextualism. For example, in DeRose's bank cases, external factors remain fixed across both situations, but the contextual content is different in B than in A. What causes the shift in Bank Case B is Keith's wife making him aware that banks sometimes change their hours. A more serious concern is externalist views of contextualism would fail to provide a solution to skepticism. Contextualism is largely motivated by its ability to account for how ordinary people have knowledge. Contextualists maintain that since ordinary people are unaware of skeptical scenarios, their epistemic position is less demanding, and therefore their true belief can meet or surpass the evidential threshold for knowledge. If one externalizes contextual shifts, then it seems one is committed to skeptical worries undermining ordinary knowledge attributions, whether or not subjects or attributors were aware of them.

> surprised him by depositing money in his account earlier in the day. Keith, being ignorant of his wife's kind deed, says to himself, "banks sometimes change their hours. I don't know the bank will be open tomorrow. I better deposit my check now."

My contention is that in this case, anti-intellectualism coupled with context externalism, gives the intuitively wrong answer, while contextualism, and perhaps purism, does.

Contextualists can explain the first case in the following way: since there's salience of error, Keith is in a high standards context. It doesn't matter that there is no real consequence for him being wrong; his beliefs about the situation are sufficient to elevate his epistemic context from a less, to a more, demanding one.

Purist epistemologists have a more straightforward response. Since only truth-related factors transform a truth belief into an instance of knowledge, whether Keith is mistaken about the practical consequences of his belief is irrelevant. If Keith's evidence (or, if one prefers, strength of epistemic position) is strong enough to meet the evidential threshold for knowledge, he knows the bank will be open on Saturday.

Anti-intellectualists, on the other hand, seem committed to the claim that Keith is in a low standards context, since nothing is *really* at stake for him if he is mistaken. If the bank does change its hours, he will still have the necessary funds to make his house payment. Further, principle-based explanations of Mistake Bank Case (and those similar to it), fail to provide a satisfactory account. Could Keith rationally act on the proposition "the bank will be open on Saturday?" Presumably he could since little is at stake if he's wrong, and therefore, given the low (or moderately low) epistemic threshold for knowledge, his belief that the bank will be open on Saturday is true. Hence, he could be rational to put off going to the bank.

One might object that even though Keith knows the bank will be open, and therefore satisfies the knowledge-action principle's constraint on rational action, his beliefs about the situation fail to make it sufficient for him not to deposit his check. In other words, an anti-intellectualist might respond that even though Keith knows that the bank will be open on Saturday, he should still deposit his check given his beliefs about the situation. However, this response conflicts with other anti-intellectualist theses. Consider the following principle from Fantl and McGrath:

> (KJ) If you know that p, then p is warranted enough to justify you in Φ-ing, for

any Φ.[16]

Given the low stakes in Mistaken Bank Case, it's plausible to think that Keith knows the bank will be open on Saturday, and therefore, he should be able to act on the proposition "I know the bank will be open on Saturday" and put off depositing his check. Instantiating the variables in KJ, we get the following argument that it's rational for Keith to wait to deposit his check:

P1. If Keith knows that the bank will be open on Saturday, then it is rational for him to wait to deposit his check until Saturday.

P2. Keith knows that the bank will be open on Saturday

C. It is rational for him (Keith) to wait to deposit his check until Saturday.

There's an additional worry that needs to be addressed. One might think that anti-intellectualists don't need to be externalists about epistemic contexts. After all, there's nothing built into the structure of the view committing them to such a position. However, there are plausible reasons why an anti-intellectualist would be unwise to rejected context externalism.

Given context externalism, anti-intellectualists are able to account for a potentially infinite number of generable cases contextualists cannot. Any case in which the stakes are high, but the subject is unaware of them, anti-intellectualists can provide an explanation for. If context externalism is rejected, then PE lacks the resources to provide a superior explanation over contextualism and purist epistemologies.

Since PE is subject, rather than attributor focused, what's at stake moves independently of awareness. If S is faced with two cups, one filled with water and the other poison, she need not be aware that anything is amiss for her to be in a high stakes situation.

Hence, while anti-intellectualists are not committed to context externalism, it's nevertheless wise for them to endorsethe view, since they are able to account for ignorant high stakes cases. Additionally, given the subjected focused nature of PE, context externalism strikes me as a natural position for anti-intellectualists to adopt.

[16] Fantl and McGrath, *Knowledge*, 66.

EVOLUTIONARY DEBUNKING: THE DEMARCATION PROBLEM

Christos KYRIACOU

ABSTRACT: Recent literature has paid considerable attention to evolutionary debunking arguments. But the cogency of evolutionary debunking arguments is compromised by a problem for such arguments that has been somewhat overlooked, namely, what we may call *'the demarcation problem.'* This is the problem of asking in virtue of what regulative metaepistemic norm evolutionary considerations either render a belief justified, or debunk it as unjustified. In this paper, I present and explain why in the absence of such a regulative metaepistemic norm any appeal to evolutionary considerations (in order to justify or debunk a belief) is bound to be ad hoc and question-begging and, therefore, ultimately unjustified.

KEYWORDS: evolutionary debunking, 'the demarcation problem,' justification

1. Introduction

Appeals to evolutionary, causal considerations that serve to construct evolutionary arguments (debunking or justifying) are rife in recent philosophical debates. Such evolutionary arguments typically have the following basic form:[1]

> "Causal Premise: S's belief that p is explained by X.
>
> Epistemic Premise: X is an (off-\on-)track process.
>
> Therefore, S's belief that p is (un-)justified."[2]

[1] See, for example, Guy Kahane, "Evolutionary Debunking Arguments," *Nous* 45, 1 (2011): 106 and Paul Griffiths and John Wilkins, "Crossing the Milvian Bridge: When Do Evolutionary Explanations of Belief Debunk Belief?" in *Darwin in the 21st Century: Nature, Humanity, and God*, eds. Phillip R. Sloan, Gerald McKenny, and Kathleen Eggleson (Notre Dame, IN: Notre Dame University Press. 2015), section 1. For a critical response to Griffths and Wilkins, see Christos Kyriacou, "Evolutionary Debunking: The Milvian Bridge Destabilized," *Synthese,* forthcoming. Online publication: DOI: 10.1007/s11229-017-1555-0.

[2] Such 'genealogical' arguments need not be evolutionary in particular, see Russ Shafer-Landau, "Evolutionary Debunking, Moral Realism and Moral Knowledge," *Journal of Ethics and Social Philosophy* 7, 1 (2012):1-2 for discussion. They could be sociological, psychological, historical etc. and understand the causal premise accordingly. See Gilbert Harman, *The Nature of Morality* (Oxford: Oxford University Press, 1977) for a sociological genealogical argument against moral beliefs and Sigmund Freud, "The Future of An Illusion," in *The Freud Reader*, ed. Peter Gay

On the one hand, appeals to evolutionary considerations that result in *debunking* arguments suggest in the ambivalent epistemic premise that X is an *off*-track process.[3] Such arguments have notably been applied to moral, religious, color, ordinary objects and even mathematics and logic beliefs as well as to various kinds of cognitive illusions beliefs (such as so-called positive illusions and thermoreceptive illusions).[4] On the other hand, appeals to evolutionary considerations that result in *justifying* arguments suggest in the epistemic premise that X is an *on*-track process. Such arguments have been applied to cognitive processes (and their doxastic output) such as induction, abduction, deduction, perception, memory, the belief in an external world, understanding of other minds and beyond.[5]

(London, Vintage, 1989), 685-721 for a psychological argument against religious beliefs.

[3] I assume an understanding of the tracking condition in terms of reliability. That is, a process is on-track if and only if it reliably tracks respective facts and produces a preponderance of true beliefs. Otherwise, it is off-track and unreliable.

[4] See Richard Joyce, *The Evolution of Morality* (Cambridge, MA: MIT Press, 2006), Philip Kitcher, "Biology and Ethics," in *The Oxford Handbook of Ethical Theory*, ed. David Copp (Oxford: Oxford University Press,2006), 163-185, Sharon Street, "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127 (2006): 109-166 for debunking of normative beliefs, Joshua Schechter, "Could Evolution Explain our Reliability about Logic?" *Oxford Studies in Epistemology* 20 (2013):214-239 for logic beliefs, Justin Clarke-Doane, "Morality and Mathematics: The Evolutionary Challenge," *Ethics* 122, 2 (2012): 313-340 for maths beliefs, Paul Boghossian and David Velleman, "Colour as a Secondary Quality," *Mind* 98, 389 (1989):81-103 for color beliefs, Daniel Korman, "Debunking Perceptual Beliefs About Ordinary Objects," *Philosophers' Imprint* 14, 13 (2014) for ordinary object beliefs, Ryan McKay and Daniel Dennett, "The Evolution of Misbelief," *Behavioral and Brain Sciences* 32 (2009): 493-513, for positive illusions beliefs, Helen De Cruz, Maarten Boudry, Johan De Smedt, and Stefaan Blancke, "Evolutionary Approaches to Epistemic Justification,"*Dialectica* (2011): 517-535 for thermoreceptive beliefs, and Richard Dawkins, *The God Delusion* (London: Bantam Press, 2006) for religious beliefs. With positive illusions McKay and Dennett ("The Evolution of Misbelief," 505) refer to "unrealistically positive self-evaluations, exaggerated perceptions of personal control or mastery, and unrealistic optimism about the future." With thermoreception De Cruz et al. ("Evolutionary Approaches," 532) refer to "the system that reacts to surface skin temperatures." This cognitive process is not very reliable since it tends, for evolutionary reasons, to represent the temperature conducive to an organism's fitness and survival, not the accurate temperature.

[5] See, for example, W.V.O. Quine, "The Nature of Natural Knowledge," in *Mind and Language: Wolfson College Lectures*, ed. Samuel Guttenplan (Oxford: Clarendon Press, 1975), 67-81 and Sharon Street, "Evolution and the Normativity of Epistemic Reasons," *Canadian Journal of Philosophy*, Supplementary Volume 35 (2009): 213-249 for induction, Alan Goldman, "Natural Selection, Justification and Inference to the Best Explanation," in *Evolution, Cognition and Realism*, ed. Nicholas Rescher (Lanham: University Press of America, 1990), 39-46 for abduction,

In this paper I present and explain a conspicuous but somewhat overlooked problem for such epistemic appeals to evolutionary considerations, what we may call *'the demarcation problem.'* This is the problem of asking in virtue of what regulative metaepistemic norm evolutionary considerations render a belief justified or debunk it as unjustified. In the absence of such a regulative norm, any appeal to evolutionary considerations (in order to justify or debunk a belief) is ad hoc and question-begging and, therefore, ultimately unjustified. Call this *'the adhocness problem.'*

2. Unpacking the Demarcation Problem

The demarcation problem has been around for at least some time and is gaining traction in recent literature.[6] Roughly, the problem is that evolutionary considerations may sometimes be used to justify beliefs (e.g. inductive, perceptual, memorial, external reality beliefs etc.) and sometimes to undermine and debunk beliefs as unjustified (e.g. normative beliefs, color beliefs, religious beliefs, ordinary

---

Joshua Schechter, ''Could Evolution Explain our Reliability about Logic?'' *Oxford Studies in Epistemology* 20 (2013): 214-239 for deduction, Steve Stewart-Williams, ''Innate Ideas as a Naturalistic Source of Metaphysical Knowledge,'' *Biology and Philosophy* 20 (2005): 791-814 for the belief in an independent external world, Ruth Garrett Millikan, *Language, Thought and Other Biological Categories* (Cambridge, MA: MIT Press, 1984) Griffiths and Wilkins, ''Crossing the Milvian Bridge," and Paul Griffiths and John Wilkins, "Evolutionary Debunking Arguments in Three Domains: Fact, Value and Religion," in *A New Science of Religion*, eds. James Maclaurin and Greg Dawes (Routledge, forthcoming) for perception\representation, David Papineau, ''The Evolution of Knowledge,'' in *The Roots of Reason* (Oxford: Oxford University Press, 2003), 39-82 for understanding of other minds, and Stephen Boulter ''The 'Evolutionary Argument' and the Metaphilosophy of Commonsense,'' *Biology and Philosophy* 22 (2007):369-382 for memory.

[6] See for example the discussion in Michael Bradie''Should Epistemologists Take Darwin Seriously?'' in *Evolution, Cognition and Realism*, ed. Rescher, 33-38), Kahane, ''Evolutionary Debunking Arguments," Shafer-Landau, ''Evolutionary Debunking, Moral Realism and Moral Knowledge,'' 35, Justin Clarke-Doane, ''Debunking Arguments: Mathematics, Logic and Modal Security,'' in *The Cambridge Companion to Evolutionary Ethics*, eds. Robert J. Richards and Michael Ruse (Cambridge: Cambridge University Press, forthcoming), section 3, and Jack Woods, ''Mathematics, Morality and Self-Effacement,'' *Nous* (2016): section 4. Michael Vlerick and Alex Broadbent, ''Evolution and Epistemic Justification,'' *Dialectica* 69, 2 (2015):185-203 come close to the problem, but their explication suffers, I think, from a basic mistake that compromises it (see ftn. 10 for the basic mistake). Shafer-Landau, ''Evolutionary Debunking, Moral Realism and Moral Knowledge,'' in particular, notes that evolutionary debunking arguments about a philosophical domain quickly over-generalize to domains that seem beyond serious epistemological doubt and, therefore, we need to disambiguate the metaepistemic norm in virtue of which debunking arguments run and confer unjustifiedness.

objects beliefs, cognitive illusions beliefs etc.) and it is difficult to discern a *principled* way of how such considerations are to be used.[7] That is, it is difficult to discern what the regulative metaepistemic norm is in virtue of which evolutionary considerations are deemed to justify or debunk beliefs.

To be sure, the first-order, epistemic norm of reliability (and truth-trackingness) is the norm that the epistemic premise is relying on, but this is of little help to our meta-problem. This is the case because mere appeal to the epistemic norm of reliability will not do as it is unclear how we demarcate between processes that are reliable and generally on-track and processes that are unreliable and generally off-track. This is again the case because, in principle, evolutionary considerations may, on the basis of the causal premise, be invoked for or against the justification of beliefs depending on how we construe the ambivalent epistemic premise (i.e. involving an on- or off-track process). This is *'the demarcation problem.*[8] Let us illustrate the problem with an example.

---

[7] Of note, is that evolutionary debunking arguments threaten to over-generalize and debunk even themselves (and virtually any other belief), which would be epistemically self-defeating. See Christos Kyriacou, "Are Evolutionary Debunking Arguments Self-Debunking?" *Philosophia* 44, 4 (2016): 1351-1366, "Expressivism, Question Substitution and Evolutionary Debunking," *Philosophical Psychology* 30, 8 (2017): 1019-1042, "Evolutionary Debunking: The Milvian Bridge Destabilized," for discussion of this theme. It is also important to note that some beliefs are rationally indubitable and, therefore, have to be exempt from any debunking. In particular, some rationally indubitable epistemic 'fixed points' have to be assumed as theoretically indispensable for any kind of rational inquiry (inl. debunking inquiry), see Christos Kyriacou, "From *Moral* Fixed Points to *Epistemic* Fixed Points," in *Metaepistemology* eds. Christos Kyriacou and Robin McKenna (London: Palgrave, 2018) for discussion. For a general introduction to metaepistemology, see Christos Kyriacou, "Metaepistemology," in *Internet Encyclopedia of Philosophy* (2016). URL=https://www.iep.utm.edu/meta-epi/.

