Logos & Episteme

an international journal of epistemology

TABLE OF CONTENTS

RESEARCH ARTICLES

GIVING UP THE ENKRATIC PRINCIPLE

Claire Field

ABSTRACT: The Enkratic Principle enjoys something of a protected status as a requirement of rationality. I argue that this status is undeserved, at least in the epistemic domain. Compliance with the principle should not be thought of as a requirement of epistemic rationality, but rather as defeasible indication of epistemic blamelessness. To show this, I present the Puzzle of Inconsistent Requirements, and argue that the best way to solve it is to distinguish two kinds of epistemic evaluation – requirement evaluations and appraisal evaluations. This allows us to solve the puzzle while accommodating traditional motivations for thinking of the Enkratic Principle as a requirement of rationality.

KEYWORDS: Enkratic Principle, coherence, evidence, rationality

1. The Enkratic Principle

The Enkratic Principle demands coherence. In the epistemic domain, it demands coherence between the agent's beliefs about which epistemic attitudes she ought to have, and her first-order[1] epistemic attitudes.[2] According to the orthodox view, the Enkratic Principle is a requirement of epistemic rationality.[3] Reading O as "rationally required," Φ as representing a doxastic attitude, and B as representing belief, the principle can be stated as follows:

Enkratic Principle: O (BOΦ → Φ)

The Enkratic Principle says that rationality requires *either* having the attitudes you believe you ought to have, *or* giving up the belief that you ought to have those attitudes. It prohibits combinations of attitudes that include the belief that believing P is required, but not the belief P.

It is not hard to see why the Enkratic Principle has been thought to be a requirement of epistemic rationality – attitudinal coherence can seem definitional

---

[1] 'First-order' is not the ideal term. By 'first-order' I mean epistemic attitudes whose content does not concern what we ought, rationally, to believe.

[2] These could include believing, disbelieving, refraining from believing, suspending, or having a particular credence in a proposition.

[3] It is also typically thought to be a requirement of practical rationality, but here I am concerned only with the epistemic version.

of what rationality demands.[4] Nevertheless, rejecting this orthodoxy is the best way to solve a stubborn puzzle that arises from the possibility of misleading evidence about what rationality requires – the Puzzle of Inconsistent Requirements. I argue that distinguishing between two distinct kinds of epistemic evaluation – evaluations of whether attitudes meet the requirements of rationality, and evaluations of whether epistemic praise or blame is deserved – offers the best way to solve the puzzle. Rather than thinking of the Enkratic Principle as a requirement of rationality, we should view the attitudinal coherence it demands as defeasible indication that the agent is epistemically blameless. The following section describes the Puzzle of Inconsistent Requirements. Section 3 diagnoses the puzzle as arising from a conflation of two distinct kinds of evaluation sometimes associated with evaluations of rationality. Section 4 shows how distinguishing these allows us to solve the puzzle. Section 5 defends the proposed strategy as the best solution available – it is the least theoretically costly of the available solutions, and it allows us to accommodate traditional motivations for the orthodox view.

2. The Puzzle of Inconsistent Requirements

Suppose that you are required to Φ, in virtue of some set of normative requirements. Then, suppose that you have some misleading evidence that in fact, you are required to refrain from Φ-ing. Are you required to Φ? Or are you required to refrain from Φ-ing, as your evidence indicates? On the one hand, you seem to be required to Φ, since this is what complying with the normative requirements demands. However, you also seem to be required to refrain from Φ-ing. This is what your evidence indicates you ought to do, and the Enkratic Principle requires coherence between your higher-order judgments about what you ought to do. You cannot rationally give up the higher-order judgment because it is supported by your evidence. So, you appear to be subject to conflicting requirements to Φ and also to refrain from Φ-ing. This generates a puzzle. The puzzle is particularly problematic when the normative requirements involved are the requirements of epistemic rationality. Here is an example:

---

[4] See, for example, John Broome, *Rationality Through Reasoning* (Wiley-Blackwell, 2013); Richard Feldman, "Respecting the Evidence," *Philosophical Perpsectives* 19, 1 (2005), Sophie Horowitz, "Epistemic Akrasia," *Noûs* 48 (2014); Niko Kolodny, "Why Be Rational?," *Mind* 114 (2005), Clayton Littlejohn "Stop Making Sense? On a Puzzle about Rationality," *Philosophy and Phenomenological Research* 93 (2015), Michael Titelbaum, "Rationality's Fixed Point (Or: In Defense of Right Reason)," in *Oxford Studies in Epistemology,* Vol. 5, eds. Tamar Szabó Gendler and John Hawthorne (Oxford: Oxford University Press: 2015), and Jonathan Way, "The Symmetry of Rational Requirements," *Philosophical Studies* 155 (2011).

Logic 101. Suppose that rationality prohibits contradictory belief. Suppose also that it is your first day of university, and you are about to take your first philosophy class. You sign up for an introductory course in logic. Unluckily, your instructor is an overzealous advocate of dialetheism.[5] He believes that rationality sometimes requires inconsistent belief, particularly in matters concerning truth,[6] and he intends to set you on the right track by introducing you to the best arguments in favour of this position. You study in depth all the best arguments for dialetheism, and by the end of the course you believe – on the basis of good but misleading evidence – that rationality sometimes requires contradictory belief. As you walk out of class you see some graffiti that is a version of the Liar Paradox ('The writing on this wall is false.').

You seem to be simultaneously required to believe and *not* believe a contradiction. The true requirements require you to refrain from believing contradictions. However, you have testimony and arguments indicating that in this case, rationality requires you to believe a contradiction. Assuming that epistemic rationality requires you to believe what your evidence supports, you are required to believe that rationality requires you to believe the contradiction. You can comply with the Enkratic Principle by either believing the contradiction, or by giving up the belief that you are rationally required to believe a contradiction. However, you should not give up this belief. Your evidence supports it, so to give it up would be to ignore your evidence. So, if you are to believe what your evidence supports, and comply with the Enkratic Principle, then you ought to believe the contradiction. Your unfortunate epistemic situation – the fact that you have been exposed to misleading arguments for a false philosophical view – does not change what rationality requires. If epistemic rationality really prohibits contradictory belief, then you are required to refrain from believing contradictions. The Puzzle of Inconsistent Requirements is that epistemic rationality appears to generate inconsistent requirements in cases such as this. It appears to require you to both believe and *not* believe the contradiction.

---

[5] As developed by Graham Priest. See his "The Logic of Paradox," *Journal of Philosophical Logic* 8 (1985), "Contradiction, Belief and Rationality," *Proceedings of the Aristotelian Society* 86 (1985), *Doubt Truth to Be a Liar* (Oxford: Oxford University Press: 2005), and *In Contradiction: A Study of the Transconsistent* (Oxford: Oxford University Press, 2006).

[6] For example, "it seems to me that anyone weighing up the state of play concerning [truth], ought rationally to be inconsistent." (Priest, *In Contradiction*, 125).

Claire Field

It is worth pointing out that generating the puzzle does not depend on the possibility of rationally doubting any particular principle of classical logic.[7] Here is another example:[8]

> *Disagreement.* Suppose that rationality requires you to remain steadfast in your beliefs when epistemic peers disagree with you. Then suppose you write your PhD thesis on the epistemology of disagreement. You exert significant effort considering arguments bearing on whether one should conciliate or remain steadfast in the face of disagreement, reaching the conclusion that rationality requires conciliation in response to disagreement from epistemic peers.[9] One evening, while discussing politics, you assert P (in which you have a credence of 0.8). Your partner (who is your epistemic peer on this matter) disagrees with you.

Epistemic rationality requires you to remain steadfast. Nevertheless, your total evidence supports the view that rationality requires you to conciliate. By the Enkratic Principle, you ought to either reduce your credence in P or give up the belief that you ought to conciliate in response to peer disagreement. However, to give up the belief that you ought to conciliate would be to ignore your evidence, so you do not seem to be rationally permitted to take this option. Even so, the fact that your evidence supports a false philosophical view does not change what rationality requires of you; it does not change the fact that you ought to remain steadfast. Again, rationality appears to make conflicting demands of you: you are required to both reduce and *not* reduce your credence in P.

The Puzzle of Inconsistent Requirements involves cases of apparent intra-domain conflict between normative requirements. It is thus importantly different

---

[7] Some have thought the irrationality of contradictions is too certain to be rationally doubted. Putnam argues that his minimal principle of contradiction ('not every statement is true and false') presupposes the possibility of debate, thought, and explanation (Hilary Putnam, "There is at least one a priori truth," *Erkenntnis* 13). Lewis declines to debate the matter because "the principles *not* in dispute are so very much less certain than non-contradiction itself that it matters little whether or not a successful defence of non-contradiction could be based on them" (David Lewis, "Letters to Priest and Beall," in *The Law of Non-Contradiction,* eds. Graham Priest, J.C. Beall, and Bradley Armour-Garb (Oxford University Press, 2004), 176). The Puzzle of Inconsistent Requirements does not turn on these issues.

[8] The puzzle has also been motivated using requirements to (not) believe lottery propositions (Littlejohn, "Stop Making Sense"); requirements governing perception, testimony, and Lewis' Principal Principle (Darren Bradley, "Are There Indefeasible Epistemic Rules?," *Philosophers' Imprint* 19 (2019): 2); and external world scepticism (Feldman, "Respecting the Evidence").