[8] Some may object that the alleged problem is trading on a conflation of different levels of analysis, namely, normative epistemological and metaepistemological. That is, when various philosophers appeal to evolutionary, causal considerations in order to help justify or debunk a belief, they are not in the metaepistemological business of providing a metaepistemic norm (or even a theory of justification). They are only in the normative business of weighing reasons for or against a belief and they see fitting to take evolutionary considerations into account. This objection, however, misses the point. The point is that such epistemic appeals to evolutionary considerations need to be grounded in a regulative metaepistemic norm, otherwise they would be merely ad hoc. It will not assuage the problem to point out that evolutionary theorists, metaethicists, cognitive scientists etc. are not in the business of traditional epistemological theorizing. The problem remains pressing. See William Alston, "Level Confusions in Epistemology," in his *Epistemic Justification* (Ithaca: Cornell University Press, 1989), 153-171 for discussion of 'level confusion' in epistemology that illustrates our point.

Some have appealed to evolutionary considerations in order to debunk moral and religious beliefs as unjustified while others have appealed to evolutionary considerations in order to justify moral and religious beliefs. The former have argued that evolution explains why we tend to have these sorts of beliefs and, given a plausible naturalistic epistemology, we have no good reasons to hold them because unreliable processes produce these beliefs.[9] The latter have argued in Reidian-Plantingian style that evolution explains why we tend to have this sort of beliefs and the fact that they have, surprisingly, evolved, given a plausible 'reformed epistemology,' renders them prima facie justified as 'properly basic.'[10] These beliefs are justified because reliable processes produce them (at least insofar as they are functioning properly in hospitable conditions without defeaters). Thus, the appeal to evolutionary considerations could, in principle, go

[9] See for instance Joyce, *The Evolution of Morality* and Street, "A Darwinian Dilemma," on moral beliefs and Dawkins, *The God Delusion* on religious beliefs.

[10] This is a distinctively Reidian, 'reformed epistemology' line of thought found in the work of Alvin Plantinga, *Warrant and Proper Function* (Oxford: Oxford University Press, 1993), William Alston, *Perceiving God* (Ithaca: Cornell University Press, 1991), Nicholas Wolterstorff, "Reformed Epistemology," in *The Oxford Handbook of Philosophy of Religion*, ed. William Wainright (Oxford: Oxford University Press, 2008), 245-271 and James Clark Kelly, "Without Evidence or Argument," in *Reason and Responsibility*, eds. Joel Feinberg and Russ Shafer-Landau (Boston MA: Cengage Learning, 2008), 164-168. Notably, Alvin Plantinga, "Is Belief in God Rational?' in *Rationality and Religious Belief*, ed. C. F. Delaney (Notre Dame: University of Notre Dame Press, 1979), 7-27 has argued that religious beliefs are justified because they are properly basic. Evolved processes that are reliable when functioning properly in hospitable conditions produce them. But the evolutionary process must have been "guided and orchestrated by God" (*Warrant and Proper Function*, 236), otherwise, given that natural selection opts for survival and reproduction and not strictly speaking truth, we would have no good reason to trust our cognitive faculties themselves, something that would lead evolutionary theory to epistemic self-defeat. This much recapitulates his 'evolutionary argument against naturalism' (see Griffiths and Wilkins, "Crossing the Milvian Bridge," section 3 for criticism). Of note, is that Plantinga ("Is Belief in God Rational?", *Warrant and Proper Function*) does not so much discuss moral beliefs, but it is rather obvious how his case for religious beliefs could carry over to the moral case. Indeed, in the same spirit, some have argued that natural selection tracks –imperfectly- moral facts (see Kevin Brosnan, "Do the Evolutionary Origins of Our Moral beliefs Undermine Moral Knowledge?" *Biology and Philosophy* 26 (2011):51-64 and William FitzPatrick "Debunking Evolutionary Debunking of Ethical Realism," *Philosophical Studies* 172 (2015): 883-904. Moreover, Michael Huemer's ("Compassionate Phenomenal Conservatism," *Philosophy and Phenomenological Research* 74 (2007):30-55, *Ethical Intuitionism* (London: Palgrave MacMillan, 2008) 'phenomenal conservatism' idea is sufficiently similar to Reidian 'reformed epistemology' and has been applied to the prima facie justification of moral beliefs (although Huemer, *Ethical Intuitionism*, 54-60 is an antireductionist realist critical of attempts to ground a realist morality to God).

either way, justifying or debunking, depending on how we construe the ambivalent epistemic premise (i.e involving an on- or off- track process).

So, unless we can provide a metaepistemic norm that regulates the application of the first-order, epistemic norm of reliability and clarifies which processes are on-track and which are off-track, the first-order epistemic appeals to evolutionary considerations would seem ad hoc and question-begging and, therefore, unjustified. They would seem ad hoc and question-begging because they could, in principle, go either way due to underdetermination by evolutionary, causal considerations (as captured by the causal premise). Call this '*the adhocness problem.'*

The adhocness problem places a *constraint* for a plausible solution to the demarcation problem. Unless it is satisfied, the proposed metaepistemic norm cannot be the missing metaepistemic norm we are looking for. For the missing metaepistemic norm should be capable of demarcating under what conditions first-order appeals to evolutionary considerations should apply. Let us spell out a bit more the adhocness problem.

It is well-known that cognition is often beset with irrational biases, heuristics and effects, such as the confirmation bias and the affect heuristic. [11]Thus, in the absence of a regulative metaepistemic norm, we could subconsciously be appealing to evolutionary considerations in order to justify beliefs we want them justified because they are consoling and debunk beliefs we want them unjustified because they are disquieting.[12]

In such a scenario, we may be caught in a coherentist, inferential circle that perpetuates a confirmation bias driven by the affect heuristic. That is, we could have a coherent belief system that approves only what confirms the belief system on the basis of what we would like to believe. This is what Paul Boghossian in his *Fear of Knowledge,*[13] calls 'norm-circular justification' and it is in essence a version of the well-known epistemic circularity problem.[14]

---

[11] See Daniel Kahneman, *Thinking, Fast and Slow* (London: Penguin, 2011), Jonathan Haidt, *The Righteous Mind* (London: Penguin, (2012)

[12] Such phenomena are well-studied by cognitive psychologists. See for example some of the discussion in Kahneman, *Thinking, Fast and Slow* and Haidt, *The Righteous Mind*. Philosophers have known of epistemic circularity since the conception of the Agrippan trilemma (or 'the problem of the wheel'). See Michael Williams, *Problems of Knowledge*. (Oxford: Oxford University Press, 2001)) for some discussion of the Agrippan trilemma about justification.

[13] Paul Boghossian, *Fear of Knowledge* (Oxford: Clarendon Press, 2006), 79.

[14] See William Alston, "Epistemic Circularity," in his *Epistemic Justification*, 319-349. An epistemic circularity problem has also been applied to reliabilism, namely, 'the bootstrapping problem' (see Jonathan Vogel, "Reliabilism Leveled," *Journal of Philosophy* 97, 11 (2000): 602-

In other words, in the absence of a regulative metaepistemic norm that could be used as a 'measure stick' of what is to be deemed justified and what is to be debunked as unjustified by appeal to evolutionary considerations, we could be trapped in a subtle confirmation bias and a vicious epistemic circle (driven by the affect heuristic) confirming as justified what suits us to think is justified and debunking what suits us to think is unjustified.[15] Inevitably, in such cognitive conditions our reasoning would be epistemically defective (because it would be unreliable or, if reliable, only coincidentally so).

Of course, this is not how things ought to work in epistemic matters. Prima facie, beliefs can be justified or not independently of whether we would like them (un-)justified and we ought to believe what is justified, given evidence, not what we would like to be justified.[16] But in order to avoid the adhocness problem, we need to ground epistemic appeals to evolutionary considerations on a metaepistemic norm that arbitrates the application of the epistemic reliability

---

623). The bootstrapping problem is, roughly, the problem that we need to assume the reliability of a process before moving on to rely on its doxastic output as reliable and at the same time we need to assume the reliability of the doxastic output if we are to rely on the reliability of the process. The upshot is that reliabilism is committed to a circularity problem. Michael Vlerick and Alex Broadbent, "Evolution and Epistemic Justification," *Dialectica* 69, 2 (2015):185-203 have concurred that a circularity problem can be found in evolutionary arguments, but they go on to propose that we can distinguish between 'virtuous, non-self-certifying' circles and 'vicious self-certifying' circles *within* the framework of naturalism. This is, however, problematic because in the paper they just *assume* naturalism, which is viciously self-certifying in the most fundamental of ways because it begs the question against antireductionism about normativity, maths, modality, logic, religion etc. Thus, the problem remains at the fundamental metaphysical level. Unfortunately, I have to forgo detailed discussion here.

[15] It is sometimes thought that theists tend to believe in some God due to the pragmatic, psychological utility this has (consoling beliefs in immortality, a Freudian father figure, life in heaven etc.). But the same 'rationalizing' style of reasoning applies to atheists' psychology of belief as well. It can as easily be said that they tend not to believe in some God due to the pragmatic, psychological utility this has (absence of an independent 'measure of all things,' an authority figure, fear of punishment etc.) Thus, opium-of-the-people style of reasoning could go either way. See Guy Kahane, "Should We Want God to Exist?" *Philosophy and Phenomenological Research* 82, 3 (2011): 674-696 for discussion of a similar point due to Thomas Nagel.

[16] See Haidt, *The Righteous Mind* for extensive empirical work confirming that our judgments are often driven by emotions, desires etc. and then reason, as a Humean obedient servant, hastens to offer post hoc rationalizations for these judgments. Of course, this empirical evidence need not vindicate any instrumentalism about moral or epistemic rationality. See McKay and Dennett, "The Evolution of Misbelief," for some discussion of how and why we are prone to deceive ourselves.

norm: which processes are reliable and on-track and which are unreliable and off-track. I conclude that we need some sort of metaepistemic measure stick for epistemic appeals to evolutionary considerations.

IS EPISTEMIC BLAME DISTINCT FROM MORAL BLAME?

Daniella MEEHAN

ABSTRACT: In contemporary epistemology, recent attempts have been made to resist the notion of epistemic blame. This view, which I refer to as 'epistemic blame skepticism,' seems to challenge the notion of epistemic blame by reducing apparent cases of the phenomenon to examples of moral or practical blame. The purpose of this paper is to defend the notion of epistemic blame against a reductionist objection to epistemic blame, offered by Trent Dougherty in "Reducing Responsibility." This paper will object to Dougherty's position by examining an account in favour of epistemic blame and demonstrate concerns over the reductionist methodology employed by Dougherty to argue for his sceptical position.

KEYWORDS: epistemology, ethics, blame, scepticism, reductionism, responsibility

1. Introduction

Talk of responsibilities, duties, and blameworthiness is a widespread phenomenon in the fields of epistemology and ethics. These fields frequently draw from one another, and the exploration of epistemic and moral blame is one of the most recent examples of this overlap.[1] The discussion of epistemic blame is not just limited to epistemology and ethics, but is also pervasive in our everyday lives and plays an important part in society. Our everyday language implies a concept of epistemic blame as we often talk of holding people accountable for their beliefs, stating that one 'should have known better' or 'they ought to believe that x.'[2] We also have special kinds of words and concepts for people who are notoriously irresponsible or bad believers, as opposed to when their beliefs are excusable. These different concepts seem to rely on the idea that we can be responsible and blameworthy believers. However, it is not entirely clear how epistemic blame is distinct from moral or instrumental blame, and whether it is a form of blame in its own right. This paper examines this distinction in depth, offering an argument for the independence of epistemic blame as a distinct concept.

[1] See Jessica Brown, "Blame and wrongdoing," *Episteme* 14, 3 (2017): 275-296.

[2] See Corey Cusimano, "Defending Epistemic Responsibility," *Arché* 4, 1 (2012): 32-59.

The plan of this paper is as follows; in section two I will briefly summarize the importance of this debate and the perceived relationship between epistemic and moral blame. I will then present Trent Dougherty's reductionist objection against the distinctiveness of epistemic blame from moral blame.[3] After setting up Dougherty's objection, the remainder of my paper will respond to his concerns. In section three, I will offer my first objection to Dougherty's position by presenting an argument in favour of a distinct form of epistemic blame, offered by Nikolaj Nottelmann.[4] I aim to weaken Dougherty's objection towards the notion of epistemic blame by assessing Nottelmann's argument and Dougherty's insufficient response to it. After providing my own objection to Dougherty's challenge to Nottelmann's position, I will also assess a potential response Dougherty could offer against my defence. Despite my charitable attempt to save Dougherty's position, I will also find this objection to be unsuccessful. In the third section of this paper, I argue against the reductionist methodology employed by Dougherty to object to the possibility of epistemic blame. I will draw upon an argument provided by Scott Stapleford who defends the existence of epistemic duties against similar reductionist arguments offered against their distinctiveness from moral or instrumental duties.[5] Developing Stapleford's argument arguably provides support for the distinctiveness of epistemic blame by demonstrating how a reductionist reasoning leads to some problematic and odd consequences. With both responses to Dougherty's epistemic blame scepticism presented, the overall conclusion of my paper will find Dougherty's argument against the distinctiveness of epistemic blame from moral and instrumental blame, unsuccessful.

2. Epistemic Blame Scepticism

We routinely make judgements about what one ought to or ought not to believe. You ought not to believe falsehoods, or believe without sufficient evidence or justification, for example. When we make these judgements, we often respond negatively when people fail to comply. We acknowledge that they have failed in some sense, or done something wrong, and we regard them blameworthy by holding them responsible for these wrongdoings. On face value, it appears that this form of blame is epistemic in its nature, in that it is an epistemic evaluation made about an epistemic action or lack of action. As Cusimano notes, philosophers

[3] Trent Dougherty, "Reducing Responsibility: An Evidentialist Account of Epistemic Blame," *European Journal of Philosophy* 20, 4 (2012): 534-547.

[4] Nikolaj Nottelmann, *Blameworthy Belief: A Study in Epistemic Deontologism* (New York: Springer, 2007).

[5] Scott Stapleford, "Why There May Be Epistemic Duties," *Dialogue* 54, 1 (2015): 63-89.

traditionally associate the goal of truth as one of the defining features of the epistemic realm and the responsibilities associated with this are also concerned with achieving the truth.[6] Arguably, it seems to naturally follow from this that if the epistemic responsibilities are epistemic in nature, the blameworthiness that we attribute is due to a failure to carry out an epistemic responsibility, so is itself epistemic.

However, despite the arguably initial appeal and popularity of epistemic blame, in recent literature, some epistemologists have questioned the notion of epistemic blame and rejected it in its entirety. I will refer to this stance as 'epistemic blame scepticism.' In short, epistemic blame sceptics reject the claim that there is a distinctive form of epistemic blame, often reducing apparent cases of such to moral or practical blame. From this reasoning, sceptics claim that the notion of epistemic blameworthiness becomes redundant, meaning there is no need for it to exist in the literature, as a distinct form of epistemic blame would over-complicate the taxonomy and direct attention away from the real type of blame at hand.

As mentioned, one prominent epistemic blame sceptic is Dougherty, who offers a reductionist objection against the notion of epistemic blame.[7] It is his objection that this paper will focus on, and we can now turn to examine his objection in more depth.

2.1 Dougherty's 'Reducing' of Epistemic Responsibility

Dougherty presents a variety of arguments in favour of epistemic blame scepticism, centredaround the key claim that epistemic responsibility can be 'reduced.' What Dougherty means by this claim, is that cases which appear to concern a distinct type of epistemic responsibility can be 'reduced,' into other types of blame. Epistemic responsibility or blame identifies with other forms of blame on a base level, so arguably, there is no need to overcomplicate matters and define these types of blame as epistemic, especially not as *distinctively* epistemic.[8]

According to Dougherty, most cases of seemingly epistemic blameworthiness are either cases of moral or instrumental blameworthiness or cases where no blame should be attributed at all. More specifically, Dougherty

[6] Cusimano, "Defending Epistemic Responsibility," 34.