[9] Perhaps you think that intellectual modesty is more important than avoiding theories that are self-undermining (in disagreement with Adam Elga, "How to Disagree about How to Disagree," in *Disagreement,* eds. Richard Feldman and Ted Warfield (Oxford University Press, 2007).

from a similar, and more benign, puzzle involving int*er*-domain conflict between normative requirements, for example:

> *Lying*. Suppose that morality requires that you act so as to protect the lives of innocents, in all situations – even if this would sometimes require you to lie. Suppose also that you have recently taken a course on ethics led by a professor who defends a somewhat extreme Kantian view on the moral permissibility of lying. According to this view, lying is morally wrong in all cases, even if by lying you could save a life. You are a good student – you do the reading, you follow the arguments you are presented with in class, and you come to believe that lying is morally impermissible in all cases. One day, an infamous axe murderer comes to your door, in pursuit of your friend Stefano, and asks you where he is. In fact, Stefano is hiding in your house. You know that you could easily save his life by lying to the murderer.[10]

Morality requires that you lie to protect Stefano. However, your total evidence supports the view that morality prohibits lying. In this instance, morality requires you to lie, but epistemic rationality seems to require you not to lie. This is much less worrying than the Puzzle of Inconsistent Requirements; the observation that morality and epistemic rationality sometimes conflict is not, itself, surprising. To resolve inter-domain conflict, we need only decide which set of normative requirements to prioritise – those of morality or rationality. While this is not always a straightforward question,[11] we can answer it without overhauling our theories of morality or rationality. The same cannot be said for the int*ra*-domain involved in Logic 101 and Disagreement. Unlike int*er*-domain conflict, int*ra*-domain conflict implies that our current theories of normative requirements have internal inconsistencies demanding resolution.

The Puzzle of Inconsistent Requirements is also importantly different from puzzles involving misleading higher-order evidence.[12] Here is a typical example:

---

[10] See Immanuel Kant, "On a Supposed Right to Lie from Philanthropy (1797)," in *Practical Philosophy,* ed. Mary Gregor (Cambridge University Press, 1996). For a sympathetic approach to Kant's rigorism about lying, see Wolfgang Schwarz, "Kant's Refutation of Charitable Lies," *Ethics* 81(1970) and Jacob Weinrib, "The Juridical Significance of Kant's 'Supposed Right to Lie'," *Kantian Review* 13 (2008). Many have attempted to save Kant from the commitment to prohibit lying in all cases, including: Christine Korsgaard, "The Right to Lie: Kant on Dealing with Evil," *Philosophy and Public Affairs* 15 (1986); Tamar Schapiro, "Kantian Rigorism and Mitigating Circumstances," *Ethics* 117 (2006); and Allen Wood, *Kantian Ethics.* (Cambridge University Press, 2008, Ch. 14).

[11] See Broome, *Rationality Through Reasoning* and "Normative Requirements," *Ratio* 12 (1999), as well as Elizabeth Harman, "The Irrelevance of Moral Uncertainty," in *Oxford Studies in Metaethics,* Vol. 10, ed. Russ Shafer-Landau (Oxford University Press, 2015) for discussion of this.

[12] This similar but distinct puzzle is discussed in David Christensen, "Higher-Order Evidence," *Philosophy and Phenomenological Research* 81 (2010); Adam Elga, "The Puzzle of the Unmarked

> *Medicine.* You are a medical doctor who diagnoses patients and prescribes appropriate treatment. After diagnosing a particular patient's condition and prescribing certain medications, you are informed by a nurse that you have been awake for 36 hours. You diagnose a patient with Disease D, forming the belief "the patient has D." In fact, you are correct. However, you know that people are prone to making cognitive errors when sleep-deprived (perhaps you even know your own poor diagnostic track-record under such circumstances), so you also believe "I ought not believe that patient has D," because it is not supported by your evidence.[13]

Your evidence appears to support both of the following beliefs:[14]

> The patient has disease D
>
> My evidence does not support that the patient has disease D.

The problem is that you have good evidence for P, but misleading evidence about your current ability to assess the evidence for P. If these beliefs are supported by your total evidence, then believing what your evidence supports means believing both. However, this seems to violate the Enkratic Principle. Misleading higher-order evidence puzzles rely on two substantive claims about epistemic rationality: that it is possible for one's total evidence to support such combinations,[15] and that what

Clock and the New Rational Reflection Principle," *Philosophical Studies* 164 (2013), Horowitz, "Epistemic Akrasia," Maria Lasonen-Aarnio, "Enkrasia or Evidentialism?," *Philosophical Studies* 177 (2020)*,* "Higher-order Evidence and the Limits of Defeat," *Philosophy and Phenomenological Research* (2014), Paulina Sliwa and Sophie Horowitz, "Respecting all the Evidence," *Philosophical Studies* 172 (2015); Brian Weatherson, *Normative Externalism* (Oxford: Oxford University Press, 2019), Timothy Williamson, "Improbable Knowing," in *Evidentialism and its Discontents*, ed. Trent Dougherty (Oxford University Press, 2011); and Alex Worsnip, "The Conflict of Evidence and Coherence," *Philosophy and Phenomenological Research* 96 (2018). For a solution to that puzzle that, like this one, accepts the occasional rationality of level-incoherence, see Lasonen-Aarnio ("Enkrasia or Evidentialism?", "Higher-Order Evidence and the Limits of Defeat"), and Weatherson (*Normative Externalism*).

[13] Christensen ("Higher Order Evidence," 186) and Weatherson (*Normative Externalism*, 130) discuss this example.

[14] At least, according to Christensen ("Higher Order Evidence"), Lasonen-Aarnio ("Enkrasia or Evidentialism?," "Higher-order Evidence and the Limits of Defeat"), and Weatherson (*Normative Externalism*).

[15] For some disagreement, see Jessica Brown (*Fallibilism: Evidence and Knowledge* (Oxford University Press, 2018), Ch. 5, 6)), Horowitz, ("Epistemic Akrasia"), and Sliwa and Horowitz ("Respecting All the Evidence"). This also turns on the question of whether higher-order evidence 'screens off' first-order evidence. For discussion, see Richard Feldman ("Reasonable Religious Disagreements," in *Philosophers Without Gods: Meditations on Atheism and the Secular Life*, ed. Louise Antony (Oxford University Press, 2007)), Brandon Fitelson ("Evidence of evidence is not (necessarily) evidence," *Analysis* 72 (2012)), William Roche (Is Evidence of Evidence Evidence?

epistemic rationality requires is completely determined by what one's total evidence supports. Although the Puzzle of Inconsistent Requirements *can* be stated in terms of misleading evidence about what epistemic rationality requires, it is independent to this and other specific claims about what epistemic rationality requires. The intra-domain conflict arises not between first- and higher-order evidence, but between what the evidence supports and what the true requirements require. All that is needed to generate it is at least one requirement of rationality (e.g. Non-Contradiction, Conciliation), a commitment to the Enkratic Principle, and some misleading evidence sufficient to support a false belief about what rationality requires. This makes possible situations in which one is rationally required (by the misleading evidence) to believe a false claim about what rationality requires, and also rationally required (by the true requirements of rationality) to adopt a first-order attitude that conflicts with that false belief.

The Puzzle of Inconsistent Requirements depends on the following three plausible claims, and can be solved by giving up at least one of them:

> *Externalism:* There are facts about what rationality requires ("the rational requirements"), and these facts are independent of agents' epistemic perspectives.[16]

> *Support*: If S's total evidence supports believing P, then S is rationally required to believe P.[17]

> *Enkratic Principle*: Rationality requires that the agent either has the attitudes she believes she ought to have, or gives up the belief that she ought to have those attitudes. [O (BOΦ → Φ)].

I argue that the right solution is to give up the Enkratic Principle, viewing attitudinal coherence not as a requirement of rationality but instead as defeasible indication of epistemic blamelessness. I argue that this solution incurs the fewest theoretical costs of the available solutions.

Screening-Off vs. No-Defeaters," *Episteme* 15 (2018), William Roche and Tomoji Shogenji ("Confirmation, transitivity, and Moore: the Screening-Off Approach," *Philosophical Studies* 3 (2013)).

[16] In other words, the requirements are 'indefeasible' – they remain binding in all possible situations. Darren Bradley ("Are There Indefeasible Epistemic Rules?" *Philosophers' Imprint* 19 (2019)) and Jonathan Ichikawa and Benjamin Jarvis, *The Rules of Thought* (Oxford University Press, 2013)) endorse this view of requirements of rationality.

[17] I express Support in terms of evidence, but this is not necessary. One might think that non-evidential forms of rational support are also relevant to what one is rationally required to believe, and thus may wish to expand Support to include, for example, epistemic reasons to suspend belief, whether propositions are produced by a reliable method, or theoretical virtues such as simplicity or elegance. This would be compatible with the puzzle, and my proposed solution.

Claire Field

It is worth briefly addressing a worry. One might wonder whether sacrificing one of the three claims really exhausts our options for solving the puzzle. One might be tempted to think that we should deny the claim that rationality requires both the false higher-order belief and the conflicting first-order attitude. After all, if rationality does not require both then there is no conflict. For example, Lord and Sylvan argue that the required attitude at both the first- and higher-order is suspension.[18] However, this means giving up *both* Externalism and Support. In the cases at hand, the agent's evidence supports the false belief that rationality requires believing P. However, due to some other true rational requirement, rationality in fact requires something else that is incompatible with the false view (e.g. *not* believing P). Responding to this conflict by saying that, in fact, rationality requires *neither* of these apparently conflicting attitudes, but suspension instead, would mean denying *both* Support (rationality does not, in this case, require believing what one's total evidence supports), *and* Externalism (in this case, one is not required to believe what the true requirements require). This strategy is unnecessarily costly – it involves giving up two of the three claims, but we need only give up one. The strategies I explore involve giving up only one of the claims.

In the following section, I argue that we can diagnose the apparent conflict in Puzzle of Inconsistent Requirements as a conflation of two kinds of evaluation – requirement and appraisal. Distinguishing these evaluations means that we need not think that all three of the claims (Externalism, Support, Enkratic Principle) are associated with the same kind of evaluation – whether of requirement or appraisal. This gives us a way to solve the puzzle, provided we can associate at least one claim with appraisal, rather than requirement evaluations. I argue that the Enkratic Principle is naturally conceived of as associated with appraisal. This allows us to both preserve the requirement to believe what our total evidence supports and any other requirement of rationality we like (bar the Enkratic Principle) as fully robust requirements that apply without exception. Even better, the traditional motivations for thinking of the Enkratic Principle as a requirement of rationality can be accommodated by a view that takes compliance with it as defeasible indication of epistemic blamelessness.