[7] Dougherty, "Reducing Responsibility," 534-547.

[8] I will infer that Dougherty's use of 'responsibility' is interchangeable with 'blameworthiness.' I am aware that the notions of responsibility and blameworthiness can come apart (for example see Thomas Michael Scanlon, *What We Owe to Each Other*, (Cambridge: Harvard University Press, 1998), however, due to word constraints I will not be discussing this material in this paper.

claims that cases of epistemic blameworthiness are not part of epistemology and should be understood as falling within the domain of applied ethics, on par with medical and business ethics in that it is an aspect of ethical theory applied to a certain domain.[9]

Dougherty summarises his reductionist argument in the form of his 'identity thesis,' understood as follows;

> *"IT: Each instance of [so-called] epistemic irresponsibility is just an instance of purely non-epistemic irresponsibility/ irrationality (either moral or instrumental)."*[10]

It is important to note here that Dougherty still believes in a form of epistemic normativity, but that it does not lead to a robust 'ethics of belief' which responsibilists believe in. The only epistemic demands, and thus epistemic forms of responsibility and blameworthiness, are ones relating to evidential fit.[11]

Dougherty argues that epistemic 'oughts' should only be understood as the following;

> *"(EO) One epistemically-ought to believe p if and only if p fits one's evidence."*[12]

Dougherty provides further support for his reductionist thesis by presenting an example to demonstrate how epistemic blame collapses into either moral or instrumental blame.[13] We can briefly sketch this example now to further illustrate how Dougherty explains away an intuitive case of epistemic blameworthiness.

Craig the Creationist

Craig is a dysfunctional agent. He believes in creationism, the view that the universe and living organisms originate from acts of divine creation, as opposed to natural processes such as evolution. Craig was raised within a community of creationist believers. His parents believed in creationism, his school taught and favoured creationism and he only read books with a creationist bias. We can now imagine that I happen to meet Craig, and upon hearing of his creationist view, offer him some books on the topic which discuss the evolutionary viewpoint.

---

[9] Trent Dougherty, "The 'Ethics of Belief' is Ethics (Period): Reassigning Responsibilism," in *The Ethics of Belief*, eds. Jonathon Matheson and Rico Vitz (Oxford: Oxford University Press, 2014),146-168.

[10] Dougherty, "Reducing Responsibility," 537.

[11] As an evidentialist, Dougherty claims that 'lack of evidential fit' is a genuine epistemic criticism which one is blameworthy for.

[12] Dougherty, "The Ethics of Belief is Ethics (Period)," 153.

[13] Dougherty, "Reducing Responsibility," 538.

However, Craig blindly refuses to read them, not wishing for his beliefs to be challenged.

From this information, it would appear that Craig's initial belief in creationism satisfies the standards for synchronic rationality (as his beliefs fit the evidence he had at the time, prior to our conversation), but he fails on diachronic rationality, i.e. an assessment of rationality across time. If we focus on the time in which I offered Craig the evolutionary books and he refused to read them, this arguably appears to be a case of epistemic irresponsibility. Craig had plenty of free time to read the books if he desired, and they are relatively short. By refusing to do so, however, he appears to be willfully ignorant, which is epistemically irresponsible. Upon closer examination, however, Dougherty argues that the irresponsibility at hand is really a case of moral or instrumental irresponsibility.[14] Dougherty argues for this statement by appealing to stakes, claiming that either there is something at stake for Craig, or not. If there is not something at stake, then Craig does nothing irresponsible or blameworthy in not being over-scrupulous in his creationist beliefs. If, on the other hand, there is something at stake for Craig, then it either relates to his own interests or the interests of others. If the former, then it would be instrumentally irresponsible and irrational for Craig to continue to sustain his beliefs in creationism, for he is actively believing in a falsehood which is a personal disadvantage to him. If the stakes regard the interests of others, as we have a duty to promote the interests of others, Craig's beliefs would be deemed morally irresponsible. As such, Dougherty explains away the intuitive attribution of epistemic irresponsibility to Craig's action by reducing it to cases of instrumental and moral irresponsibility. The form of blameworthiness which we would attribute here would be either instrumental or moral, as it would only be appropriate to blame Craig epistemically if there was something epistemically at stake, which there is not.

In summary, Dougherty is claiming that perceived cases of epistemic blame can be reduced to cases of moral or instrumental blame. Applying a form of Ockham's razor, there is no need to overcomplicate matters by arguing for a new species of blame, which arguably only distracts from the other types of blame we should be really focusing on. The remainder of this essay will aim to resist Dougherty's claims, arguing against his identity thesis.

[14] Dougherty, "Reducing Responsibility," 540.

3. Nottelmann's Account of Epistemic Blame

Having outlined both epistemic blame and epistemic blame scepticism, we can now turn to critically assess the argument put forth by Dougherty. We can begin by presenting an argument in favour of a distinct form of epistemic blame, offered by Nottelmann which is discussed and dismissed briefly by Dougherty.[15] I will critically assess Dougherty's objection to Nottelmann's stance, in turn providing a novel defence of Nottelmann argument for epistemic blame.

Nottelmann argues for the distinctiveness of epistemic blame by establishing a theory of epistemic deontologism built upon epistemic blame. By appealing to legal considerations, Nottelmann makes the claim that moral culpability presupposes epistemic culpability, which demonstrates how moral and epistemic blame are distinct. Nottelmann opens his argument for this by detailing a historic rape case from 1975, which caused widespread controversy when three men were not deemed blameworthy for their act of rape.[16] The case consisted of three men, who were invited by their friend, Mr. Morgan, to have sexual intercourse with his wife. Mr. Morgan informed his friends that his wife was 'kinky' and would feign protest. When arriving at the Morgan household, all four men forcibly dragged Mrs. Morgan from her son's bed where she was sleeping, and each had forcible intercourse without her consent whilst the other men held her down. Mrs. Morgan attempted to scream for her son to call the police but was choked by the men. At the trial, the three men pleaded that they believed Mrs. Morgan had consented to sexual intercourse. In conclusion, The House of Lords held that the men made an honest, but mistaken, belief that Mrs. Morgan was consenting, which provided a complete defence.

However, Nottelmann claims that the men should have been considered blameworthy for their actions by arguing for a distinctive form of epistemic blameworthiness. From this, he argues that if epistemic blameworthiness is not reducible to moral blameworthiness, moral blameworthiness must presuppose epistemic blameworthiness. Nottelmann locates the blameworthiness of the rape in the men's belief that Mrs. Morgan consented to sexual intercourse, stating it has "epistemically undesirable properties (such as unreasonableness)."[17] It is this unreasonable belief which motivates the immoral act of rape, which leads Nottelmann to make the claim that epistemic culpability is presupposed by moral culpability. He appeals to a classic distinction in law known as the *actus reus* and

[15] Dougherty, "Reducing Responsibility," 536.
[16] Nottelmann, *Blameworthy Belief*, 3-5.
[17] Nottelmann, *Blameworthy Belief*, 10.

the *mens rea* distinction, to further this presupposition.[18] The *actus reus* refers to the conduct element of a crime, which the defendant must have proven to have done. The *mens rea* is the psychological element of the crime, the intention or forethought which makes one morally culpable. Nottelmann compares the moral blameworthiness to the *actus reus*, and epistemic blameworthiness to the *mens rea.* As the intention comes prior to the action, this means that an agent must hold an epistemically undesirable belief prior to carrying the immoral action. This demonstrates how a clear-cut distinction can be made between the two forms of blame.

It is worth noting here that so far, Nottelmann appears to have demonstrated that there are cases in which the basis for blameworthiness is epistemic, but only with regards to the rape case. It may be true that this is not always the case, and Nottelmann offers little insight as to what other types of cases he also believes the basis for blame is epistemic. However, I do not take this as a concern of Nottelmann's argument, for he arguably does not need more than this modest claim to make his point. If there are examples where epistemic blame comes prior to moral blame, it simply cannot be the case that it reduces to moral blame. An agent must hold an epistemically unreasonable belief prior to the immoral act which the belief stems from, meaning epistemic blame must come prior to moral blame.[19]

In summary, Nottelmann has argued for a distinctive form of epistemic blame by locating blameworthiness in an agent's unreasonable belief. With an appeal to legal considerations, Nottelmann has argued that moral culpability presupposes epistemic culpability, which demonstrates how moral and epistemic blame are distinct.

3.1 Dougherty's Objection to Nottelmann's Position

Having briefly summarized Nottelmann's main argument for the distinctiveness of epistemic blame we can now turn to examine the concerns raised with his view by Dougherty.

Dougherty rejects Nottelmann's position by arguing that just because the target of the blameworthiness is the belief, it does not follow that the nature of the blame is epistemic; beliefs can also be governed by moral, prudential norms.[20] Additionally, Dougherty claims that blame is located in the moral consequences of

[18] Nottelmann, *Blameworthy Belief*, 10

[19] I thank Mona Simion for raising this point in personal conversation.

[20] Dougherty, "Reducing Responsibility," 537.

the act itself, and this is distinctively moral, not epistemic.[21] Taking both of Dougherty's concerns into consideration, it seems Nottelmann fails to locate epistemic blameworthiness in the belief of a guilty agent or demonstrate how the blameworthiness we speak of is distinctively epistemic. It thus appears that Nottelmann fails to successfully argue for the distinctiveness of epistemic blame by appealing to the priority of epistemic blame over moral blame.

Despite Dougherty's concern, I believe we can resist his objection by claiming that the denial of epistemic irresponsibility results in the eradication of any moral irresponsibility too. If this commitment is correct, then it demonstrates how epistemic responsibility must come prior to moral responsibility, as Nottelmann originally claimed. So how can one deny the existence of epistemic irresponsibility? One could argue that the rapists may have searched for more evidence about Mr. Morgan's claim that his wife wanted to partake in sexual intercourse and found positive reasons to believe it. Alternatively, perhaps they had no way to improve their epistemic situation, for example, they had no epistemic defeaters against the claim. Despite the intuition that the three men were aware Mr. Morgan was lying, these epistemic situations do not seem too far-fetched. With this in mind, how does denying any claims of epistemic responsibility deny claims of moral responsibility? If we argue that there was nothing the men could do to better their situation and were therefore truly justified in believing that Mrs. Morgan enjoyed non-consensual sexual intercourse, there no longer seems to be any attribution of blame, moral or epistemic. Their epistemic situation may, at most, make them ignorant, but not culpably ignorant.

We can apply Goldman's case of the 'benighted cognizer'here to explain this point further, which I believe strengthens my response to Dougherty.[22] Goldman details a society which uses unreliable methods to form beliefs about the future. The society uses astrology and oracles to assist in belief formation, thus ignoring proper scientific practice. We can imagine that a member of this society forms a belief about the outcome of an upcoming battle based on zodiacal signs. Goldman refers to this individual as a benighted cognizer, someone who has formed a belief via bad methods but knows no better way to inform himself.[23] Arguably, it seems wrong to attribute any type of blame to the benighted cognizer for his faulty belief formation, despite the potentially disastrous consequences, for the individual has good reason to trust his cultural peers and has no way of acquiring better belief

[21] Dougherty, "Reducing Responsibility," 537.

[22] Alvin Ira Goldman, "Strong and Weak Justification," *Philosophical Perspectives* 2 (1988): 51-5-3.

[23] Goldman, "Strong and Weak Justification," 57.

formation methods. We therefore find it hard to fault or blame them for believing what they do.

Bringing our argument back to Dougherty's objection, we can argue in defence of Nottelmann that blameworthiness is not located in the moral consequences of the act itself, for all moral consequences are eradicated if epistemic responsibility is also eradicated. The cognizer appears to be epistemically justified in their belief, and this excuses any sort of epistemic blameworthiness. It thus appears that blame can be distinctively epistemic and presuppose moral blameworthiness, for the men escape any attributions of moral blameworthiness if they are not deemed epistemically blameworthy.

It is worth addressing here however, a possible attempt Dougherty could present to deny our above objection. One way Dougherty may respond could be to claim the benighted cognizer is not morally blameworthy. Goldman's case of the benighted cognizer is similar to Dougherty's own case of Craig the creationist, where Craig also formed faulty beliefs under bad epistemic situations. With this in mind, perhaps it is possible for Dougherty to appeal to the same argument for this and claim that the benighted cognizer was not morally blameworthy, as nothing was at stake for him. This way, the reason we do not intuitively want to attribute blame does not rest upon there not being any attribution of epistemic blame. However, I think it seems quite clear that there is something at stake for the benighted cognizer, (e.g. the battle could go wrong), and yet, we still do not attribute blame. It seems then that Dougherty would be wrong to argue that cases which are not blameworthy are cases where nothing is at stake, meaning blame is not necessary located in what is at stake morally or practically, for there are cases of such where we do not attribute blame.

4. A Concern for the Reductionist Methodology

One way to resist Dougherty' scepticism is to demonstrate how his reductionist methodology results in some odd and worrisome consequences. It is worth reminding ourselves that Dougherty offers a reductionist argument in favour of epistemic scepticism, claiming that epistemic blameworthiness is a disguised form of moral or instrumental blameworthiness, and therefore is not a distinct field of blame. Examining the literature on epistemic dentologism can be helpful to demonstrate how taking this reductionist approach to the normative domains, can be problematic. Drawing from an argument offered in defence of epistemic deontology against reductionism, I will now outline how this raises concerns for Dougherty's methodology.

Epistemic dentologism is the view that there are certain duties pertaining to a distinct epistemic domain which we are subject to qua rational beings.[24]Sceptical arguments, similar to those offered by Dougherty are used to object against the possibility of distinct epistemic duties. Taking the same form of argument, epistemic duty sceptics argue that epistemic duties can be reduced to moral or practical duties, meaning there is no need for a distinct epistemic deontology.[25] Whilst no parallels have been made between the literature to date, I believe the similarities between the reductionist objection of epistemic duties bears a clear resemblance to Dougherty's reductionist objection towards epistemic blameworthiness. With this in mind, objections made against the reductionist objection to epistemic duties may be valuable in defending epistemic blame against reductionist approaches to epistemic blame scepticism. We can now turn to assess such an objection, offered by Stapleford in his "Why There May Be Epistemic Duties" who defends the distinctiveness of epistemic by demonstrating how a reductionist reasoning leads to some problematic and odd consequences.[26]

Whilst Stapleford does not offer a positive argument for the possibility of epistemic duties, he arguably highlights how the reductionist reasoning is ineffective in dismissing the possibility of epistemic duties. The epistemic sceptic (now understood in both senses of duty and blameworthiness) argues that all cases of epistemic blame or epistemic duties can be reduced to moral blameworthiness or moral duties. However, Stapleford argues that cases where there is a legal duty or blame, which also imposes a moral duty or blame, should be reduced to just cases of moral duties or blame by the reductionist methodology.[27] For example, it seems to be the case that situations which pose a legal duty to do x, also imposes a moral duty to do x, in the sense that laws are often perceived as providing guidance for promoting fairness.[28] However, it seems right that we want to keep legal and moral

[24] See Anthony Robert Booth, "Deontology in Ethics and Epistemology," *Metaphilosophy* 39, 4-5 (2008): 530-545.

[25] The main proponent for this view is Wrenn who, in short, argues that if distinct forms of epistemic duties existed then they would conflict with our other type of obligations, such as our moral, legal and prudential duties. When it appears to be that we have an epistemic duty conflicting with another source of obligation, what we really have is a disguised moral duty competing with some other non-epistemic requirement. Thus, epistemic obligations simply do not exist. See Chase Wrenn, "Why There Are No Epistemic Duties," *Dialogue: Canadian Philosophical Review / Revue Canadienne De Philosophie* 46, 1 (2007): 115–136.

[26] Stapleford, "Why There May Be Epistemic Duties," 63-89.

[27] Stapleford, "Why There May Be Epistemic Duties," 70.

[28] See Andrei Marmor, "Authority, Equality and Democracy," *Ratio Juris* 18, 3 (2005): 315–345, and Andrei Marmor, "How Law is Like Chess," *Legal Theory* 12, 4 (2006): 347–371.

duties distinct; what is considered legal is not always considered to be moral. For example, I may have a legal duty to pay my parking ticket fines, however, it would be odd to claim that my legal duty is also a moral one.

For the reductionist, however, it cannot be true that we have both legal and moral obligations, for reductionism demands that we simplify legal duties or legal forms of blameworthiness into moral duties and blameworthiness. Stapleford argues that this line of reasoning also applies to instrumental duties.[29] Instrumental duties can be understood as legal duties in that it is beneficial to conform one's actions to the law. Take for example paying taxes, not speeding or running red lights, here it is instrumentally good to conform to one's legal duties to avoid fines or imprisonment. This seems puzzling then when we realize that cases of what seems like a prudential duty can be collapsed into legal duties, and legal duties can be reduced to moral duties. The same applies to blameworthiness. Failing to carry out one of these practical duties may seem practically blameworthy, which in turn can be reduced to legal blameworthiness, which can be understood even further as moral blameworthiness.

It appears then that the very same reductionist reasoning employed by epistemic blame and normative sceptics creates a total collapse of the normative realms. Stapleford argues that this is extremely concerning for the epistemic normativity sceptic, for they need to preserve the autonomy of the moral realm to make the claim that epistemic obligations are really disguised moral requirements. For this claim to be considered as credible, it cannot preclude genuine legal and prudential requirements, for we readily do recognise these as independent sources of obligation.

Arguably the epistemic sceptic may attempt to resist these consequences would be to bite the bullet and accept that only moral sources of blameworthiness or obligations exist. However, this is arguably a commitment Dougherty would struggle to accept, for as mentioned previously, Dougherty prescribes to an evidentialist viewpoint, and is therefore committed to the view that there are epistemic ought's regarding evidential fit. Dougherty would therefore be strongly against the idea that these epistemic duties should be collapsed into purely moral obligations, meaning this concern over his reductionist methodology creates a worrying objection to his view.

[29] Stapleford, "Why There May Be Epistemic Duties," 74.

5. Conclusion

I have examined the distinction between epistemic and moral blame, with regards to Dougherty's reductionist argument against the existence of epistemic blame. I found his argument to be unsuccessful, particularly when pitted against the arguments offered by Nottelmann and Stapleford. The main aim of this paper has been to defuse Dougherty's reductionist argument. With this aim in mind, I have not yet attempted to provide a positive reason for the distinctive of epistemic blame. However, a denial of Dougherty's reductionist argument provides a necessary preliminary to make way for prospective arguments for the possibility of epistemic blame.

KNOWING HOW ONE KNOWS

Giovanni ROLLA

ABSTRACT: In this paper, I argue that knowledge is dimly luminous. That is: if a person knows that *p*, she knows how she knows that *p*. The argument depends on a safety-based account of propositional knowledge, which is salient in Williamson's critique of the 'KK' principle. I combine that account with non-intellectualism about knowledge-how – according to which, if a person knows how to φ, then in nearly all (if not all) nearby possible worlds in which she φes in the same way as in the actual world, she only φes successfully. Thus, the possession of first-order propositional knowledge implies second-order practical knowledge, and this can be iterated. Because of the assumed non-intellectualism about know-how, dim luminosity does not imply bright luminosity about knowledge, which is expressed by the traditional KK principle. I conclude by considering some potential counterexamples to the view that knowledge is dimly luminous.

KEYWORDS: luminosity, knowledge-that, knowledge-how, safety

1. The Story So Far

Since Hintikka's[1] ground-breaking work, the thesis known as the 'KK' principle has been under dispute. It says that if one knows that *p*, one knows (or is in a position to know) that one knows that *p*. The interest on the principle quickly reached debates beyond that of its acceptability for systems of epistemic logic. The internalist-externalist dispute on knowledge is perhaps the most prolific discussion on that matter. Internalists, following the traditional view that knowledge is an internally accessible state, were inclined to endorse something like the KK principle, whereas externalists were confident about its rejection.

Many externalists reject the principle based on the following argument:[2] If knowing that *p* implies having a true belief that *p* that also satisfies an external condition – say, being reliably produced – then knowing that one knows that *p* implies knowing that the belief in question is produced by a reliable process. However, most people who know many mundane propositions seldom entertain beliefs about the reliability of their belief-forming processes, so they certainly do

[1] Jakko Hintikka, *Knowledge and Belief* (Ithaca, New York: Oxford University Press, 1962); Jakko Hintikka, "Knowing That One Knows," *Synthese* 21 (1970): 141–62.

[2] For a run-down of these instances, see Samir Okasha, "On a Flawed Argument against the KK Principle," *Analysis* 73, 1 (2013): 80–86.

not know that they know. Thus, the externalist argument goes, the KK principle is false.

Samir Okasha[3] has made a very compelling point that the argument above is logically flawed. The problem is that externalists have systematically committed an intentional fallacy. The second 'K' in the KK principle occurs in an intentional context, so knowing that one knows does not imply knowing the following (implausible, anyway) conjunction: that one believes that *p*, *p* and one's belief that *p* was reliably formed. Instead, second-order knowledge implies only that: one believes that one knows that *p*, it is true that one knows that *p* and one's belief that one knows that *p* was reliably formed. This is a much more plausible take on the KK principle, one that is at least prima facie compatible with externalists views on knowledge.

But there are more compelling arguments against the principle. Tim Williamson's[4] in particular makes a convincing case against the putative luminosity of knowledge. Williamson defines luminosity through the following condition:[5]

> (L) For every case α, if in α C obtains, then in α one is in a position to know that C obtains.

For reasons that will soon become apparent, I take the condition above to express a specific kind of luminosity, one that I call "bright luminosity." Williamson's argument against bright luminosity for knowledge depends on a scenario such as this:

> *Mrs. Coldhot:*
>
> Mrs. Coldhot wants to keep track of whether she is felling cold or hot at each millisecond during a whole morning. Imagine that, at the beginning of the morning, Mrs. Coldhot feels cold. Imagine also that, at noon, she feels hot. Now suppose for reduction that feeling cold or hot satisfies the (bright) luminosity condition. If knowledge implies that one could not easily be mistaken about one's target belief, then if one knows that *p* at any given time, it follows that *p* is true not only at the time one knows, but also true at the next millisecond, for these moments are very similar. Therefore, when Mrs. Coldhot iterates the process of judging how she feels, she knows that for any millisecond that comes after the beginning of the morning, that she feels cold. Then, she feels cold at noon –but that contradicts our initial supposition, so feeling cold or hot is not (brightly) luminous.

[3] Okasha, "On a Flawed Argument."

[4] Timothy Wiliamson, *Knowledge and Its Limits* (Oxford: Oxford University Press, 2000).

[5] Cf. Williamson, *Knowledge and Its Limits*, 95.

According to Williamson, the argument applies equally well to every non-trivial condition. Since knowledge is a non-trivial condition, it follows that knowledge is not (brightly) luminous.

That is a very plausible outcome, but we may not be ready to give up on the luminosity of knowledge just yet. I argue that there is a conceptual connection between some of our mental states and our knowledge of them. Thus, I follow what Stalnaker[6] claims to be the standard way of reconciling the KK thesis, or something akin to it, with externalists conceptions of mind and knowledge. What is specific about this strategy is that the relevant knowledge of our mental states is practical rather than propositional or direct. So first-order propositional knowledge implies second-order practical or procedural knowledge. In other words, knowing that *p* implies knowing how to know that *p*. This, however, depends on a non-intellectualist construal of know-how. I will not argue for that latter position, nor will I argue, at least not directly, against the intellectualism proposed by Stanley, Stanley and Williamson, Brogaard or any others.[7] The results advanced here, however, do depend on a non-reductive view on knowledge-how. Despite the fact that the account of know-how presented here is non-reductive, it allows for a conceptual analogy between propositional and practical knowledge, thus, it satisfies a desideratum of explanatory unity of knowledge without implying intellectualism.

2. Dim Luminosity and Knowing How to Know

Not everything that is luminous is bright. A condition C is dimly luminous if and only if (D-L) obtains:

(D-L) For every case α, if in α C obtains, then in α one knows how C obtains.

At least some types of knowledge are dimly luminous, namely, know-how and know-that.[8] The argument for that claim depends on two related views on

[6] Robert Stalnaker, "Luminosity and the KK Thesis," in *Externalism, Self-Knowledge, and Skepticism*, ed. Sanford C. Goldberg (Cambridge: Cambridge University Press, 2015), 21.

[7] Jason Stanley, *Know How* (Oxford University Press, 2011); Jason Stanley and Timothy Williamson, "Knowing How," *The Journal of Philosophy* 98, 8 (2001): 411–44; Berit Brogaard. "What Mary Did Yesterday: Reflections on Knowledge-Wh," *Philosophy and Phenomenological Research* 78, 2 (2009): 439–67.

[8] It is more controversial whether, if one knows a place or an object, one knows how one knows it. The same point applies to knowledge-wh (when, where, who). But if these types of knowledge could be described as knowledge-that, then they are dimly luminous. It is implausible, however, that we can reduce knowledge-wh to knowledge-that without committing knowledge-how to the same fate.

knowledge, the first of which is already present in Williamson's argument against bright luminosity:

> (Safety) if S knows that p, then in nearly all (if not all) nearby possible worlds in which S forms her belief that *p* in the same way as in the actual world, S only believes that *p* when *p* is true.[9]
>
> (Stability) if S knows how to φ, then in nearly all (if not all) nearby possible worlds in which S does φ in the same way as in the actual world, S only φes successfully.

Safety and Stability arise from the same concern, viz., that knowing is incompatible with getting it right by luck.[10] As the post-Gettier literature made abundantly clear, subjects in Gettier-style cases lack propositional knowledge because they could easily be mistaken. Hence, Safety is a necessary condition for knowledge. It may not be a sufficient condition – as shown by cases where the subject arrives at safe true beliefs because an epistemic angel changes the world to fit her beliefs.[11] However we attempt to explain cases like that – say, by adding credit for arriving at true beliefs or by taking cognitive achievement as a additional necessary conditions for knowledge – Safety remains a plausible necessary condition. In Williamson's original argument against luminosity, Safety is the basis for the claim that, if Mrs. Coldhot is feeling cold at any given time, then she is cold at the *next millisecond*, given that these instants are modally close.

Stability is simply one way to develop the same insight underlying safety when it comes to know-how. The basic idea is that knowing how to do something implies having an ability, or set of abilities, that, when properly exercised in

---

[9] The definition is due to Duncan Pritchard, *Epistemic Luck* (Oxford: Clarendon Press, 2005), 163, which is intended to deal with a number of counterexamples. I have omitted the mention of the contingency of the target-belief for simplicity, since it does not affect my point here. Also, safety is not without its worries: cf. Ram Neta and Guy Rohrbaugh, "Luminosity and the Safety of Knowledge," *Pacific Philosophical Quarterly* 85 (2004): 396–506; Avram Hiller and Ram Neta, "Safety and Epistemic Luck," *Synthese* 158, 3 (2007): 303–13. It is beyond the scope of this paper to save safety from these criticisms.

[10] I am not relying on any specific taxonomy of luck. Even though having a true belief through luck – what is sometimes called veritic luck, and can be divided into intervening or environmental luck - is clearly a case of getting it right by luck, not all lucky achievements are veritic. If a non-intellectualism about know-how of the kind assumed here is correct, the kind of luck that affects know-how is not veritic because there may not be beliefs involved in practical knowledge at all, a fortiori, there may not be beliefs which are true due to luck doing any negative epistemic work in practical cases of lucky achievement.

[11] Duncan Pritchard, "Safety-Based Epistemology: Whither Now?" *Journal of Philosophical Research* 34 (2009): 33–45.

sufficiently similar circumstances, guarantees success. A professional footballer that hits the perfect cross knows how to do so because she is able to perform similarly in sufficiently similar situations. A beginner might even, on occasion, hit the perfect cross as well, but she might fail to do so on many other occasions. The difference between the professional and the beginner is that the former is able, in an epistemically robust sense, to hit the perfect cross, whereas the later lacks the requisite abilities.[12] Analogously to Safety, Stability is intended to be a plausible necessary condition for know-how, even though other necessary conditions can be added. Note, moreover, that because having and exercising an ability is necessarily creditable to an individual, we do not need to add another condition to rule out cases that are analogous to that of epistemic angels that fix the world in order for it to coincide with one's beliefs. Also, my argument here works as long as believing that there is a correct way of φying, in a manner which is sufficient for knowledge, is not another necessary condition for φying successfully. We may sometimes, or even always, form corresponding beliefs that describe the correct or best ways of doing something, and we may even do that on a sufficiently strong epistemic basis, but that is completely incidental to whether we know how to do it.

Thus, knowing-how is being able to successfully do something in a stable manner. That might raise a familiar concern, namely: we are fallible creatures but that account left no room for error. We could, it seems, be able to φ and still fail systematically at φying. If a notion of ability cannot account for that apparent fact, then it is inherently flawed. The objection assumes that the only conceptual room for error lies between exercising an ability and getting it right. But that is false: we may fail because we failed to exercise the correct ability (or set of abilities) for doing something, despite our best intentions, or because we failed at exercising it according to what the situation demands. Millar[13] makes that same point regarding recognitional abilities, and it seems that his view also applies to the notion of ability in general. So the notion of ability developed here does make room for error.

The claim that knowing how to φ implies having the ability to φ is prima facie quite plausible, but it has been challenged. Consider the following case, which is adapted from Carr:[14]

---

[12] Ability possession is certainly a matter of degree - one can be more or less able to dosomething, but my argument is indifferent to that.

[13] Alan Millar, "How Visual Perception Yields Reason for Belief," *Philosophical Issues* 21, The Epistemology of Perception (2011): 332–51.

[14] David Carr, "Knowledge in Practice," *American Philosophical Quarterly* 18, 1 (1981): 53–61.

> *Arthritic Ana:*
>
> Ana is an arthritic piano teacher. Ana knows how to play the piano because she has played it her whole life, but since her rheumatoid condition worsened, she became unable to play. However, she can still teach her pupils on how to play any given piece. Ana knows how to play the piano, but she does not have the ability to do so.

The case seems to show that knowing-how does not imply possessing and exercising certain abilities. We can, however, resist that conclusion, for Arthritic Ana's case is underdescribed. Instead of ascribing her know-how, it is possible to construe her situation as follows: her expertise in piano playing enables her to describe the correct way to play a given piece in a very nuanced and precise manner, but that is a matter of a propositionally articulated epistemic status (either knowing-that and understanding-that). That explains how she can teach her pupils without implying that she knows how to play, because she is unable to do so. The point is that she has the relevant piano-playing abilities, but her actual condition is one in which she cannot exercise her abilities successfully across a sufficiently large class of nearby possible worlds, including the actual one. So her abilities are, as it were, pragmatically defeated, which in turn defeats her know-how. Remove the pragmatic defeater (her rheumatoid condition), and her practical knowledge is restored.

On the other hand, it seems we can imagine a case where one is able to φ without knowing how to φ. Consider the following case, also from Carr:[15]

> *Lucy the Lucky Gymnast:*
>
> Lucy is a beginner gymnast. After only a couple of classes, and without previous training, she attempts to do a challenging somersault, one that could only be achieved by a more proficient gymnast. Surpassing all expectations, she succeeds.