[18] Errol Lord and Kurt Sylvan, "Suspension, Higher-Order Evidence, and Defeat," forthcoming in *Reasons, Justification, and Defeat*, eds. Mona Simion and Jessica Brown (Oxford University Press).

3. Two Kinds of Evaluation

This section diagnoses the apparent conflict in the Puzzle of Inconsistent Requirements as the conflation of two kinds of evaluation: requirement and appraisal. Distinguishing these permits a solution to the puzzle.

Judgments of epistemic rationality may involve either or both of the following evaluations:

> *Requirement*: S has the attitudes required by the requirements of rationality.

> *Appraisal*: S is epistemically blameless.

When an evaluation of rationality is a claim about requirement, it focuses on whether the agent has the attitudes she is required to have. Agents are evaluated as having met the requirements of rationality when they have the attitudes required by rationality. When an evaluation of rationality is a claim about appraisal, it focuses on whether and to what extent the agent deserves epistemic praise or blame. Agents are evaluated as epistemically blameless when they do not deserve epistemic blame, or they deserve epistemic praise. Various considerations can contribute to an agent's being epistemically blameless. For example, that she exhibits epistemic virtues or avoids epistemic vices,[19] that she responds appropriately given her epistemic perspective,[20] that she demonstrates the right kind of concern for epistemic reasons,[21] or that she manifests success-conducive dispositions.[22] Additionally various excuses, when applicable, can mitigate epistemic blame she might otherwise deserve – for example, that she was misled, deceived, did as well as she could given her circumstances, or lacked the relevant capacity.[23] Importantly, appraisal and

---

[19] Quassim Cassam, "Vice Epistemology," *The Monist* 99 (2016); *Vices of the Mind: from the Intellectual to the Political* (Oxford University Press, 2019).

[20] Jonathan Kvanvig, *Rationality and Reflection: How to Think About What to Think*. (Oxford University Press, 2014); Errol Lord, *The Importance of Being Rational*. (Oxford University Press, 2018); Michael Zimmerman, *Living with Uncertainty: The Moral Significance of Ignorance*. (Cambridge University Press, 2008).

[21] Cameron Boult, "Excuses, Exemptions, and Derivative Norms," *Ratio* 32 (2019).

[22] Maria Lasonen-Aarnio, "Enkrasia or Evidentialism?"; Timothy Williamson, "Justifications, Excuses, and Sceptical Scenarios," forthcoming in *The New Evil Demon Problem*, eds. Julien Dutant and Fabian Dorsch (Oxford University Press).

[23] One might think that lacking capacities constitutes an exemption, rather than an excuse, although there is debate to be had over whether exemption is a *sui generis* category, or merely a full excuse. Nothing here turns on whether there is a distinction between excuse and exemption. For further discussion of the conditions for blameworthy belief see Pamela Hieronymi, "Responsibility for Believing," *Synthese* 161 (2008); Miriam McCormick, "Taking Control of Belief," *Philosophical Explorations* 14 (2011); Conor McHugh, "Epistemic Responsibility and

requirement evaluations can come apart – agents should not always be praised (or escape blame) for doing what is required, and they do not always deserve blame for failing to do what is required. The distinction between requirement and appraisal evaluations is well-established in ethics. For example:

> *Kant's Shopkeeper.*[24] A shopkeeper prices his wares fairly, as morality requires him to do. However, he does this not out of a motivation to do what is fair, kind, or morally right, but out of a motivation to maximise his profits. He knows that if he does not price his wares fairly, his customers will go elsewhere. If he could make more profit by pricing his wares unfairly, then he would do this instead.

Kant's shopkeeper does what is required – he succeeds in complying with the requirements, but he does so in such a way that does not deserve praise. So, he deserves a positive requirement evaluation, and a neutral appraisal evaluation.[25] Likewise, failing to meet requirements does not always deserve blame:

> *Toes.* I step on your toe in a crowded lift, and in doing so cause you pain. Suppose that, *ceteris paribus*, causing others pain for no good reason is prohibited by the requirements of morality. However, I step on your toe not out of any intention to cause you pain, but because the lift is crowded and I am not aware of where your toe is. Had I known your toe was there, I would not have stepped on it.

Here, I fail to comply with the requirement to avoid causing pain to others, but I do not deserve blame.[26] In ethics, considerations relevant to agent appraisal are usually distinguished from considerations of whether the agent does what is required, and making this distinction is also useful in epistemology.[27]

Doxastic Agency," *Philosophical Issues* 23 (2013); Nikolaj Nottelmann, *Blameworthy Belief: A Study in Epistemic Deontologism* (Springer, 2007); Rik Peels, *Responsible Belief: A Theory in Ethics and Epistemology* (Oxford University Press, 2016); Angela Smith, "Responsibility for Attitudes: Activity and Passivity in Mental Life," *Ethics* 115 (2005). For discussion of the practice of epistemic blaming, see Jessica Brown, "What is Epistemic Blame?," *Noûs* 54 (2018).

[24] For a discussion of Kant's own example, see Jens Timmerman, "Acting from Duty: Inclination, Reason and Moral Worth" in *Kant's "Groundwork of the Metaphysics of Morals:" A Critical Guide,* ed. Jens Timmerman (Cambridge University Press, 2007).

[25] This does not imply that he deserves blame. Other examples are possible in which the agent does what is required, but deserves blame. For example, a politician who makes a large donation to charity only to detract attention from his seriously corrupt activities. Positive requirement evaluations do not imply positive appraisal evaluations.

[26] Again, it does not necessarily deserve praise. I may deserve a neutral evaluation, or an excuse.

[27] Various others distinguish appraisal and requirement in both ethics and epistemology. For applications of the distinction in ethics, see Nomy Arpaly, *Unprincipled Virtue: An Inquiry into Moral Agency* (Oxford University Press, 2002) and Peter Graham, "In Defense of Objectivism About Moral Obligation," *Ethics* 121 (2010). For applications of the distinction to epistemology, see John Hawthorne and Amia Srinivasan, "Disagreement Without Transparency: Some Bleak

Consider the students of Logic 101. If they believe the contradiction, then they violate the requirements of rationality. However, various considerations suggest that they would nevertheless be epistemically blameless for doing so. For example, the forbidden first-order attitude (believing the contradiction) is a result of usually successful epistemic practices, such as trusting the testimony of an apparent expert and believing the conclusions of convincing arguments. We might think that they manifest praiseworthy epistemic virtues when they follow their evidence where it leads (even when that destination is highly counter-intuitive), and when they take seriously their own higher-order judgments. Were they instead to *not* believe the contradiction, we might think that this would manifest epistemic vice – stubbornness, or close-mindedness – and perhaps a lack of respect for the evidence and their own epistemic commitments.[28] They can also appeal to excuses for violating the requirements. For instance, they have been misled by convincing, but unsound, arguments. They are philosophical novices, and have limited capacities to work out what is wrong with the arguments they are given. These considerations suggest that students in Logic 101 who come to believe the contradiction are epistemically blameless, despite violating a requirement of rationality.

Some have denied the distinction between the requirement evaluations and appraisal evaluations, implying that evaluations that the agent is epistemically blameless and evaluations that the agent has met the requirements of rationality are one and the same.[29] However, anti-luminosity considerations offer an important reason to distinguish requirement and appraisal evaluations in epistemic rationality. Anti-luminosity says that there is no non-trivial condition for which it is always possible to know whether or not one has met that condition.[30] If this is right, then it

---

Thoughts" in *The Epistemology of Disagreement*. (Oxford University Press, 2013); Maria Lasonen-Aarnio, "Enkrasia or Evidentialism?," "Unreasonable Knowledge," *Philosophical Perspectives* 24 (2010); Clayton Littlejohn, *Justification and the Truth-Connection*. (Cambridge University Press, 2012); Jonathan Sutton, "Stick to what you know," *Noûs* 39 (2005), *Without Justification*. (MIT Press, 2007); Timothy Williamson, "Justifications, Excuses, and Sceptical Scenarios."

[28] This is true even if we have *prima facie* justification for the basic laws of logic, as some have argued (Ichikawa and Jarvis, *The Rules of Thought*; Titelbaum, "Rationality's Fixed Point (Or: In Defense of Right Reason);" and Crispin Wright "Intuition, Entitlement and the Epistemology of Logical Laws," *Dialectica* 58 (2004)). In Logic 101, the students are exposed to counter-arguments that they are not able to rationally reject, so it would be epistemically vicious for them to make use of this justification.

[29] See Kvanvig, *Rationality and Reflection: How to Think About What to Think*; Lord, *The Importance of Being Rational*; Ralph Wedgwood, *The Value of Rationality* (Oxford University Press, 2017).

[30] Trivial conditions immune to anti-luminosity are those that hold in either all or no cases, and conditions for which one cannot change from being in a position to know that it obtains to not

will not always be possible to know when one has met the requirements of rationality. Logic 101 and Disagreement are examples of how one can fail to be in a position to know whether one has met the requirements. Failure to know what is required means that complying with what is required is not always under one's control, because we cannot always tell what we ought to do in order to comply with the requirements.[31] In such cases, it is implausible that agents deserve epistemic blame for failing to meet requirements. Making meeting the requirements necessary for avoiding epistemic blame would mean that agents would sometimes deserve epistemic blame for failing to do what they were in no position to know they were failing to do. This would be implausibly harsh. To avoid this, many views of epistemic normativity recognise some blameless failures to meet requirements,[32] thus implicitly separating meeting requirements from epistemic appraisal.