Lucky Lucy's case suggests that she is able to do a somersault – for, after all, she has done it – without knowing how to do so. We may resist that conclusion too, because there is an ambiguity at bay. We consider her able to do a somersault only in hindsight, and this is a weaker sense of "being able" than the one in play in Stability. Of course, in a sense, we are able to do many things that we are actually unable to do in another, more robust sense. I am able, in the weakest sense, to write an enthralling, best seller novel – I have a bunch of loose ideas and I know how to write. But I am not able to do so in a more robust sense: it would be a surprise for anyone if I were to write a best seller novel, myself included. In the relevant sense, being able to do something generates stable expectations of success,

[15] Carr, "Knowledge in Practice."

something that Lucy clearly lacks. The answer whose question shows whether one is able to φ is not "did he or she φed?," but "is he or she capable of φying again?"

Finally, here is how knowing that *p* satisfies (D-L): given Safety, if a person knows that *p*, she could not be easily mistaken about *p*. That is, in all, if not all, nearby possible worlds where she believes that *p* through the same means as in the actual world, *p* is true. Given Stability, if she knows how to arrive at knowledge that *p* in the actual world, then she successfully does so in all, if not all, nearby possible worlds where she attempts to do so through the same abilities as in the actual world. So the possible worlds where a person knows that *p* are the ones where she knows how to know that *p*.[16] Thus, for every case α, if in α one knows that *p*, the n in α one is in a position to know how one knows that *p*.

Knowing-how to know-that, in the sense I am advancing here, is being able to perform the same procedures, whatever they are, that generate first-order propositional knowledge in sufficiently similar circumstances. This view is neutral regarding which specific account of propositional knowledge is endorsed, as long as it makes room for Safety. Accounts that deny the necessity of ruling out certain cases of lucky achievement, such as Turri's,[17] do not satisfy Safety, and do not meet (prima facie at least) the requirements for dim luminosity of knowledge.

Moreover, knowing-how also trivially satisfies (D-L). Whenever one knows how to φ, one knows how one knows how to φ. This implies that knowing how one knows at an order *n* can be iterated in a *n+1* order. Because I am assuming a non-intellectualism about know-how, iteration is not an issue. It would be problematic only if it required the individual to form higher-order beliefs at every new step, with the sufficient epistemic support for knowledge; but this account of know-how does not implies believing, nor believing in an epistemically robust way. It implies only that the individual who has practical or procedural knowledge is able to perform in a sufficiently stable way in sufficiently similar circumstances. Accordingly, because knowing how to φ does not imply believing, in a sufficiently robust epistemic manner, that there is a correct way to φ, and because knowing-

[16] There might be possible worlds where she knows how to know that *p* but does not exercise the relevant abilities, so, in these worlds, she would not know that *p*. However, we could, in principle, advance the more contentious claims that the possible worlds where a person knows that *p* and the the ones where she knows how to know that *p* are coextensive. In order to do so, we would have to claim that not only knowing how to φ is a matter of being able to φ successfully in a sufficiently stable manner, but also that a person is only able to φ if she effectively φes. My case for dim luminosity does not rely on that claim. I thank Luis Rosa for that suggestion.

[17] John Turri, "Manifest Failure: The Gettier Problem Solved," *Philosopher's Imprint* 11, 8 (2011): 1–11.

that implies believing, it follows that dim luminosity does not imply bright luminosity. That is, if one is in a case where one knows that *p* – which implies, via dim luminosity, that one knows how to know that *p* – it does not follow that one knows that one knows that *p*, nor that one knows that one knows how to know that *p*.

One advantage of dim luminosity is that it secures the intuition shared among virtue reliabilists that propositional knowledge implies cognitive achievement.[18] According to these authors, if one knows that *p*, one has credit for arriving at the true belief that *p* – alternatively, the true belief that *p* is creditable to one's reliable epistemic dispositions. Thus, knowledge is a matter of achieving a true belief. Now, setting the notion of cognitive achievement at the center of one's conception of knowledge is notoriously tricky, for it seems to exclude testimonial knowledge, for it could be acquired effortlessly by the one receiving the testimony.[19] The view developed here is able to salvage the intuition regarding credit without incompatibility with testimonial knowledge, for credit lies at the second-order epistemic status of the knower. The idea is that if one comes to know that *p* because one acquires the belief that *p* on the basis of a testimony from a reliable source, then one could not easily belief falsely that *p* based on that source. That being the case, one knows how to acquire that same piece of knowledge in similar situations – for one is able to base one's belief that *p* on the testimony of that knowledge source. This solution depends on the plausible premise that being able to φ, in the sense discussed above, is a cognitive achievement.

3. Against Dim Luminosity

It seems that we can conceive of cases where an individual knows that *p* but does not know how he knows that *p*, so that knowledge is not even dimly luminous. Consider BonJour's Norman case.[20] Norman has clairvoyant powers that enable him to reliably track the president's actual location – say, the president is in a hotel in Moscow – despite the fact that all available evidences points towards him being

[18] See Wayne Riggs, "Reliability and the Value of Knowledge," *Philosophy and Phenomenological Research* 64, 1 (2002): 79–96; Ernest Sosa, *A Virtue Epistemology – Apt Belief and Reflective Knowledge*. Vol. 1, (Oxford, New York: Clarendon Press, 2007); John Greco, "Knowledge as Credit for True Belief," in *Intellectual Virtudes: Perspectives from Ethics and Epistemology*, eds. Michael DePaul and Linda Zagzbeski (Oxford: Oxford University Press, 2003). I am thankful to Gregory Gaboardi for this observation.

[19] Jennifer Lackey, "Why We Don't Deserve Credit for Everything We Know," *Synthese* 158, 3 (2007): 345–61.

[20] Laurence BonJour, "Externalists Theories of Empirical Knowledge," *Midwest Studies in Philosophy* 5 (1980): 53–73.

in Washington. Norman has no prior belief, in favor or against, his clairvoyance. The verdict is that there is something wrong with Norman's epistemic situation. Internalists, such as BonJour himself, claim that even though Norman satisfies an externalists condition, he lacks knowledge because he lacks reasons in favor of his belief regarding the president's actual location. Now, if we are going to assume that Norman's powers are capable of generating true beliefs in a safe manner, we may be tempted to say that he knows that *p* (where *p* is the relevant propositional about presidential whereabouts). But what is troubling is that there is a clear sense that Norman *does not know how* he knows that *p*.[21]

That case, however, can be explained away because clairvoyance is a bizarre epistemic ability. Once we consider whether Normans knows how he knows that *p*, matters become muddled, because it is not clear whether there are many, or any for that matter, nearby possible worlds in which Norman is able to arrive at his first-order knowledge in a similar manner. And this explains why we might be tempted to revoke his putative first-order knowledge. The problem, therefore, lies not with dim luminosity, but on the epistemic ability in question, for intuitions about it tend to colapse. Consider, in contrast, the less controversial ability of chicken-sexing. If the chicken-sexer is in indeed able to form safe beliefs about the subject matter, then there is no doubt she is able to achieve second-order procedural knowledge, even if she fails to meet an internalist condition.

Now, can we use Williamson's argument, which was directed at bright luminosity, to show that one can know that *p* without knowing how one knows? After all, knowing how to φ is a non-trivial state, so it seems that dim luminosity for knowledge fares no better than its brighter cousin. Take Mrs. Coldhot case again. If she knows that she is feeling cold at a time *t*, then she knows how to arrive at her first-order knowledge at *t*. Also, if she knows that she is feeling cold at *t*, then she is cold at *t+1*, given Safety (and factivity of knolwedge). Also, if she knows how to arrive at her first-order knowledge that she is feeling cold at *t*, then she knows that she is feeling cold at *t+1*. However, that would lead to the same contradiction as before, because if she iterates the same procedure throughout the morning, she is feeling cold at noon, contrary to one of our initial assumptions.

The problem of that argument lies in the premise that, if she knows how to arrive at her first-order knowledge that she is feeling cold at *t*, then she knows that she is feeling cold at *t+1*. Given that there is a close connection between knowing how to arrive at a knowledgeable belief and arriving at that belief, at the last state when Mrs. Coldhot feels cold, there are many nearby possible worlds where she no longer feels cold. If, in these possible worlds, she tries to arrive at a knowledgeable

[21]I am thankful to Gregory Gaboardi for these comments.

belief that she feels cold, but fails to do so (because she no longer feels cold), then she did not meet the Stability condition. And failing to do so, she does not know that she feels cold at *t+1*. The crucial point here is that know-how implies stable success, just like know-that implies safe true believing.

There seems to be some empirical evidence against the view advanced here and against the value and efficiency of reflection more generally.[22] In particular, consider the following experiment. Data gathered by Halberstadt and Wilson[23] shows that verbalized reasoning is objectively less reliable than what they call *affect heuristics*, viz., "domain-specific relationships between feeling states and objective states of the world."[24] Affect heuristics are cognitive procedures that use subjectively available cues in decision making, judgements, inferences and so on. One clear example of affect heuristics is that of assessing whether A is more familiar, for the individual, than B in order to establish some objective results (e.g.: whether A is more popular than B). Experiments mentioned by Halberstadt and Wilson also show that reasoners – individuals who try to explicitly explain how they arrive at certain predictions – often impair the objective quality of their judgements, whereas those that simply use affect heuristic are more likely to attain the correct results. For instance, in a study conducted by Halberstadt andLevine,[25] 71 subjects were asked to predict 8 basketball games in 1995 (the same experiment was replicated with 52 participants in 1996). Half of the experimenters were asked to jot down the reasons for their predictions, whereas the other half followed their instinct without analysing their reasons (the so called *nonreasoners*). Overall, nonreasoners got the right outcome for 70,4% of the matches, almost 5% more than reasoners. There is a lot going on in the case above, but the important point is this: both groups had the same knowledge basis for predicting the outcome of basketball games – but when reasoners tried to show how they know that one team was more likely to win than the other, they ended up affecting their first-order knowledge negatively. So it seems that not only knowing that *p* does not imply knowing how one knows, but also that attempting to show how one's first-order beliefs or knowledge are arrived at is epistemically hindering, albeit marginally so.

---

[22] For a battery of arguments against reflection, see Hilary Kornblith, *On Reflection* (Oxford: Oxford University Press, 2012).

[23] Jamin Halberstadt and Timothy Wilson, "Reflections on Conscious Reflection: Mechanisms of Impairment and Analysis," in *Reasoning: Studies of Human Interfaces and Its Foundations*, eds. Jonathan E. Adler and Lance J. Rips (Cambridge: Cambridge University Press, 2008), 548–65.

[24] Halberstadt and Wilson, "Reflections on Conscious Reflection," 554.

[25] Jamin Halberstadt and Gary Levine, "Effects on Reason Analysis on the Accuracy of Predicting Basketball Games," *Journal of Applied Social Pyschology* 29, 3 (1999): 517–30.

The issue could be avoided if we keep in mind that knowing how one knows is not reducible to knowing that one knows. We need to add to that the prima facie plausible premise that making one's knowledge explicit through reasons is essentially a matter of articulating pieces of propositional knowledge (or justification), which also involves exhibiting the correct abilities in reasoning. That is, it is plausible that reasoning requires both having reasons in support of one's beliefs and being able to use these reasons competently. The underlying view here is a "two-components" explication of rational capacities.[26] If reasoning involves both know-how and know-that, then it is possible that what is doing all the negative epistemic work for the reasoners, but not for the non-reasoners, is the formers' propositional second-order justification. That is, when one tries to make one's reasons explicit in order to articulate one's first-order knowledge pieces, one might fail to achieve propositional second-order knowledge. Clearly, more (empirical) explanatory work has to be done in order to show how knowing that one knows goes awry in cases like the above, but I am not going to pursue this matter here.

4. Conclusion

The intuition that knowing implies some sort of reflective attitude has persisted through much of the history of epistemology. One way to cash-out its details is through the infamous KK principle. Although recently that principle has been under attack, especially from the externalist camp, it seems we can rescue the original intuition by carefully distinguishing between knowing-that and knowing-how. These are conceptually distinct epistemic attitudes, but are understood minimally as implying Safety and Stability, thus they share a common core. Accordingly, whenever one knows that *p*, one knows how one knows that *p*. Knowledge is luminous, but only dimly so.

[26] Luis Rosa, "In order to be rational you need to know how to reason," *Philosophical Inquiries* 4, 1 (2016): 25–40.

QUINE AND THE INCOHERENCE OF THE INDISPENSABILITY ARGUMENT

Michael J. SHAFFER

ABSTRACT: It is an under-appreciated fact that Quine's rejection of the analytic/synthetic distinction—when coupled with some other plausible and related views—implies that there are serious difficulties in demarcating empirical theories from pure mathematical theories within the Quinean framework. This is a serious problem because there seems to be a principled difference between the two disciplines that cannot apparently be captured in the orthodox Quienan framework. For the purpose of simplicity let us call this *Quine's problem of demarcation*. In this paper this problem will be articulated and it will be shown that the typical sorts of responses to this problem are all unworkable within the Quinean framework. It will then be shown that the lack of resources to solve this problem within the Quinean framework implies that Quine's version of the indispensability argument cannot get off the ground, for it presupposes the possibility of making such a distinction.

KEYWORDS: mathematics, indispensability, knowledge, holism, confirmation, ontology

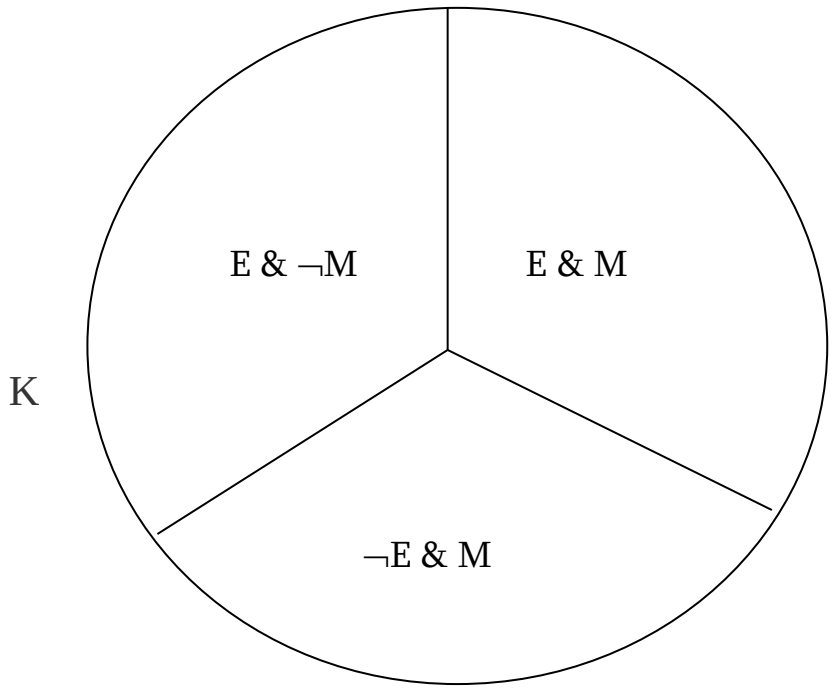
1. Introduction

It is an under-appreciated fact that Quine's rejection of the analytic/synthetic distinction—when coupled with some other plausible and related views—implies that there are serious difficulties in demarcating empirical theories from pure mathematical theories within the Quinean framework.[1] This is a serious problem because there seems to be a principled difference between the two disciplines that cannot apparently be captured in the orthodox Quinean framework. For the purpose of simplicity let us call this *Quine's problem of demarcation*. In this paper this problem will be articulated and it will be shown that the typical sorts of responses to this problem are all unworkable within the Quinean framework. It will then be shown that the lack of resources to solve this problem within the Quinean framework undermines Quine's version of the indispensability argument, for it presupposes the possibility of such a distinction.