It might have been thought that such harsh results could be avoided by adjusting what rationality requires so that the requirements of rationality are themselves luminous, and so can never be violated blamelessly. We might have them depend closely on the agent's mental states, or on how things seem to her by her own lights. However, this would not be sufficient. Anti-luminosity arguments show that no non-trivial condition is luminous, not even conditions for which compliance depends on how things seem from our own lights, such as feeling cold. Any genuinely luminous condition would need to be extremely trivial to ensure that agents were always in a position to know whether they were meeting it,[33] and requirements of rationality could not be this trivial.[34] The pressure to avoid both

---

being in such a position (Timothy Williamson, *Knowledge and its Limits* (Oxford University Press, 2002), 108). See also Amia Srinivasan, "Are We Luminous?," *Philosophy and Phenomenological Research* 90 (2015); "Normativity without Cartesian Privilege," *Philosophical Issues* 25 (2015).

[31] Srinivasan, "Normativity without Cartesian Privilege," 278.

[32] For example, those who endorse a truth or knowledge norm for belief typically respond to the New Evil Demon argument by evaluating envatted subjects who have false but responsibly formed beliefs as blameless. For discussion, see Brown, *Fallibilism: Evidence and Knowledge*; Stewart Cohen, "Justification and Truth," *Philosophical Studies* 46 (1984), Jeremy Fantl and Matthew McGrath, *Knowledge in an Uncertain World* (Oxford University Press, 2009), Christoph Kelp, "Justified Belief: Knowledge First-Style," *Philosophy and Phenomenological Research* 92 (2016), Clayton Littlejohn, "The Externalist's Demon," *Canadian Journal of Philosophy* 39 (2009), Sutton, *Without Justification*; Williamson, "Justifications, Excuses, and Sceptical Scenarios."

[33] Furthermore, it might be even more difficult for us to discern from the inside when we meet conditions that require us to be in particular mental states than when we meet external conditions (Eric Schwitzgebel, "The Unreliability of Naive Introspection," *Philosophical Review* 117 (2006); Srinivasan, "Normativity without Cartesian Privilege," 275.

[34] Even if such requirements were possible, they would need to be very different from traditional requirements of rationality. It is difficult to see how such a requirement could help adjudicate

overly demanding and impossibly trivial requirements of rationality is a reason to distinguish requirement and appraisal.

4. Solving the Puzzle of Inconsistent Requirements

This section outlines how distinguishing requirement and appraisal allows us to solve the puzzle. Distinguishing requirement evaluations and appraisal evaluations means that we need not think that agents in cases like Logic 101 and Disagreement are subject to inconsistent requirements. Instead, when agents have misleading evidence for false beliefs about what rationality requires, we can distinguish the question of which epistemic attitudes are rationally required from the question of whether the agent deserves epistemic blame for adopting those attitudes. Since these distinct evaluations are determined by different kinds of considerations, it will sometimes be possible for agents to blamelessly adopt attitudes prohibited by rationality, and vice versa. This allows us to explain why agents who have misleading evidence for false beliefs about what rationality requires seem to be subject to conflicting requirements of rationality – these apparent requirements reflect the distinct claims that can be involved in evaluations of rationality. So, when rationality prohibits believing P, agents would be irrational in the sense of failing to meet the requirements of rationality were they to believe P. However, depending on the circumstances, they might also be rational in the sense of being epistemically blameless, if their belief P is held blamelessly.

Distinguishing these evaluations means that our three claims (Externalism, Support, Enkratic Principle) need not conflict. In the following section I argue that the Enkratic Principle can be associated with appraisal rather than requirement with minimal theoretical cost. This would mean that rationality requires, in all cases, that agents believe what their total evidence supports, and that they refrain from adopting any rationally prohibited attitudes. When agents have misleading evidence that rationality requires *not* believing P, when it in fact requires believing P, rationality requires the following beliefs:

I am rationally required to *not* believe P

P.

However, agents in such circumstances may be epistemically blameless if they refrain from believing P, provided they do so in an epistemically blameless way. Equally, in not believing P, they would be failing to meet the requirement to believe

---

between competing consistent belief sets, or guide agents towards more rational belief sets (see Lasonen-Aarnio, "Unreasonable Knowledge").

P. By distinguishing requirement evaluations from appraisal evaluations, we need not also think that they deserve blame for this failure.

At this point, one might wonder whether the proposed cure is worse than the disease. The problem that we are attempting to solve is that given a plausible set of commitments, and the possibility of misleading evidence about what rationality requires, rationality issues inconsistent requirements: for some P, it requires agents to both believe and *not* believe P. If we give up the Enkratic Principle, agents are never subject to inconsistent requirements but they are sometimes required to be level-incoherent. One might wonder if this is any better. However, although neither situation is ideal, there are reasons to think that inconsistent requirements of rationality are worse than requirements to be level-incoherent.

Firstly, it is impossible to both (fully) believe and not believe P. So, the inconsistent requirements are impossible to fulfil. Although some have thought that level-incoherent belief is impossible, but this is less clear.[35] Of course, provided we remain neutral on whether the Ought of rationality implies Can, rationality may sometimes require the impossible. However, whatever we think about this, inconsistent requirements are not only a problem for the agent subject to them. If rationality generates inconsistent requirements then rationality itself contains logical contradictions. This is a problem – contradictions entail triviality, at least assuming Standard Deontic Logic. Here is the derivation:[36]

1. OBp

2. O ¬Bp

3. O (Bp & ¬Bp) (by Agglomeration)

4. O (Bp & ¬Bp) → Oq (by Ought Entailment, Ex Falso Quodlibet)

5. Oq, for any q.

From inconsistent requirements ((1) and (2)), we can derive anything and everything as a rational requirement. A theory that permits this is unacceptably trivial. Of course, we could reject Standard Deontic Logic.[37] While some have taken

---

[35] At least some have thought it both possible and rational (Lasonen-Aarnio, "Enkrasia or Evidentialism?," "Higher-order Evidence and the Limits of Defeat;" Weatherson, *Normative Externalism*). For arguments that it is impossible, see (see Jonathan Adler, "Akratic Believing?," *Philosophical Studies* 110 (2002), Daniel Greco, "A Puzzle about Epistemic Akrasia," *Philosophical Studies* 167 (2014), David Owens, "Epistemic Akrasia," *The Monist* 85 (2002), Philip Pettit and Michael Smith, "Freedom in Belief and Desire," *Journal of Philosophy* 93 (1996), 448.

[36] *Ought Entailment*: Necessarily ((p → q) → (Op → Oq)); *Agglomeration*: Op & Oq → O (p & q); *Ex Falso Quodlibet*: (p & ¬p) → q.

[37] Indeed, some have thought there are independent reasons to do this. See G. Pigozzi, J. Hansen,

this option,[38] it is unnecessary to solve the puzzle, and sufficiently drastic that we have reason to seek others. A theory that gives up the Enkratic Principle offers recommendations for the agent that are no worse than inconsistent requirements, and it is significantly better for theoretical consistency. Not only this, but as the following section argues, it is the least theoretically costly of the available solutions and it can accommodate the usual motivations for the Enkratic Principle.

5. Giving Up the Enkratic Principle

This section defends the strategy of solving the Puzzle of Inconsistent Requirements by giving up the Enkratic Principle. First, this solution to the puzzle is the least theoretically costly option (§5.1), and, second, the main motivations for thinking of attitudinal coherence as a requirement of rationality are equally well, if not better, accommodated by a view that associates it with appraisal (§5.2).

5.1 A Comparative Bargain

First, giving up the Enkratic Principle is the least theoretically costly option. One might wonder whether it would be better to give up one of the other claims – Externalism or Support – than the Enkratic Principle. This subsection addresses this concern, arguing that while the puzzle could be solved by giving up any of the three claims, rejecting Externalism or Support involves significant theoretical costs.[39] Giving up the Enkratic Principle is a comparative bargain.

First, the costs of giving up Externalism. According to Externalism, the requirements hold independently of what the agents subject to them believe, what evidence they have, what they are in a position to know, and all other features of the agents' epistemic perspectives. So, if rationality requires you to refrain from believing contradictions, then it requires you to refrain from believing contradictions regardless of whether you *believe* that you are required to refrain

---

and L. van der Torre, "Ten philosophical problems in deontic logic. Normative Multi-Agent Systems," *Dagstuhl Seminar Proceedings* (2007); Sayre-McCord, "Deontic logic and the priority of moral theory," *Noûs* 20, 2 (1986). I have no quarrel with these arguments, since nothing I say here depends on the truth of SDL. Rather, my point is that the puzzle itself is not a reason to give up SDL.

[38] For example, David Alexander, "The Problem of Respecting Higher-Order Doubt," *Philosophers' Imprint* 13 (2013), Christensen, "Higher-Order Evidence," and Nicholas Hughes, "Dilemmic Epistemology," *Synthese* 196 (2019) propose thinking of cases such as Logic 101 and Disagreement as epistemic dilemmas. Not only does this involve an unnecessary rejection of SDL, it is less a solution to the puzzle than a way of saying that the puzzle does not require a solution.

[39] As previously mentioned, giving up more than one would involve even greater theoretical cost (the sum of the costs involved in giving up each of the individual claims).

from believing contradictions. Rejecting Externalism would mean denying this, and holding instead that what rationality requires depends on your perspective, such as how things appear to you, what you already believe, and the evidence you have available.[40] There are significant costs to accepting such a view.

First, denying Externalism implies that any attitude at all could, in principle, count as rational, provided that one has sufficient perspectival support for it. This is a problem because it makes the epistemic value of rationality mysterious, particularly for agents who already have many false beliefs. If being rational is merely a matter of making sense from your own perspective, there is no reason to think that being rational will lead you to valuable epistemic goods such as truth and knowledge.[41] It may even exasperate the epistemically negative consequences of false belief by further isolating us from epistemic goods.[42]

Second, rejecting Externalism would mean endorsing an error theory about traditional requirements, such that a statement of what rationality requires would be impossible to make in advance of considering the agent's precise situation. This would, surprisingly, turn cases of misleading evidence about what rationality requires into counterexamples to particular putative requirements of rationality.

---

[40] See Hartry Field, "Epistemology without Metaphysics," *Philosophical Studies* 143 (2009); John Gibbons, *The Norm of Belief* (Oxford University Press, 2013); Lord, *The Importance of Being Rational*; Kolodny, "Why Be Rational?;" Kiesewetter "'Ought' and the Perspective of the Agent," *Journal of Ethics and Social Philosophy* 5, 3 (2011); Benjamin Kiesewetter, *The Normativity of Rationality* (Oxford University Press, 2017); Kvanvig, *Rationality and Reflection: How to Think About What to Think*; Joseph Raz, "The Myth of Instrumental Rationality," *Journal of Ethics and Social Philosophy* 1; Daniel Whiting, "Keep Things in Perspective: Reasons, Rationality, and the A Priori," *Journal of Ethics and Social Philosophy* 8 (2014); Jonathan Way and Daniel Whiting, "Perspectivism and the Argument from Guidance," *Ethical Theory and Moral Practice* 20 (2017); Zimmerman, *Living with Uncertainty: The Moral Significance of Ignorance*.

[41] This mysteriousness invites the question of whether one should be rational, a question that has been well explored, without clear resolution, for the case of practical rationality (Broome, *Rationality Through Reasoning*; Kolodny, "Why Be Rational;" Raz, "The Myth of Instrumental Rationality." It thus may be considered an advantage if our solution precludes this question for epistemic rationality.

[42] This is a version of a familiar objection to coherentist theories of justification. Entirely false belief sets could count as justified if all that is required for justification is coherence (see Ernest Sosa, "The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge," *Midwest Studies in Philosophy* 5 (1980), 19, and particularly worrying in light of the observation that those who believe conspiracy theories often have beliefs that are largely consistent and well-supported from their perspective (see C. Thi Nguyen, "Echo Chambers and Epistemic Bubbles," *Episteme* 17 (2020), "Cognitive Islands and Runaway Echo Chambers: Problems for Epistemic Dependence on Experts," *Synthese* 197 (2018)).

This would constitute a significant revision to how we ordinarily think of rational requirements.[43]

Third, rejecting Externalism would prevent epistemic rationality from guiding agents towards epistemically better belief sets. If we think that rationality depends entirely on how things seem from one's perspective then no particular requirement of rationality will be genuinely binding for all agents. One unwelcome consequence of this is that requirements of rationality cannot provide information about which epistemic attitudes are rational, and so cannot be used by agents to guide their epistemic activities.

Similarly serious theoretical costs come with giving up Support. Various otherwise distinct accounts of epistemic rationality retain a commitment to it[44] – giving it up would be a radical overhaul for which we would need a very good reason. Moreover, while some have taken this option, their key motivation has typically been to preserve the Enkratic Principle.[45] For this move to be plausible, we would require independent reason to think that the Enkratic Principle expresses a more important aspect of epistemic rationality than Support. This independent reason has not been forthcoming. Furthermore, as the following subsection notes, some have motivated the Enkratic Principle by appeal to evidentialist principles related to Support,[46] suggesting that Support in fact has theoretical primacy.

So, there are significant theoretical costs to solving the puzzle by giving up either Externalism or Support. If there are comparatively minimal costs involved in giving up the Enkratic Principle, this would be a reason to prefer giving it up. The following subsection argues that the costs involved are indeed minimal, since the traditional motivations for thinking that the Enkratic Principle is a requirement of rationality can be accommodated by a view that thinks of it as associated with appraisal evaluations.

---

[43] See Lasonen-Aarnio, "Higher-Order Evidence and the Limits of Defeat," 332 on this point.

[44] See, amongst many others, Richard Feldman and Earl Conee, "Evidentialism," *Philosophical Studies* 48 (1985); Michael Huemer, "Does Probability Theory Refute Coherentism?," *Journal of Philosophy* 108 (2011); James Joyce, "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief," in *Degrees of Belief*, eds. Franz Huber and Christoph Schmidt-Petri (*Synthese*, 2009); Thomas Kelly, "The Rationality of Belief and Other Propositional Attitudes," *Philosophical Studies* 110 (2002); Lasonen-Aarnio, "Higher-Order Evidence and the Limits of Defeat;" Declan Smithies, "Moore's Paradox and the Accessibility of Justification," *Philosophy and Phenomenological Research* 85 (2012); Williamson, *Knowledge and its Limits*; Weatherson, *Normative Externalism*.

[45] Lasonen-Aarnio, "Enkrasia of Evidentialism?;" Littlejohn, "Stop Making Sense."

[46] For example, Worsnip, "The Conflict of Evidence and Coherence."

5.2 Accommodating Traditional Motivations

This subsection shows how traditional motivations for thinking that the Enkratic Principle is a requirement of rationality can be accommodated by a view that takes compliance with the principle to be defeasibly indicative of epistemic blamelessness.

Perhaps the most obvious such motivation is the idea that rationality demands attitudinal coherence, including the kind of inter-level coherence prohibited by the Enkratic Principle. In favour of this idea, it is tempting to appeal to the Moorean quality of belief combinations such as:

> P, but I ought not believe that P.

These are prohibited by the Enkratic Principle, and their prima facie irrationality may seem to suggest that the Enkratic Principle is a requirement of rationality,[47] or that level-incoherent belief could only be rational for fragmented minds.[48] Some have been so confident that the Enkratic Principle expresses a genuine requirement of rationality that they have relied on it to argue for far more controversial conclusions.[49] However, the intuition these claims rely on is insufficiently fine-grained to establish that the Enkratic Principle is a requirement of rationality. We cannot be sure of what it tracks – compliance with rational requirements, or some non-negative epistemic status, such as blamelessness.[50]

Another motivation for thinking that the Enkratic Principle is a requirement of rationality is the idea that it follows from an independent commitment to evidentialism, because one's total evidence could never support level-incoherent belief combinations. In this vein, Davidson suggests that epistemic akrasia – a kind of level-incoherence – manifests the epistemic "sin" of failing to believe what one's total evidence supports.[51] However, while it may be true that in most cases, when the evidence supports P it does not also support "I ought not believe P," there appear to be rare, but not impossible, exceptions. For example, when evidence is misleading about itself, the evidence would appear to support level-incoherent states. Typical

---

[47] See Smithies, "Moore's Paradox and the Accessibility of Justification."

[48] See Greco, "A Puzzle about Epistemic Akrasia," and Donald Davidson "Deception and Division," in *Problems of Rationality* (Oxford University Press, 2004).

[49] See Littlejohn's argument against Evidentialism ("Stop Making Sense"), and Titelbaum's argument for the impossibility of rational mistakes about rationality ("Rationality's Fixed Point").

[50] A similar point is often made about intuitions of justification and blamelessness (see Clayton Littlejohn, "A Plea for Epistemic Excuses," forthcoming in *The New Evil Demon Problem*, eds. Julien Dutant and Fabian Dorsch (Oxford University Press); Sutton, *Without Justification*; Williamson, "Justifications, Excuses, and Sceptical Scenarios," "Ambiguous Rationality," *Episteme* 14 (2017).

[51] Davidson, "Deception and Division," 201.

examples are cases in which agents have good evidence for P, but misleading evidence about their abilities to assess the evidence for P effectively (see Medicine, above). In such cases, your evidence appears to support both P and "my evidence does not support P." So, a commitment to evidentialism does not show that the Enkratic Principle is a requirement of rationality, although the rarity of such cases might show that it is usually good epistemic practice to avoid level-incoherent combinations, since they are usually not supported by one's total evidence. This is, however, compatible with the view that compliance with the Enkratic Principle is defeasible indication of blamelessness.

Another traditional motivation for thinking that the Enkratic Principle is a requirement of rationality is appeal to the idea that rationality requires good belief management, and good belief management requires compliance with the Enkratic Principle. For example, one might think that even if one's total evidence can sometimes support level-incoherent beliefs,[52] violating the Enkratic Principle could never be rational because rationality exerts normative pressure on us to revise our beliefs in response to our own beliefs about what we ought to believe.[53] Perhaps this is because doing so involves misusing one's higher-order beliefs,[54] failing to take seriously one's "conspicuous reasons,"[55] failing to reason appropriately from one's beliefs,[56] or failing to respect one's evidence by taking one's own judgments about what it supports seriously.[57]

However, complying with the Enkratic Principle is not always an example of good epistemic conduct. Complying with the Enkratic Principle only usually coincides with good epistemic conduct. Again, this is entirely compatible with thinking of the Enkratic Principle as a principle of appraisal, rather than requirement. Compare two possible students who take Logic 101. The first, call her Lazy, does not pay much attention in class or do the homework exercises. She exerts minimal intellectual effort, ignoring the arguments and testimony she receives in

---

[52] Some have argued that it cannot, because the levels cannot be separated so easily – P is itself evidence for "I ought to believe that P." For reasons to resist this view, and adopt a partially level-incoherent position (see Sliwa and Horowitz, "Respecting All the Evidence").

[53] For example, Kvanvig (*Rationality and Reflection*) sees this feature of the "egocentric predicament" as a central concern of epistemic rationality. This leads him to endorse a highly perspectivist account of rational requirement, which I have argued here comes with serious costs.

[54] David Christensen "Higher-Order Evidence;" "Rational Reflection," *Philosophical Perspectives* 24 (2010); "Disagreement as Evidence: The Epistemology of Controversy," *Philosophy Compass* 45 (2009); Littlejohn, "Stop Making Sense."

[55] Lasonen-Aarnio, "Enkrasia or Evidentialism."

[56] Horowitz, "Epistemic Akrasia."