[1] See W.V.O. Quine "Two Dogmas of Empiricism," *Philosophical Review* 60 (1951): 20-43.

2. Quine and the Problem of Demarcating Science and Mathematics

Quine is duly famous for his critique of the analytic/synthetic distinction in his 1951 article "Two Dogmas of Empiricism," despite persistent disagreement about the significance of this work. Nevertheless, given Quine's criticism of the analytic/synthetic distinction, Quine (and those who follow in his naturalistic footsteps) regard *all* statements as being empirical in character in some important sense. Thus, every statement is supposed to be subject to revision in light of empirical evidence. Of course, this is not news to us today, but what this implies about teasing apart pure mathematics from empirical theory has not been properly appreciated. What it immediately and most obviously implies is that Quineans cannot discriminate pure mathematics from empirical science by asserting that the propositions that make up mathematical theories are analytic, whereas the propositions that make up empirical theories are synthetic. This is troubling because in practice there appears to be a quite sharp distinction between the practices of mathematics and the empirical sciences. Presumably we would then like to be able to partition the complete body of known statements K into M—the mathematical statements—and E—the empirical statements, such that $E \subset K$ and $M \subset K$. But, in order to accomplish this task, a plausible criterion C that grounds the distinction between the elements of E and M must exist. This figure captures the necessary distinction:



Here, according to the Quinean view, we are *only* to adopt ontological commitment to those claims that lay within the E regions of K. What we then need, in the spirit of Quine's "no entity without identity" dictum, is a criterion of identity that could demarcate mathematical statements from empirical statements (and thereby also a related criterion to demarcate the statements of pure mathematics from those of applied mathematics). Without such an identity criterion, Quineans would violate their own ontological scruples. Quine's criticism of the analytic/synthetic distinction immediately shows that Quineans cannot use analyticity as the criterion for demarcating mathematical statements from empirical statements in K and so we must look elsewhere if we are to solve the Quinean problem of demarcation

This inability to discriminate pure mathematics from empirical science on the basis of the analytic/synthetic distinction in the Quinean framework is then exacerbated further when it is also recognized that traditional methodological accounts of pure mathematics are wildly unrealistic and depend on the viability of a workable notion of a priority. Those accounts typically treat the methodology of pure mathematics as the development of necessarily true axiomatic systems where theorems are proved on the basis of the axioms by the use of a priori methods, but this is simply not true of the actual practice of mathematics as numerous philosophers of mathematics have now come to realize. Lakatos in particular is largely responsible for this recognition.[2]

More importantly, Quine's rejection of the analytic/synthetic distinction also simultaneously eliminates the viability of using either the a priori/a posteriori distinction or the necessary/contingent distinction as the basis for demarcating empirical science from pure mathematics because, at least for Quineans, as a result of the collapse of the analytic/synthetic distinction, there is no a priori knowledge and there are no necessarily true propositions. Quine held that the class of analytic truths is the just the class of a priori knowable truths and the class of a priori knowable truth is just the class of necessary truths. Similarly, the class of synthetic truths is just the class of a posteriori knowable truths and the class of a posteriori knowable truths is just the class of contingent truths. But according to Quine there

[2] See Imre Lakatos, "Proofs and Refutations," *British Journal for the Philosophy of Science* 14 (1963-4): 1-25, 120-139, 221-243, 296, 342, Imre Lakatos, *Proofs and Refutations* (Cambridge, MA: Cambridge University Press, 1976) and Imre Lakatos, "A Renaissance of Empiricism in the Recent Philosophy of Mathematics," *British Journal for the Philosophy of Science* 27 (1976): 201-223.

are no analytic truths and so there are no truths that are knowable a priori and there are no necessary truths. So, as we saw earlier Quineans cannot use the analytic/synthetic distinction to do the work of C and for the same basic reason they can use neither the a priori/a posteriori distinction nor the necessary/contingent distinction to do the work of C. Moreover, they cannot use any criterion that employs any of these conceptual distinctions, for there are no such distinctions according to Quine.[3]

Yet more troubling still is the recognition that Quine's rejection of the analytic/synthetic distinction also implies a radical sort of confirmational holism, whereby the whole of our system of beliefs is the proper unit of confirmation relative to observational data.[4] In other words the whole body of our beliefs is at issue when it comes to the issue of confirmation. This means that when we consider the acceptability of our beliefs it must be done in a global manner and when our system of beliefs conflicts with observational data we are then confronted with the infamous Quine/Duhem thesis.[5] This is the assertion that

[3] Quine makes an alternative and more mature attempt to specify C later on in W. V. O. Quine, *From Stimulus to Science* (Cambridge, MA: Harvard University Press, 1995), 52-57. There he argues that the distinction between empirical statements and those of pure mathematics can perhaps be established as follows. Empirical statements and sets of empirical statements imply observation categoricals, whereas the statements of pure mathematics individually and jointly do not. This suggestion is variously problematic as follows. First, this approach fails to specify a defining feature of specifically mathematical statements, as opposed to other non-empirical statements. Moreover, it makes the statements of pure mathematics meaningless and devoid of truth values given his adherence to the view that all semantic content is ultimately grounded empirically. Finally, semantic content is typically generated only by sets of statements that must involve at least some empirical claims, but in Quine's system there is no way in practice tease apart which statements in a testable set with semantic content are specifically conferring the empirical content on that set. This is due to his subscription to a form of semantic holism. We cannot in most cases selectively and sequentially delete statements and then check to see if semantic content remains and this is simply because critical semantic mass is not a property of single sentences (*From Stimulus to Science*, 48-49). Elsewhere, he famously tells us that "It is misleading to speak of the empirical content of an individual statement ("Two Dogmas," 43)," and also that "the unit of empirical significance is the whole of science ("Two Dogmas," 42)." In fact, Quine admits all of these charges (*From Stimulus to Science*, 55-57) and so even he acknowledges that this approach is problematic.

[4] In Quine's later work this more extreme view is relaxed and Quine holds that "large" chunks of our systems of belief are the units of confirmation. See W.V.O. Quine, "Five Milestones of Empiricism," (1975) reprinted in *Theories and Things* (Cambridge: Harvard University Press, 1981).

[5] See W.V.O. Quine, *Philosophy of Logic* (Englewood, NJ: Prentice-Hall, 1970) and W.V.O. Quine and Joseph Ullian, *The Web of Belief* (New York: Random House, 1970).

when our holistic system of beliefs is faced with empirical falsification, we must give up something to restore consistency, but that there is nothing *in particular* we must give up. We can give up the observation statement itself or one or more theoretical beliefs that give rise to the contradiction. As a result, when our system of beliefs is faced with falsification we must give *something* up, but we can typically restore consistency in a number of ways by adjusting our beliefs. This view then also undermines the principled possibility of using the concept of revisability to make the distinction between the statements of mathematics and those of science.

What is the of great interest is that in discussing the role of mathematics and its ontology in the context of its application in empirical theory, Quine—along with Putnam—subscribed to the infamous indispensability arguments.[6] Such indispensability arguments take the following generic form:

> P1: We ought to have ontological commitments to all and only entities that are indispensible to our best scientific theories.
>
> P2: Mathematical entities are indispensible to our best scientific theories.
>
> Therefore, we ought to have ontological commitment to mathematical entities.

These arguments essentially conclude that our ontological commitments to the existence of mathematical entities should be on a par with our ontological commitments to the theoretical entities appealed to in empirical theories because indispensable mathematical propositions employed in such theorizing accrue confirmation holistically when the empirical propositions in question are confirmed. That is to say, all of the propositions used in some given empirical endeavor—including those that are a part of mathematics—accrue confirmation *jointly* because no proposition can be confirmed in isolation and mathematics is in some crucially important sense indispensable to the conduct of the work of the empirical sciences.

But this also means that the difference between the statements of pure mathematics and the statements of the empirical sciences also cannot be grounded

---

[6] For the origin of the indispensability arguments see W.V.O. Quine, "On What There Is," *Review of Metaphysics* 2 (1948): 21-38, "Two Dogmas," *Word and Object* (New York, MIT Press, 1960), "Carnap and Logical Truth," *Synthese* 12 (1960): 350-374, *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969), *Theories and Things* (Cambridge: Harvard University Press, 1981), Hilary Putnam, *Philosophy of Logic* (New York: Allen and Unwin, 1972), *Mathematics, Matter and Method*, 2nd ed. (Cambridge, MA: Cambridge University Press, 1979). For discussion of these arguments see Mark Colyvan, *The Indispensability of Mathematics* (Oxford: Oxford University Press, 2003) and Penelope Maddy, *Realism in Mathematics* (Oxford: Clarendon Press, 1990).

in the differing ontological attitudes towards the referents of mathematical statements and the referents of empirical statements, for what we should take to exist is just what our best total theory of the world indicates as existing.[7] So given the Quinean stance on ontology the criterion C cannot be difference in ontological attitude, for there are no such differences. But, commitment to the indispensability arguments strongly indicates that Quine and those who follow his lead must identify some principled manner by which mathematical propositions can be distinguished from other propositions. Otherwise, the indispensability arguments verge on being nonsensical, for their very formulation depends on such a distinction. In the indispensability argument both P1 and P2 *assume* such a distinction. In order to make the indispensability argument work we must already have in hand a viable criterion C by which we can distinguish statements involving mathematical entities from statements of empirical science. Otherwise we would have no way to establish P2, i.e. that some *specifically* mathematical entities are indispensable to any scientific theory. This is because we would have no way to determine which statements concern mathematical entities and it might turn out that unbeknownst to us there *are no mathematical entities involved in science at all*. If that were true, then the indispensability argument would just be pointless. Similarly, absent some criterion to ground the principled difference between mathematical and scientific statements we would have no way to determine that we ought to be ontologically committed to. This is because we could not follow P1—the claim that the only entities that we should be committed to are those that are indispensable only to our scientific theories—in practice if we do not know what count as specifically scientific claims and what count as statements of pure mathematics.

What all of this ultimately appears to imply about discriminating pure mathematics from empirical science within the Quinean framework should now be apparent and it is deeply troubling. Namely, there is no obvious way to ground the clearly real and important distinction between the empirical sciences and pure mathematics in light of orthodox Quinean principles. Quine explicitly recognizes in his 1951 article that the statements of pure mathematics are as revisable as those of any empirical science, but does not appear to see how deeply that this threatens the very *distinction* between the two types of statements and between the two disciplines. What he does not recognize is that his views, when taken together, threaten the identity conditions for mathematical and empirical statements. Quine himself seems only to have recognized that at the superficial level the difference between the mathematical elements of our belief systems and the more properly

[7] See Quine, "On What There Is."

empirical elements is one of degree of entrenchment in our belief system. More specifically, typically then the propositions of mathematics are supposed to be more deeply entrenched than the propositions of empirical theories. But, that is just to say that when our belief system faces falsification we are (typically) *psychologically* less inclined to give up the propositions of pure mathematics than we are to give up observation statements or more properly empirical statements of the empirical sciences. But, this is nothing more than a difference in psychological attitudes towards those types of statements which (1) is itself totally ungrounded on a theoretical level—although it does serve to explain the "felt" necessity of mathematics, and (2) which may vary across individuals. Nevertheless, this simply and directly implies that there is no obvious, principled and theoretically grounded criterion C that can be used to make distinction between the statements of pure mathematics and the statements of the empirical sciences within the Quinean corpus. So, for Quineans it appears to be the case that there really is no identifiable principled difference between science and pure mathematics. But, this conclusion is wildly implausible given the practice of both mathematics and the empirical sciences and it renders the indispensability argument incoherent. Moreover, if Quineans simply fall back to the view that conformational holism is ultimately the source of support for all statements in the web of belief—including those of mathematics—and thus fixes ontological commitment, it would simply be a concession that the indispensability argument is simply irrelevant because that would require conceding P1. Doing so would also presumably entail the kind of indiscriminate and comprehensive Platonism that is at odds with the conservative nature of Quine's attitude towards the existence of mathematical entities in particular and abstract objects in general as forcefully argued for in his 1960a and which is clearly precluded by P1. Quine wants us to commit *only* to the existence of those mathematical entities that are in fact indispensable to science and not indiscriminately to the existence of all mathematical entities. So, especially given the "no entity without identity" principle, Quineans need to acknowledge the problem and to provide a principled manner in which pure mathematics can be distinguished from empirical science that retains all—or at least most—of the basic Quinean views and which allows us to makes sense of the indispensability arguments. However, it is not at all clear how this might be accomplished while retaining the core of Quine's argument against the viability of the analytic/ synthetic distinction. So, Quine's views may simply be inconsistent when it comes to this particular issue.

DISCUSSION NOTES/DEBATE

GETTIERED BELIEFS ARE GENUINE BELIEFS: A REPLY TO GAULTIER AND BIRO

Gábor FORRAI

ABSTRACT. In recent articles in this journal Benoit Gaultier and John Biro have argued that the original Gettier cases and the ones closely modelled on them fail, and the reason for the failure is that the subject in these cases does not actually have the belief that would serve as a counterexample to the justified-true-belief analysis of knowledge. They claim that if our evidence pertains to a particular individual (as in the first case) or to the truth of one of the disjuncts (as in the second case), we do not genuinely believe the existential generalization or the disjunction which logically follows. I will challenge their arguments and suggest that our unwillingness to assert the existential generalization or the disjunction under these conditions does not stem from lack of belief but from pragmatic principles.

KEYWORDS: belief, Benoit Gaultier, Gettier cases, Paul Grice, Horn scales, John Biro

In recent articles in this journal Benoit Gaultier[1] and John Biro[2] propose a new way of doing away with Gettier's original counterexamples to the justified-true-belief analysis of knowledge and the counterexamples closely modelled on them, like Lehrer's Nogot/Havit case.[3] In these counterexamples the subject has a justified belief, from which he infers by existential generalization or by addition to another proposition and comes to believe it. It is stipulated that the belief he starts out from is false, but the belief he arrives at by inference happens to be true. Gettier claims that the belief the subject arrives at – which, following Gaultier, I will call Gettiered belief – satisfies the justified-true-belief analysis of knowledge but does not constitute knowledge.[4]

Gaultier and Biro respond by denying that the Gettiered belief is a genuine belief. Gaultier says that attributing the subject a Gettiered belief is just a vague

[1] Benoit Gaultier, "An Argument Against the Possibility of Gettiered Beliefs," *Logos & Episteme* V, 3 (2014): 265-272. Hereafter: Gaultier.

[2] John Biro, "Non-Pickwickian Belief and 'the Gettier Problem'," *Logos & Episteme* VIII, 3 (2017): 47-69. Hereafter: Biro.

[3] Keith Lehrer, "Knowledge, Truth and Evidence," *Analysis* 25 (1965): 168-75.

[4] Edmund Gettier, "Is Justified True Belief Knowledge?" *Analysis* 23 (1963): 121-123.

and imprecise characterization of what he believes;[5] Biro says that it is a belief only in a pickwickian sense rather than a serious belief.[6] Even though they argue differently,[7] their diagnosis of what goes wrong is the same: if one has evidence only for a singular proposition or one of the disjuncts, and has *no independent evidence* for the existential generalization or the inclusive disjunction which follows from it, one cannot form a genuine belief in the existential generalization or the inclusive disjunction.

Gaultier articulates this point by putting forward a general claim called (THESIS), which he explains as follows:

> … when, and *only when*, the *only* evidence one has in favour of something weaker than p – namely, p^* – is the evidence one has in favour of p and that led one to believe that p, we cannot be in presence of two different beliefs – the belief that p and the belief that p^* – but only in presence of one single belief, the belief that p.[8]

Since the existential generalization/disjunction is weaker than the singular proposition/one of the disjuncts, if we only have evidence for the latter but possess no independent evidence for the former, (THESIS) entails that we do not believe the former.