[57] Worsnip, "The Conflict of Evidence and Coherence."

class, and would have nothing to say in response to them. When she does consider what rationality requires of her with respect to believing contradictions, she finds the idea that rationality could ever require her to believe contradictions "silly," and so refuses to believe it.[58] Although her beliefs are true and comply with the Enkratic Principle, this is not an example of good epistemic conduct. In reasoning "upstream"[59] she disregards her evidence, and she does for insuffcient reason (that she thinks the view is "silly"[60]). In this case the agent complies, but is not epistemically blameless.

Not only is Lazy not manifesting good epistemic conduct, she is doing significantly worse, epistemically, than other possible students who do comply with the Enkratic Principle. A second student, call her Diligent, considers what her teacher says and the arguments studied in class. She sees how they lead to the conclusion that rationality sometimes requires contradictory belief, and so she believes this. However, when she tries to believe the contradiction itself, she finds this difficult[61] – it seems so very counterintuitive. So, she has level-incoherent beliefs. However, she has arrived at these beliefs by managing her beliefs well. She believed the conclusion of a convincing argument, and she refrained from believing what seems counterintuitive. Perhaps she could have done better, but she has certainly managed her beliefs better than her fellow student, Lazy. Here, the agent is epistemically blameless despite not complying with the Enkratic Principle.

Not only this, but there are possible cases in which managing one's beliefs as well as one can means ending up with beliefs that violate the Enkratic Principle. Suppose that epistemologists in the future develop a device, call it the Excellent Evidence Evaluator, that can perfectly evaluate what one's evidence supports in any scenario. Everyone uses these devices and comes to depend on them. Your great-granddaughter has one of these devices, and her higher-order evidence suggests that her first-order evidence is misleading, when in fact it is not. In this case, her total

---

[58] Compare cases of reliable clairvoyants (see Bonjour, *The Structure of Empirical Knowledge* (Harvard University Press, 1985) – Lazy gets it right, but does not have good reason to believe that she is getting it right.

[59] As Kolodny ("Why Be Rational," 529) puts it. See also Schroeder's 'symmetry' objection to thinking of the practical Enkratic Principle as wide scope (Mark Schroeder, "The Scope of Instrumental Reason," *Philosophical Perspectives* 18 (2004), 339, which points out that only some of the ways one could bring oneself in line with the Enkratic Principle intuitively seem rational.

[60] While Lewis ("Letters to Priest and Beall") also responds to the arguments of dialetheism in this way, we can charitably assume that his philosophical experience means that he would more to say in response to the dialetheist's arguments.

[61] Likely, if, as some have argued, contradictory belief is impossible (see David Lewis, "Logic for Equivocators," *Noûs* 16 (1982); Worsnip, "The Conflict of Evidence and Coherence").

evidence appears to support both P and "my evidence does not support P." Believing what you believe to be supported by your evidence is an example of managing your beliefs well, so in trusting her Excellent Evidence Evaluator, your great-granddaughter is managing her beliefs well, and so is blamelessly violating the Enkratic Principle.[62]

Conformity to the Enkratic Principle is thus not necessary for good belief management. However, we can preserve the intuitive idea that good epistemic agents *usually* comply with the Enkratic Principle by understanding compliance as defeasible indication of epistemic blamelessness. Consideration of Lazy shows us that if compliance is such an indication, it must be defeasible. Consideration of Diligent and the Excellent Evidence Evaluator shows us that compliance with the Enkratic Principle is not necessary for epistemic blamelessness.

Finally, a key motivation for thinking that the Enkratic Principle is a requirement of rationality – one that is much better accommodated by the proposed solution than the orthodox view – is the idea that to be rational *is* to be epistemically blameless.[63] A tempting thought is that agents who comply with the Enkratic Principle are immune from epistemic blame – perhaps because maintaining level-coherence constitutes good epistemic conduct, manifests epistemic virtues, or is a reasonable thing to expect of rational agents. However, the cases of Diligent and the Excellent Evidence Evaluator show this to be not quite true. Compliance with the Enkratic Principle is only usually, not always, indicative of epistemic blamelessness. So, the grain of truth in this motivation is easily accommodated by a view that takes the Enkratic Principle to be indicative of epistemic blamelessness.

In sum, giving up the Enkratic Principle involves only minimal theoretical costs. This makes it a preferable solution to giving up either of the other two claims (Externalism and Support). The final section outlines the correct evaluation of agents in cases such as Logic 101 and Disagreement.

---

[62] This possibility is consistent with what some defenders of the orthodox view say. Horowitz, for example, argues that level-coherence is necessary for rational belief in the majority of cases, but concedes that there are some cases in which the higher- and first-order evidence support incompatible propositions, and rationality does not require level-coherence ("Epistemic Akrasia," 735-40). However, this position is not consistent with the view that the Enkratic Principle is a requirement of rationality, holding in all cases.

[63] See Kvanvig, *Rationality and Reflection*; Lord, *The Importance of Being Rational*; Wedgwood, *The Value of Rationality*.

Claire Field

6. Evaluating Misled Agents

This section illustrates the evaluation of agents with misleading evidence about what rationality requires offered by the proposed solution. When agents have misleading evidence that rationality requires *not* believing P, but in fact rationality requires believing P, then rationality requires the following beliefs:

> I am rationally required to *not* believe P
>
> P.

The agent's total evidence supports that rationality requires not believing P, so she should believe this. Rationality also requires believing P, so she should believe P. However, she may be epistemically blameless is she does not believe P, depending on how she conducts herself epistemically. The agent is required to have a combination of beliefs that violates the Enkratic Principle.

We avoid the possibility of rationality generating inconsistent requirements because the Enkratic Principle is not a requirement of rationality. If the agent does comply with the Enkratic Principle, this may indicate epistemic blamelessness, although we cannot be sure of this without knowing more about her and her situation.

I have argued that the best way to solve the Puzzle of Inconsistent Requirements is to distinguish requirement evaluations and appraisal evaluations, and think of compliance with the Enkratic Principle not as a requirement of rationality, but rather defeasible indication that the agent is epistemically blameless. This solves the puzzle with minimal theoretical cost, while vindicating the motivations that have contributed to the orthodox view of the Enkratic Principle's status.[64]

DISAGREEMENT AND DEEP AGNOSTICISM

Eric GILBERTSON

ABSTRACT: One defense of the "steadfast" position in cases of peer disagreement appeals to the idea that it's rational for you to remain *deeply agnostic* about relevant propositions concerning your peer's judgment, that is, to assign no credence value at all to such propositions. Thus, according to this view, since you need not assign any value to the proposition that your peer's judgment is likely to be correct, you need not conciliate, since you can remain deeply agnostic on the question of how the likelihood of your peer's judgment bears on the likelihood of your own. This paper argues that the case for deep agnosticism as a response to peer disagreement fails. Deep agnosticism (as a general thesis) implies that it is sometimes permissible to withhold judgment about whether there is a *non-zero* chance of a proposition's being true. However, in cases of disagreement where deep agnosticism is supposed to support the steadfast position, such withholding isn't rational. This is because of constraints placed on rational credence by objective probability or chance, which ensure that rational credence adequately reflects strength of evidence.

KEYWORDS: epistemic peerhood, disagreement, deep agnosticism, credence

1. Introduction

Consider the following case (from Grundmann[1]):

> *Election*: There is a very close presidential race between two competing candidates, X and Y. After the people have voted, two equally renowned election forecasters using different, but equally representative samples come up with conflicting predictions regarding who won. Forecaster A predicts that X has won by a narrow margin. Forecaster B predicts that Y has won by a narrow margin.

In this case it certainly seems as if A, upon recognizing that B has made a different judgment, acquires evidence that defeats her initial justification for the belief that X has won. Thus it seems to be a case in which disagreement with an acknowledged *epistemic peer* (roughly, someone who has equally good evidence and is equally good at reasoning on the basis of such evidence) requires one to reduce confidence in one's initial judgment, and perhaps even to suspend belief.

In suspending belief on the matter of who won the election, A would give equal weight to the possibility that A's judgment is incorrect and to the possibility that B's judgment is incorrect.[2] Thus *Election* and other similar cases involving peer

---

disagreement are taken by some to support the *equal weight view*: roughly, that in cases of disagreement with epistemic peers you should give equal credence to the possibility that you have made a mistake as to the possibility that your peer has.[3] Depending on precisely how the equal weight view is understood, the view known as *conciliationism* may in one sense go further, as it holds that you are rationally required to adjust your credence in the disputed proposition in the direction of your peer in such cases.[4]

But is it really rationally required to suspend judgment, or to adjust your credence, in such cases? These positions have seemed to some to depend on a principle of indifference which is suspect.[5] That is indeed a legitimate concern. Roger White, however, provides a strong defense of principles of indifference, arguing that, at least in the case of a multiple partitions problem, the apparently absurd results do not depend on them.[6] So these principles may support the equal

Vol. 4, ed. T. S. Gendler and J. Hawthorne (Oxford: Oxford University Press, 2013), 57-81, 57) argues that there are cases where suspension of belief is clearly epistemically permissible regardless of whether it's rational to assign any credence to the relevant propositions. However, these are cases where the subject is completely ignorant about the matter in question. For example, Friedman suggests that it's permissible to suspend belief about what the price of copper will be in 100 years, or whether the Hill 50 Gold Mine was Australia's most profitable mine between 1955 and 1961. In the sort of case of peer disagreement we are considering, however, one is not completely in the dark. In such circumstances, suspension of belief seems to require giving equal weight.

[3] It should be noted that the equal weight view as stated is somewhat imprecise. In particular, as Jehle and Fitelson ("What is the 'Equal weight view'?," *Episteme* 6, 3 (2009): 280–293) have shown, there are a number of different possible precisifications of the view which clearly specify how someone should update her credence upon discovery of disagreement with an epistemic peer. I will however rely on the less precise notion of giving equal weight, as I think that however the notion is precisified it will not affect my argument; the crucial claim that I defend is that whether or not *regarding someone as an epistemic peer* entails *giving equal weight* to the respective judgments, it entails assigning some (perhaps indeterminate) credence value to relevant propositions concerning your and your peer's judgments. Hence, precisely what the rule for updating should be on this view is a further question.

[4] If it is possible – and rationally permissible – to give equal credence to the possibility that your and your peer's judgments are incorrect without adjusting your credence in the disputed proposition, then conciliationism is not implied by equal weight. But if this is not rationally permissible – perhaps because giving equal weight requires "splitting the difference" between degrees of belief – then conciliationism is implied by equal weight.

[5] See, for instance, D. Gillies, *Philosophical Theories of Probability* (London: Routledge, 2000); B. van Fraassen, *Laws and Symmetry* (Oxford: Oxford University Press, 1989).

[6] Roger White, "Evidential Symmetry and Mushy Credence," in *Oxford Studies in Epistemology*, eds. T. S. Gendler and J. Hawthorne (Oxford: Oxford University Press, 2013), 161–186.

weight view or conciliationism, after all. Moreover, Adam Elga presents a "bootstrapping" argument for the equal weight view,[7] and thus against the view that it's rationally permissible to remain steadfast in the face of disagreement with epistemic peers[8]—again, suggesting that some form of conciliationism is correct.

The case for the equal weight view and conciliationism has been challenged in another way, however, by Jason Decker, who argues that the arguments for these views overlook the possibility of *deep agnosticism*, by which he means a kind of agnostic credal state which involves a partially-defined ("gappy") function that fails to assign any value to certain propositions.[9] Thus Decker defends the rationality of your reacting to peer disagreement with such agnosticism—about the likelihood that your peer's judgment, rather than your own, is correct—rather than with conciliation or suspension of belief. By withholding judgment on this question, Decker argues, you can remain steadfast in your initial belief: because you take no stand on how likely it is that you peer's judgment is correct, your recognition of that judgment need not diminish your confidence in your own.

In this paper, I argue that the case for deep agnosticism as a response to peer disagreement fails. First, deep agnosticism (DA)[10] implies that it is sometimes rationally permissible to refrain from assigning any positive credence value at all, whether determinate or not, to a proposition. Given plausible constraints placed on rational credence by objective probability or chance, this implies that it is permissible to withhold judgment about whether there is a *non-zero* chance of the proposition's being true. However, in those cases of disagreement in which deep agnosticism is supposed to be rational, such withholding isn't rational. Second, I argue that it is inconsistent to assign a high credence to the proposition that another

---

[7] Adam Elga, "Reflection and Disagreement," *Noûs* (2007): 478–502.

[8] Decker ("Disagreement, Evidence, and Agnosticism," *Synthese* 187 (2012): 753-783) labels this view "Millianism."

[9] Decker, "Disagreement, Evidence, and Agnosticism;" Decker, "Conciliation and Self-incrimination," *Erkenntnis* 79 (2014): 1099–1134.

[10] It's important to understand that Decker sometimes uses the term 'deep agnosticism' to refer not to this general position (which has nothing essentially to do with disagreement), but instead to a more specific position that licenses credal gaps with respect to specific propositions concerning a peer's evidence or her judgment. To make matters yet more confusing, Decker also uses the term to refer to the *credal state* one is in with respect to propositions which one's credence function has assigned no value. In what follows, I'll frequently rely on context to indicate which sense of the term is intended. However, I'll use 'DA' to refer to the more general position and 'DA$_d$' to refer to the more specific position Decker defends, that is, that it's rationally permissible to assign no credence value at all to relevant propositions concerning an epistemic peer's judgment in cases of disagreement. I think it is clear, moreover, in discussion of certain instances of peer disagreement, when I intend to refer to the more specific thesis and when I intend the "credal state" sense.

person is an epistemic peer while assigning no credence at all to the proposition that her evidence, and her ability to handle that evidence, is as good as your own. Since (I argue) the view that deep agnosticism is a rational response to peer disagreement implies that this is rational, it should be rejected.

The failure of deep agnosticism shows that the proponent of the "steadfast" position, according to which it is rationally permissible to stick to one's initial judgment in cases of peer disagreement (i.e., to refrain from conciliation), must somehow accommodate the fact that in such cases your acknowledgement of an epistemic peer as such requires you to accept that there is a significant (objective) chance that your peer's judgment is correct, and that this provides you with some reason to think that your own judgment is incorrect.[11] The challenge for the proponent of the steadfast position is to show that such an acknowledgement does not in fact require conciliation. If I am right, then this challenge cannot be avoided simply by appealing to the possibility of withholding judgment about the status of epistemic peers' judgments, as deep agnosticism maintains.

The paper proceeds as follows. First, in §2, I explain Elga's bootstrapping argument in support of the equal weight view. In §3, I explain Decker's criticism of Elga and his defense of deep agnosticism (DA$_d$), and I present my main argument against the view. In §4, I consider whether deep agnosticism may provide a satisfactory solution to the problem of multiple partitions, and I argue that it does not. In particular, I argue that deep agnosticism has certain unacceptable implications in this context. Moreover, I argue that Lewis's Principal Principle, which holds that a rational agent conforms her credences to known objective chances, provides further support for the claim that it's rationally impermissible to assign no credence value to the proposition that your epistemic peer is more likely to be correct than you. Finally, in §5, I respond to an objection to my argument's reliance on the notion of epistemic peerhood.

Before proceeding I should explain more precisely how I am understanding the notion of an epistemic peer. First, I take epistemic peerhood to be *domain-relative*. That is, X is an epistemic peer of Y with respect to a certain issue or subject matter (the morality of abortion, climate change, mathematics, modal logic, etc.). As issues in different domains may be independent of one another, it's possible to regard someone as a an epistemic peer in some cases of disagreement and not in others. I'll sometimes speak of X's being an epistemic peer of Y with respect to a certain

---

[11] Of course, it may be that mere recognition of your fallibility on the relevant matter also provides you with such a reason. Whether or not this is so, the point here is that disagreement with an acknowledged epistemic peer provides an independent reason to think that your judgment is incorrect.

proposition *p*. It would be more accurate to attribute peerhood with respect to some relevant domain of truths which includes *p*/*not-p*. But I assume then that epistemic peerhood holds relative to propositions, if it holds relative to domains. So it is not strictly improper to formulate things as I do, and it is convenient to do so. It is also important to understand that epistemic peerhood, as I understand it, doesn't require—although it is compatible with—*identical evidence*; it is sufficient, as far as the disputants' first-order evidence on the matter in dispute goes, that their evidence is *equally good, though not identical*.[12] In the election case, for example, the two forecasters rely on different evidence, but their evidence involves equally representative samples of voters. Finally, epistemic peerhood is understood as involving *equality of higher-order evidence*, in the following sense. Provided X and Y are epistemic peers, the dispute-independent evidence regarding X's and Y's epistemic credentials and trustworthiness on the matter in question no more supports the claim that X has made an error than that Y has. This condition (sometimes referred to as the Independence principle) is important, as it requires that you assess others for epistemic peerhood without relying on your particular reasoning or evidence regarding the disputed matter itself. Reliance on such reasoning or evidence would seem question-begging, as it would allow you to dismiss someone's judgment, and regard them as an epistemic inferior, simply because it conflicts with your own, when you have independent reasons to regard them as being in an equally good position to judge (on the matter in question).[13]

So my understanding of epistemic peerhood is this:

> X and Y are *epistemic peers* with respect to a proposition *p* iff (i) X and Y have the same, or equally good, relevant (first-order) evidence regarding *p* and (ii) X and Y are equally competent in reasoning based on this evidence and (iii) the higher-order evidence available to X and Y concerning X's and Y's epistemic credentials

---

[12] Cf. Grundmann, "Why Disagreement-Based Skepticism Cannot Escape the Challenge of Self-Defeat," 3.

[13] Cf. David Christensen, "Conciliation, Uniqueness, and Rational Toxicity," *Noûs* 50, 3 (2016), fn. 7: "Conciliatory views require an agent to determine the seriousness with which she takes another's disagreement in a way that's *independent* of the agent's own particular reasoning on the disputed matter. The thought behind this is that insofar as disagreement can raise doubts about the cogency of one's reasoning on a certain matter, it would be question-begging to dismiss those doubts in a way that relied on the very reasoning under dispute." It should be noted that not everyone accepts this. See, for instance, Jennifer Lackey, "What Should We Do When We Disagree?," in *Oxford Studies in Epistemology*, ed. T.S. Gendler and J. Hawthorne (Oxford: Oxford University Press, 2010); S. Benjamin, "Questionable Peers and Spinelessness," *Canadian Journal of Philosophy* 45, 4 (2015): 425–444; T. Kelly, "Disagreement and the burdens of judgment," in *The Epistemology of Disagreement: New Essays,* ed. D. Christensen and J. Lackey (Oxford: Oxford University Press, 2013): 121–166.

and trustworthiness (and specifically, the fact that X and Y believe *p* or *not-p* based on the same, or equally good, evidence) equally supports the proposition that X's judgment regarding *p* is mistaken and the proposition that Y's judgment regarding *p* is mistaken.[14]

Finally, let me be clear about my aims. I do not intend to resolve the dispute between proponents of the steadfast position and conciliationists. While my argument, supposing it succeeds, may provide some support for conciliationism, I don't think that such support is very strong. The point of the argument is to show that one particular attempt to defend the steadfast position, which depends on deep agnosticism, fails; and to show that this follows from certain assumptions about what is entailed by regarding someone with whom one disagrees as an epistemic peer. If this is right, it clearly does not mean that the steadfast position is simply indefensible, or that it has any untoward consequences which are independent of deep agnosticism. It may be rationally permissible to remain steadfast in cases of peer disagreement, though not because this is entailed by deep agnosticism. The position that I defend here is that deep agnosticism does not explain why remaining steadfast is coherent and rational, because deep agnosticism is (in these circumstances) irrational.