Biro develops the same point by assuming "that believing something in a non-pickwickian sense means being prepared to assert it seriously,"[9] and then proposing conditions under which existential generalization and disjunction can be seriously asserted.[10] The condition for the serious assertion of an existential generalization is that we should be also prepared to assert that if one particular individual does not have the property in question, another individual does.[11] In case of the disjunction the condition is that we should be also prepared to assert that if one of the disjuncts is false, the other one is true.[12] As these conditions are not satisfied if we only have evidence about a particular individual or the truth of one of the disjuncts, they, in effect, require possession of independent evidence for the existential generalization or the disjunction.

---

[5] Gaultier, 270.

[6] Biro, 53, 67.

[7] Biro does not approve of Gaultier's solution, even though he thinks it is on the right track, 67-68.

[8] Gaultier, 270.

[9] Biro, 53.

[10] I do not discuss Biro's special objection to Gettier's original first example on p. 52, because it does not generalize to all examples using existential generalization, like the Nogot/Havit case.

[11] Biro, 56.

[12] Biro, 54.

Let us now see how this idea can be used to refute the Nogot/Havit counterexample, which Gaultier and Biro both analyze. Suppose I have excellent evidence to believe that Nogot, who is in my office, owns a Ford: I saw him getting out of a Ford, he told me he had just purchased it, showed me the certificate, and I also know him to be honest and reliable. From my belief it follows by existential generalization that someone in my office owns a Ford; that is the Gettiered belief. Now suppose that Nogot lied, the certificate is forged, the car he got out of is not his own, and also suppose that someone else in my office, Havit, does happen to own a Ford. The Gettiered belief would then be both justified and true and thus a counterexample to the justified-true-belief account of knowledge. However, in this situation the only evidence we have is what supports that Nogot owns a Ford, and we have *no independent evidence* suggesting that someone in my office owns a Ford. (Such evidence could consist, for instance, in regularly seeing a Ford in the parking lot reserved for those working in the office.) If genuine belief in the existential generalization requires independent evidence, it follows that the Gettiered belief is not a genuine belief.

Gaultier and Biro lay out this point as follows. Gaultier says that since the only evidence I have in favor of the weaker proposition "Someone in my office owns a Ford" is the evidence I have in favor of the stronger proposition "Nogot, who is in my office, owns a Ford," I cannot have two beliefs, and I believe only the stronger proposition. Biro says that in these circumstances I would not be prepared to assert "If Nogot does not own a Ford, someone else in the office does," and since being prepared to assert that is required for the serious assertion of "Someone in my office owns a Ford," I cannot seriously assert the latter. What we are not willing to assert seriously we do not genuinely believe, therefore, I do not genuinely believe that someone in my office owns a Ford.

Let us now see how they argue for this kind of solution. Gaultier first offers a general argument for his (THESIS):

> …it seems plausible to claim that evidence directly constrains belief – more specifically, that one's beliefs formed at *t* directly inherit their content from the evidence one judges at *t* to have for them. It even seems that this *has to be so*, because if evidence constrained one's beliefs only through such epistemic aims or norms [truth, knowledge, or justification], one would always believe something as weak as possible on the basis of the evidence one has, in order to satisfy these aims or norms – which is clearly not the case.[13]

I do not find this convincing. First, I am not sure what to make of the claim that one's beliefs directly inherit their content from what one takes to be the

[13] Gaultier, 268.

evidence one for them. Suppose one police officer believes someone to be the murderer because he found the blood of the victim on his clothes and his fingerprints on the murder weapon and another police officer believes the same person to be the murderer based on eyewitness report. Would their beliefs have different content? Or suppose my neighbor asks me to look after his cat while he is away and says he would compensate me for taking the trouble, and I come to believe through wishful thinking that by 'compensation' he meant a huge amount of money. What do I take to be the evidence for my belief and how does it constrain the content of my belief?

Second, it seems wrong that pursuing such epistemic aims such as truth, knowledge and justification[14] would make us adopt the weakest possible beliefs the evidence allows. We are not concerned here with cases in which we lack sufficient evidence, but with cases in which the evidence supports both a stronger claim and weaker claim – e.g. "Nogot, who is in my office, owns a Ford" and "Someone in my office owns a Ford." If we come to believe the weaker claims only, we end up believing fewer truths, having less knowledge and having fewer justified beliefs. It is not the pursuit of truth, knowledge or justification that may keep us from accepting the stronger claim, but the aim of avoiding falsehood: by believing the stronger claims as well we run a greater risk of having false beliefs. However, even if avoidance of falsehood is one of our epistemic aims, the pursuit of truth, knowledge and justification still advises us to accept the risk. So we do not need Gaultier's (THESIS) to prompt us to accept the stronger but justified claims.

Gaultier offers a second reason as well.[15] Let us take the Nogot scenario without Havit, i.e. let us suppose that no one in my office owns a Ford. If it turns out that the evidence is misleading, I will be surprised to learn that Nogot does not own a Ford, but, as Gaultier correctly observes, I would not feel surprised once again if I also learned that no one in my office does. Gaultier suggests that the simplest explanation of why I would not be doubly surprised is that I did not have the Gettiered belief that someone in my office owns a Ford.

There is, however, another explanation which is just as simple and goes like this. When a belief is justified by a single inference, and one of the premises proves false, we give up the belief. If we are informed later that what we used to believe on the basis of the inference is false, we cannot be surprised, because we no longer believe that. Here is an example. I know that a car-loving colleague has recently bought a Porsche and is quite crazy about it. One day I drive by his house and see a Porsche just like his, which is badly smashed. Arriving at work next morning I do

[14] Gaultier, 267.

[15] Gaultier, 268-9.

not see the Porsche in the parking lot, and when I enter the office, I hear the colleague talking on the phone about a car insurance claim. As a result, I come to believe that his Porsche got crashed. Knowing his temperament and how much he adored his car, I infer that he must be pretty upset. I ask another colleague, who tells me that the reason the Porsche is not in the parking lot and the owner is having a conversation about an insurance claim is that his wife's car got crashed, and he had to lend the Porsche to her. Hearing this, I will no longer believe that the Porsche owner is upset, so I will not be surprised when I find out that he is not. This explanation applies to the Nogot scenario as well, in which the Gettiered belief is also based on a single inference the premise of which turns out false. This renders Gaultier's explanation superfluous: we can explain the lack of double surprise by saying that the inferentially justified belief is abandoned when the premise proves false, hence we do not have to suppose the subject never had that belief.

Biro does not provide a general argument either for his view that genuine belief is marked by serious assertion or that the conditions for serious assertion of existential generalization and disjunction are the ones he suggests, but he offers two parallel arguments. The first is this:

> Suppose Poirot says "someone in this room is the murderer" because he believes that the nephew killed the uncle. On subsequently discovering that the nephew has a cast-iron alibi and it was the butler, also present, who committed the dastardly deed, we would not allow Poirot to get away with saying (not that he would), "I was right all along!"[16]

This is clearly right, and it indeed attests to Biro's solution supposing the reason we find Poirot's reaction disingenuous is that he did not really believe "someone in this room is the murderer" but believed only that "the nephew in the room is the murderer." But there is another, more general explanation of why we do not allow Poirot to take credit for his true belief: we do not allow people to take credit for true beliefs if they are based on wrong reasons. To see this, let me revert to the case of the Porsche lover. Suppose that the Porsche lover finishes the telephone conversation about the insurance claim and is pretty upset, because the insurance company would not pay for the damage to his wife's car. I cannot turn to the colleague who told me that there is nothing wrong with the Porsche and say "See, I was right all along: he *is* upset," because my belief that the Porsche lover is upset was based on the false reason that his Porsche had been crashed.

The second parallel argument is this:

[16] Biro, 67. Gaultier also gestures toward an argument like this, 266.

> A mark of seriousness in a belief is that it guides action. If I am in the market for a used Ford and believe that Havit owns the one in the parking lot, it would not be rational for me to go around asking who owns it, as it would be if what I seriously believed was that someone or other in the building did.[17]

Believing that "Someone or other in the building owns the Ford" amounts believing "Someone in the building owns the Ford" and not believing anything concerning who that person might be, and under these conditions it makes sense to ask who owns it. But the reason why we do not ask this when we believe that the Ford belongs to Havit is not that we do not believe that someone in the building owns a Ford but that we *also* believe that Havit owns it. Actions are guided by constellations of beliefs and not a single belief.[18] Here is an example. Suppose I want to buy a used Ford and believe Havit's Ford is up for sale. It would then be perfectly rational to talk to him about buying it. However, if I *also* believe that Havit would not sell me his car for twice the market price because he hates my guts, I will not talk to him. The reason I do not talk him is not that I do not seriously believe that his car is up for sale but that I also believe something else.

As a final note let me indicate what the correct observation is that underlies Biro's solution and possibly also Gaultier's: we normally *assert* the existential generalization – the weaker claim as Gaultier would put it – if and only if the evidence available supports the existential generalization but does not support any particular singular proposition implying it, i.e. if only if we have *independent evidence* for the existential generalization *but nothing else*. If, on the other hand, we have evidence that a certain individual has the property in question, it is the singular proposition we will assert. So if the only evidence suggesting that someone in my office owns a Ford is that we regularly see a Ford in the parking lot reserved for those working in the office, we will assert the existential generalization. However, if the evidence suggests that a particular person in the office owns a Ford, we will assert the singular proposition instead. The same consideration applies *mutatis mutandis* to disjunction: we will normally assert it if and only if our evidence supports the disjunction but does not support either of the disjuncts; if the evidence supports one of the disjuncts, it is that disjunct we will assert. From our unwillingness to assert the existential generalization/disjunction when and only when we only have evidence pertaining to a particular individual/one of the disjuncts Biro – and possibly also Gaultier – infers that we do not genuinely believe the existential generalization/disjunction under these conditions. But that is

[17] Biro, 68.

[18] Even if in explaining actions we only mention the most salient belief, because we trust that the audience can figure out the rest.

wrong: our willingness is explained by pragmatic considerations rather than lack of belief.

This kind of phenomenon was first noted by Grice, in whose example A and B are discussing whether to visit C while in France. If A asks "Where does C live?" and B responds "Somewhere in the South of France," this implicates that B cannot specify exactly where; the reason why B violates the maxim that one should provide as much information as needed is that he cannot do so.[19] This observation spawned a special field of research in pragmatics which is called *scalar implicatures*. The name comes from the classic treatment by Laurence Horn, which relied on what are now called Horn scales.[20] Horn scales are groups of expressions arranged in order of decreasing informativity, like:

<all, most, many, some, few >

<excellent, good >

<hot, warm>

<always, often, sometimes>

<and, or>[21]

Horn regards it a general rule (subject to exceptions in special contexts) that employing a weaker term implicates that the speaker does not believe that using the more informative term is correct. Thus saying "I have read some of the papers" implicates that I have not read all of them. Saying this when I have read all of the papers would be pragmatically inappropriate, because it would violate the pragmatic principle "Say as much as you can" and would give rise to the false implication that I have not read all of them.

Taking our cue from pragmatics, the reason why we do not assert the existential generalization/disjunction when we possess evidence about a particular individual having the property in question/the truth of one of disjuncts is that pragmatic principles demand us to provide as much information as we can. So the reason why we do not assert the weaker claim is not that we do not believe it but that it would be pragmatically inappropriate to do so: we would thereby violate the

[19] Paul Grice, *Studies in the Way of Words* (Cambridge, Mass. and London: Harvard University Press, 1989), 32-33.

[20] Laurence R. Horn, *On the Semantic Properties of Logical Operators in English* (Ph.D. thesis, UCLA, Los Angeles, 1972). For a critical exposition and more current literature see: Bart Geurts, *Quantity Implicatures* (Cambridge: Cambridge University Press, 2010), 50-66.

[21] The examples are taken from Stephen C. Levinson, *Pragmatics* (Cambridge: Cambridge University Press, 1983), 134.

pragmatic principle and generate false implications.[22] "Someone in my office owns a Ford" is indeed be a strange thing to say if I believe that it is Nogot, but it is not a strange thing to believe.

[22] This consideration also reveals what underlies Biro's argument for his claim that disjunction as applied to serious belief is always exclusive: when we believe both disjuncts, we should use "and," 54. "And" and "or" also constitute a Horn scale. If we say "or" this implicates that we do not believe "and," which makes the disjunction exclusive. Laurence R. Horn, "Implicature" in the *Handbook of Pragmatics*, eds. Laurence S. Horn and Gregory Ward (Oxford: Blackwell, 2006), 9.

YOU CAN'T HANDLE THE TRUTH: KNOWLEDGE = EPISTEMIC CERTAINTY

Moti MIZRAHI

ABSTRACT: In this discussion note, I put forth an argument from the factivity of knowledge for the conclusion that knowledge is *epistemic* certainty. If this argument is sound, then epistemologists who think that knowledge is factive are thereby also committed to the view that knowledge is *epistemic* certainty.

KEYWORDS: epistemic certainty, factivity, fallibilism, knowledge

It is a commonly held view among contemporary epistemologists that knowledge is factive.[1] To say that knowledge is factive is to say that, if *S* knows that *p*, then *p* is true.[2] In other words, if *S* knows that *p*, then *p* cannot be false. But if *p* cannot be false, then *p* is *epistemically* certain, i.e., certainty as an epistemic property of propositions rather than a property of subjects.[3] For to say that *p* is epistemically certain is to say that *p* is guaranteed to be true; it cannot be false.[4] As Klein puts it, "if e makes p certain, the *truth* of p must be guaranteed by e" (emphasis in original).[5] Therefore, if *S* knows that *p* on the grounds that *e*, then *e* makes *p* epistemically certain.

Accordingly, the argument from the factivity of knowledge for the conclusion that knowledge is epistemic certainty runs as follows:

1) If *S* knows that *p* on the grounds that *e*, then *p* cannot be false given *e*.
2) If *p* cannot be false given *e*, then *e* makes *p* epistemically certain.
3) Therefore, if *S* knows that *p* on the grounds that *e*, then *e* makes *p* epistemically certain.

---

[1] See, for example, Timothy Williamson, *Knowledge and Its Limits* (New York: Oxford University Press, 2000), 131.

[2] Jennifer Nagel, "Knowledge as a Mental State," in *Oxford Studies in Epistemology Vol. 4*, eds. T. Szabó Gendler and J. Hawthorne, 273-308 (Oxford: Oxford University Press, 2013), 277.

[3] Baron Reed, "Certainty," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta (Winter 2011 Edition, 2011). https://plato.stanford.edu/archives/win2011/entries/certainty.

[4] Roderick Firth, "The Anatomy of Certainty," *The Philosophical Review* 76 (1967): 3-27.

[5] P.D. Klein, *Certainty: A Refutation of Scepticism* (Minneapolis: University of Minnesota Press, 1981), 185.

If hypothetical syllogism is a valid logical form, then this argument is deductively valid. Premise (1) is simply a statement of the thesis that knowledge is factive, which contemporary epistemologists generally accept. Premise (2) is true by definition, since to say that a proposition is *epistemically* certain (again, not to be confused with psychological certainty, which is often cached out in terms of indubitability[6]) is to say that it cannot be false; a proposition that is epistemically certain is guaranteed to be true.[7] Now, "if e makes p certain, the *truth* of p must be guaranteed by e,"[8] and *S* knows that *p* on the grounds that *e*, then *e* makes *p* epistemically certain. So it appears that this argument is sound.

It might be objected that there is a scope ambiguity in this argument. That is, the thesis that knowledge is factive is a wide-scope claim, i.e., $\Box(K_p \rightarrow p)$, whereas the thesis about epistemic certainty is a narrow-scope claim, i.e., $\Box p \rightarrow EC_p$. But this is mistaken, for what makes *p* epistemically certain is the evidence *e*, not that *p* is a necessary truth. In other words, a proposition is epistemically certain given the evidence for it, whereas a proposition can be necessarily true independently of the evidence for it. Now, if *e* is such that *p* cannot be false given *e*, then *p* is epistemically certain given *e*. And if *S* knows that *p* on the grounds that *e*, then *p* is epistemically certain given *e*. That is:

1) $K_{p|e} \rightarrow \Box p|e$

2) $\Box p|e \rightarrow EC_{p|e}$

3) $K_{p|e} \rightarrow EC_{p|e}$

Some might also object to this argument by saying that, if it is sound, then it means that we know very little (or maybe even nothing at all) because few of our beliefs (if any) are epistemically certain.

To this objection, "so what?" seems to be an appropriate response. The fact that a truth is difficult for people to accept is *not* evidence against it. Some religious believers find it difficult to accept the theory of evolution by natural selection, since they think that the theory is inconsistent with their religious beliefs. But the mere fact that those religious believers find it difficult to accept the theory of evolution by natural selection is *not* evidence against the theory itself. Similarly, if the thesis that knowledge is epistemic certainty has consequences that we find difficult to accept, then that fact alone does not count as evidence against the thesis that knowledge is epistemic certainty.

---

[6] Cf. Ludwig Wittgenstein, *On Certainty*, eds. G.E.M. Anscombe and G.H. von Wright (New York: Harper & Row, 1969).

[7] Reed, "Certainty," https://plato.stanford.edu/archives/win2011/entries/certainty.

[8] Klein, *Certainty*, 185.

Perhaps with few exceptions, such as Fred Dretske,[9] who argued that knowledge requires conclusive justification,[10] most contemporary epistemologists want to allow for knowledge even in cases where the justification is less than conclusive in order to block skeptical arguments. On this view, fallible knowledge of *p* requires sufficient evidence for *p* (how much is sufficient, exactly?), albeit not the sort of evidence that entails *p*. For instance, one can have fallible – but not infallible – knowledge that one has two hands.[11]

If the argument sketched above is sound, however, then the following theses would be inconsistent:

A. Knowledge is factive.

B. Knowledge is fallible.

As I have argued above, since contemporary epistemologists generally accept (A), they are also committed to the conclusion that knowledge is epistemic certainty, i.e., (3), since this conclusion follows from (A) by the argument sketched above. If we accept both (A) and (3), however, we must reject (B). For to say that knowledge is fallible is to say that *S* can know that *p* on the grounds that *e* even if the truth of *p* is not guaranteed by *e*.[12] But this is contrary to (A) and (3), which state that, if *S* knows that *p* on the grounds that *e*, then *p* is guaranteed to be true by *e*; it cannot be false.[13] If the argument sketched above is sound, then (A) implies (3), which would mean, in turn, that we should hold on to (A) and (3), and reject (B).

[9] Fred Dretske, *Perception, Knowledge, and Belief* (Cambridge: Cambridge University Press, 2000).

[10] Fred Dretske, "Gettier and Justified True Belief: 50 Years On," in *The Philosophers' Magazine* (23 January, 2015), https://www.philosophersmag.com/essays/10-gettier-and-justified-true-belief-50-years-on.

[11] See, for example, Ben Bronner, "Assertions Only?" *Thought: A Journal of Philosophy* 2 (2013): 44-52.

[12] Jessica Brown, *Fallibilism: Evidence and Knowledge* (Oxford: Oxford University Press, 2018), 2.

[13] Reed, "Certainty," https://plato.stanford.edu/archives/win2011/entries/certainty.

NOTES ON THE CONTRIBUTORS

Adam Michael Bricker is a postdoctoral researcher at the University of Oulu. He completed his PhD at the University of Edinburgh in 2018 with a thesis arguing against the factivity of knowledge. His research efforts are presently split between the application of epistemology to the philosophy of historiography and the application of electrophysiological methods to epistemology. His present empirical research is conducted in the EEG lab at the University of Turku, in conjunction with the Turku Brain and Mind Centre. He has also published on the philosophy of risk. Contact: Adam.Michael.Bricker@gmail.com.

James Cargile retired in Spring 2018 after 53 years on the faculty of the Corcoran Department of Philosophy at the University of Virginia. Contact: jc7y@virginia.edu.

David Coss is currently an Instructor of Philosophy at Indiana University Kokomo. His primary research interests lie in Epistemology, with an emphasis on Contextualism and Pragmatic Encroachment, as well as issues in social epistemology (esp. epistemic injustice). His recent papers include "Contextualism and Context Internalism" published in this journal as well as "Interest-Relative Invariantism and Indifference Problems" published in *Acta Analytica*. He is currently co-authoring a paper on the relationship between practical interests and theories well-being as it relates to pragmatic encroachment. He may be reached at davidcossphilosopher@gmail.com.

Gábor Forrai obtained his PhD at the University of Notre Dame and is currently Professor at the Institute of Business Economics of Eötvös Loránd University, Budapest. He has published on epistemology, philosophy of language, philosophy of mind and philosophy of science. His works include *Reference, Truth and Conceptual Schemes* (Kluwer, 2001) and *Intentionality: Past and Future* (coedited with George Kampis, Rodopi, 2005). Contact: forrai@gti.elte.hu.

Christos Kyriacou is a Lecturer at the University of Cyprus. His main research interests lie in epistemology, metaethics and their intersection. He has published in *Synthese*, *Erkenntnis*, *Episteme*, *Philosophical Pyschology*, *Journal of Ethics and*

Social Philosophy and elsewhere and in volumes by Oxford University Press and DeGruyter. He has also edited the volume *Metaepistemology* (2018, Palgrave) with Robin McKenna and currently preparing *Skeptical Invariantism Reconsidered* for Routledge (with Kevin Wallbridge). Contact: ckiriakou@gmail.com.

Daniella Meehan is a PhD student in the subject of Philosophy at the University of Glasgow and a member of the university's epistemology research centre, 'COGITO.' Her primary research lies in the intersection between epistemology and ethics, specifically vice epistemology and epistemic responsibility. She is particularly interested in the analysis of intellectual vices such as ignorance, close-mindedness and epistemic injustice, and what it takes for people who possess these to be blameworthy. Her aim is to form a novel account of intellectual vices which has blameworthiness as a constitute and central feature, where blameworthiness is understood as a distinct form of epistemic blame. Contact: d.meehan.1@gla.ac.uk.

Moti Mizrahi is Associate Professor of Philosophy in the School of Arts and Communication at the Florida Institute of Technology. He is also an Associate and Book Reviews Editor of *Philosophia* and the editor of *The Kuhnian Image of Science: Time for a Decisive Transformation?* (Rowman & Littlefield, 2018). He has teaching and research interests in argumentation, epistemology, ethics, logic, philosophy of religion, and philosophy of science. He has published extensively on the philosophy of science, the scientific realism/anti-realism debate, the epistemology of philosophy, and argumentation. His work has appeared in journals such as *Argumentation*, *Erkenntnis*, *Philosophical Studies*, *Studies in History and Philosophy of Science*, and *Synthese*. Contact: motimizra@gmail.com.

Giovanni Rolla is an associate professor at Federal University of Bahia (Brazil) and a permanent member of the Graduate Program of Philosophy at the Federal University of Piauí (Brazil). His works deal with philosophy of cognition, focusing on embodiment and enactivism, and with traditional epistemology, including analysis of knowledge, know-how, normativity and skepticism. He is also secretly an experimentalist musician and loves cats. Contact: rollagiovanni@gmail.com.

Michael J. Shaffer is currently professor of philosophy at St. Cloud State University. He is also a fellow of the center for formal epistemology at Carnegie-Mellon University, a fellow of the Rotman Institute for Science and Values at the University of Western Ontario, a Lakatos fellow at the London School of Economics, a fellow of the University of Cologne's summer institute for

epistemology and an NEH fellow at the University of Utah. His main areas of research interest are in epistemology, logic and the philosophy of science, and he has published nearly fifty articles on various topics in these areas. He is co-editor of *What Place for the A Priori?* (Open Court, 2011) and is the author of *Counterfactuals and Scientific Realism* (Palgrave-MacMillan, 2012), *Quasi-factive Belief and Knowledge-like States* (Lexington, forthcoming) and *The Experimental Turn and the Methods of Philosophy* (Routledge, forthcoming). He is also currently working on a co-authored book criticizing alternative medicine. Contact: shaffermphil@hotmail.com.

LOGOS & EPISTEME: AIMS & SCOPE

Logos & Episteme is a quarterly open-access international journal of epistemology that appears at the end of March, June, September, and December. Its fundamental mission is to support philosophical research on human knowledgein all its aspects, forms, types, dimensions or practices.

For this purpose, the journal publishes articles, reviews or discussion notes focused as well on problems concerning the general theory of knowledge, as on problems specific to the philosophy, methodology and ethics of science, philosophical logic, metaphilosophy, moral epistemology, epistemology of art, epistemology of religion, social or political epistemology, epistemology of communication. Studies in the history of science and of the philosophy of knowledge, or studies in the sociology of knowledge, cognitive psychology, and cognitive science are also welcome.

The journal promotes all methods, perspectives and traditions in the philosophical analysis of knowledge, from the normative to the naturalistic and experimental, and from the Anglo-American to the Continental or Eastern.

The journal accepts for publication texts in English, French and German, which satisfy the norms of clarity and rigour in exposition and argumentation.

Logos & Episteme is published and financed by the "Gheorghe Zane" Institute for Economic and Social Research of The Romanian Academy, Iasi Branch. The publication is free of any fees or charges.

For further information, please see the Notes to Contributors.

Contact: logosandepisteme@yahoo.com.

NOTES TO CONTRIBUTORS

1. Accepted Submissions

The journal accepts for publication articles, discussion notes and book reviews.

Please submit your manuscripts electronically at: logosandepisteme@yahoo.com. Authors will receive an e-mail confirming the submission. All subsequent correspondence with the authors will be carried via e-mail. When a paper is co-written, only one author should be identified as the corresponding author.

There are no submission fees or page charges for our journal.

2. Publication Ethics

The journal accepts for publication papers submitted exclusively to *Logos & Episteme* and not published, in whole or substantial part, elsewhere. The submitted papers should be the author's own work. All (and only) persons who have a reasonable claim to authorship must be named as co-authors.

The papers suspected of plagiarism, self-plagiarism, redundant publications, unwarranted ('honorary') authorship, unwarranted citations, omitting relevant citations, citing sources that were not read, participation in citation groups (and/or other forms of scholarly misconduct) or the papers containing racist and sexist (or any other kind of offensive, abusive, defamatory, obscene or fraudulent) opinions will be rejected. The authors will be informed about the reasons of the rejection. The editors of *Logos & Episteme* reserve the right to take any other legitimate sanctions against the authors proven of scholarly misconduct (such as refusing all future submissions belonging to these authors).

3. Paper Size

The articles should normally not exceed 12000 words in length, including footnotes and references. Articles exceeding 12000 words will be accepted only occasionally and upon a reasonable justification from their authors. The discussion notes must be no longer than 3000 words and the book reviews must not exceed 4000 words, including footnotes and references. The editors reserve the right to ask the authors to shorten their texts when necessary.

4. Manuscript Format

Manuscripts should be formatted in Rich Text Format file (*rtf) or Microsoft Word document (*docx) and must be double-spaced, including quotes and footnotes, in 12 point Times New Roman font. Where manuscripts contain special symbols, characters and diagrams, the authors are advised to also submit their paper in PDF format. Each page must be numbered and footnotes should be numbered consecutively in the main body of the text and appear at footer of page. For all references authors must use the Humanities style, as it is presented in The Chicago Manual of Style, 15th edition. Large quotations should be set off clearly, by indenting the left margin of the manuscript or by using a smaller font size. Double quotation marks should be used for direct quotations and single quotation marks should be used for quotations within quotations and for words or phrases used in a special sense.

5. Official Languages

The official languages of the journal are: English, French and German. Authors who submit papers not written in their native language are advised to have the article checked for style and grammar by a native speaker. Articles which are not linguistically acceptable may be rejected.

6. Abstract

All submitted articles must have a short abstract not exceeding 200 words in English and 3 to 6 keywords. The abstract must not contain any undefined abbreviations or unspecified references. Authors are asked to compile their manuscripts in the following order: title; abstract; keywords; main text; appendices (as appropriate); references.

7. Author's CV

A short CV including the author`s affiliation and professional postal and email address must be sent in a separate file. All special acknowledgements on behalf of the authors must not appear in the submitted text and should be sent in the separate file. When the manuscript is accepted for publication in the journal, the special acknowledgement will be included in a footnote on the first page of the paper.

8. Review Process

The reason for these requests is that all articles which pass the editorial review, with the exception of articles from the invited contributors, will be subject to a strict double anonymous-review process. Therefore the authors should avoid in

their manuscripts any mention to their previous work or use an impersonal or neutral form when referring to it.

The submissions will be sent to at least two reviewers recognized as specialists in their topics. The editors will take the necessary measures to assure that no conflict of interest is involved in the review process.

The review process is intended to be as quick as possible and to take no more than three months. Authors not receiving any answer during the mentioned period are kindly asked to get in contact with the editors.

The authors will be notified by the editors via e-mail about the acceptance or rejection of their papers.

The editors reserve their right to ask the authors to revise their papers and the right to require reformatting of accepted manuscripts if they do not meet the norms of the journal.

9. Acceptance of the Papers

The editorial committee has the final decision on the acceptance of the papers. Papers accepted will be published, as far as possible, in the order in which they are received and they will appear in the journal in the alphabetical order of their authors.

10. Responsibilities

Authors bear full responsibility for the contents of their own contributions. The opinions expressed in the texts published do not necessarily express the views of the editors. It is the responsibility of the author to obtain written permission for quotations from unpublished material, or for all quotations that exceed the limits provided in the copyright regulations.

11. Checking Proofs

Authors should retain a copy of their paper against which to check proofs. The final proofs will be sent to the corresponding author in PDF format. The author must send an answer within 3 days. Only minor corrections are accepted and should be sent in a separate file as an e-mail attachment.

12. Reviews

Authors who wish to have their books reviewed in the journal should send them at the following address: Institutul de CercetăriEconomiceşiSociale „Gh. Zane" Academia Română, FilialaIaşi, Str. Teodor Codrescu, Nr. 2, 700481, Iaşi, România.

The authors of the books are asked to give a valid e-mail address where they will be notified concerning the publishing of a review of their book in our journal. The editors do not guarantee that all the books sent will be reviewed in the journal. The books sent for reviews will not be returned.

13. Property & Royalties

Articles accepted for publication will become the property of *Logos & Episteme* and may not be reprinted or translated without the previous notification to the editors. No manuscripts will be returned to their authors. The journal does not pay royalties.

14. Permissions

Authors have the right to use their papers in whole and in part for non-commercial purposes. They do not need to ask permission to re-publish their papers but they are kindly asked to inform the Editorial Board of their intention and to provide acknowledgement of the original publication in *Logos & Episteme*, including the title of the article, the journal name, volume, issue number, page number and year of publication. All articles are free for anybody to read and download. They can also be distributed, copied and transmitted on the web, but only for non-commercial purposes, and provided that the journal copyright is acknowledged.

15. Electronic Archives

The journal is archived on the Romanian Academy, Iasi Branch web page. The electronic archives of *Logos & Episteme are* also freely available on Philosophy Documentation Center web page.