2. Elga's Bootstrapping Argument for the Equal Weight View

Elga's argument for the equal weight view appeals to the following case:

> *The horse race.* You and an acknowledged epistemic peer are watching a horse race. After the race is run, you're convinced that Cavonnier won and your peer is convinced that Grindstone won.

Elga writes:

> When you learn of your friend's opposing judgment, you should think that the two of you are equally likely to be correct. For suppose not – suppose it were reasonable for you to be, say, 70% confident that you are correct. Then you would have gotten some evidence that you are a better judge than your friend, since you would have gotten some evidence that you judged this race correctly, while she misjudged it. But that is absurd.[15]

---

[14] It is also worth noting that it is important that the notion of epistemic peerhood, in particular (iii), be understood in terms of the bearing of higher-order evidence, rather than in terms of *total* evidence, on likelihood of correctness. For if it's understood in terms of (Xs or Ys) total evidence, then the equal weight view would follow trivially from evidentialism. (Thanks to an anonymous referee for helping me see this point.)

[15] Elga, "Reflection and Disagreement," 486.

Elga's thought is that if we consider a series of such events, where in each it is thought to be reasonable for you to be more confident that you got it right than that your peer did, this leads to the conclusion that you get inductive evidence that you're a better judge, thus bootstrapping your way to expertise.[16] Since it is absurd to think you could do this, it follows that it's not rational, after all, to be more confident in your own judgment than your peer's. Supposing, for instance, that you were more confident in your own judgment because you judged that you have a better track record than your peer, you would thereby take the fact that you disagree as evidence that s/he's made more mistakes. That would violate Independence. Without some antecedent reason to think you're a better judge, the fact that you and your peer disagree isn't evidence that s/he's made more mistakes.

So, Elga concludes, you should instead give equal weight to both judgments and think that you're no more likely to be right than she is. And the same goes for other disagreements with epistemic peers. Any case in which you and someone you take to be an epistemic peer disagree on a claim based on the same evidence is one in which you have just as much reason to think you have erred as that she has. For if it were reasonable for you to give more weight to the proposition that your peer has erred in such a case, then it would be possible for you to gain knowledge, based on a series of such disagreements, that you are a better judge than your peer. Since this isn't possible, it's only reasonable for you to give equal weight.

3. Decker's Criticism of Elga and Defense of DA: Avoiding the Bootstrapping Problem without Giving Equal Weight

Let's suppose that Elga is right that it isn't possible to bootstrap your way to expertise by sticking to your guns in the face of peer disagreement. Decker agrees. He does not agree, however, that giving equal weight (i.e. judging that your and your peer's judgments are equally probable) is required in order to avoid the problem. In his view, you can stick to your initial judgment and take yourself to be justified in doing so without being committed to thinking that you're *more likely to be correct* than your peer; you can simply withhold judgment about this. Decker claims that it's a

---

[16] It may be useful to compare the "easy knowledge objection" to reliabilism: According to reliabilism, it's possible to acquire knowledge e.g. that a thing is red (by using your reliable faculty of color vision) even if you don't know that the method by which you arrived at such knowledge is reliable. But, according to reliabilism, in coming to know that x is red you thereby come to know that your color vision has operated correctly. So, by acquiring such knowledge on many occasions, you come to know that your color vision is reliable. But, the objection goes, it is absurd to think you can acquire knowledge of the reliability of your color vision in this way. (Cf. Elga, "Reflection and Disagreement," 488.)

mistake to think that to remain steadfast is to give any *extra* weight to your judgment, beyond what's involved in simply making the judgment. Consequently, even if there's a series of similar cases, and you confidently stick to your judgment throughout, you need not take this as inductive evidence that you're a better judge than your peer.

As I indicated earlier, Decker's argument rests on a position he calls *deep agnosticism* (DA). According to it, some propositions are such that our credence functions assign no value (or value range) at all. In the present case, the thought is that you might have no view at all on how to assess the judgments of you and your epistemic peers. Consequently, your credence function might assign no value to the proposition that your peer's judgment is as likely, or more likely, to be correct as yours. (Decker writes, "You might well feel that one would need a kind of external God's-eye perspective of the situation in order to make such an assessment, and, lacking that perspective, you might find yourself *deeply agnostic* on who has done a better job at the task of judging."[17]) In Elga's view, sticking to your guns in these cases entails taking such a stand: if you stick with your first-round judgment in the face of peer disagreement, you thereby take your judgment to be more likely to be correct than your peer's. This involves a second-order judgment. In Decker's view, sticking to your guns need not involve any such second-order judgment, as you can simply withhold judgment about who's more likely to be correct. Moreover, if you withhold such judgment, this does not mean you are required to give *equal* weight to you and your peer's being correct; for you can withhold judgment on whether they are *equally* likely to be correct, too.

So the issue is whether deep agnosticism provides an adequate defense of the steadfast position. Is it indeed rational to be deeply agnostic about relevant propositions concerning the judgments of your epistemic peers? First, I propose that

---

[17] Decker, "Disagreement in Philosophy," in *The Palgrave Handbook of Philosophical Methods*, ed. Chris Daly (MacMillan, 2015), 151. Decker presents another case in which you and a peer are official judges at a diving competition: You've been given the option of scoring the other judges on *their* performances, in addition to scoring the divers on theirs. Moreover, you have the opportunity to revise your scores for the divers on the basis of your knowledge of the other judges' scores. You might, for instance, after finding out your peer's score differs from your own, split the difference and revise your score accordingly. Indeed this might seem like the only rational response to your peer's score. Decker argues otherwise:

When one sticks with one's first-round judgment in the face of peer disagreement, one is not somehow getting evidence (or taking oneself to have somehow gotten evidence) that one is a better judge than one's peer. One isn't taking a stand on that issue one way or another. And that's fine. Contrary to the impression that is sometimes given in the literature on disagreement, it's okay to be deeply agnostic on propositions. ("Disagreement in Philosophy," 151.)

we grant that it may be reasonable to be deeply agnostic about *some* propositions regarding you and your epistemic peers.[18] It's important to see that even supposing this is the case, it doesn't follow that you can take someone to be your epistemic peer, in circumstances in which the relevant higher-order evidence no more supports the claim that you've made an error than that she has, while at the same time withholding judgment about *whether her judgment is reasonable*. In fact it is plausible that to take someone to be an epistemic peer in these circumstances *is* to take a stand on this question.[19] And it is even more plausible that in order for you to regard someone as an epistemic peer you must assign *some positive credence value* to the proposition that her judgment is reasonable. Moreover, this apparently gives you some reason to think that your peer's judgment is *correct*. For in judging that your peer's judgment (that *p*) is reasonable, you apparently acknowledge that your peer has some reason (given her evidence and her starting points) to think that *p* is true. Thus you appear to to be committed to thinking that there is *some* reason to think *p is* true.[20] But, if you take a stand on *that*, then you can't rationally refrain from judgment about how her judgment bears on the likelihood that yours is correct,

---

[18] I don't know whether it *is* reasonable, but I'll grant that it is, since I'm only interested in whether DA can provide sound defense of the steadfast position, not whether deep agnosticism is *ever* rational. For what it's worth, one proposition to which it seems it's rationally permissible for your credence function to assign no value (or no value range), is the proposition that your peer is deeply agnostic about some proposition *p*.

[19] More precisely, it is plausible that regarding someone as an epistemic peer requires taking such a stand *provided you take your own judgment to be reasonable*. I am assuming that in asking what our response to peer disagreement should be we are interested only in such cases. (Thanks to an anonymous reviewer for pointing out the need to clarify this point.)

[20] It might be argued that (rightly) regarding another as an epistemic peer does not entail that you have some reason to think your peer's judgment is correct, at least if some form of permissivism is true. Suppose, for example, two ancient Greek philosophers, who both thought the world is finite in duration, disagreed about whether it is more than 10,000 years old. Because they lacked significant evidence that would bear on the issue, both views appear reasonable, and each is reasonable to think the other's view is, from his point of view, reasonable. Is this a reason for either to think that the other's view is at all likely to be correct? It might seem it is not. (Thanks to an anonymous reviewer for this point and for offering this example to illustrate it.) But while I grant that neither individual has a reason to think that the other's judgment is *more likely than not* to be correct, I claim that it is irrational for either individual to regard the other's view as reasonable while assigning *no credence value* to the proposition that it is correct. To do so would be to withhold judgment not only on whether the view is correct but on whether it is reasonable. In committing to the view that the other's view is reasonable, each individual rejects the claim that there's no reason at all to think the other's judgment is correct – even if (from their point of view) there may be no such *weighty* reason, or any reason that is particularly worth considering in certain contexts.

since confidence in her judgment implies (some degree of) doubt about your own.[21] Taking a stand on the epistemic standing of your peer requires taking a stand on this question.

My suggestion, then, is that the proponent of the equal weight view should deny that it is coherent to hold both (i) that it's rational for you to withhold judgment on who is a better judge[22] in this case (and on whether both judges are equally good), *and* (ii) that it is *not* rational to withhold judgment on *whether the other disputant is an epistemic peer*. Hence, if it is rationally required to regard the other disputant as an epistemic peer[23] this constrains rational credence in related propositions about her, and your, epistemic standing, including the proposition that your peer's judgment is more likely to be correct. In particular, it would be

---

[21] Here it may seem I am assuming Uniqueness—i.e., that given one's total evidence, there is a unique rational doxastic attitude that one can take to any proposition. Since uniqueness is controversial, this assumption would be unwarranted. But my claim that taking a stand on the likelihood that your peer's judgment is correct rationally requires taking a stand on the likelihood that yours is incorrect does not imply uniqueness. I do not assume that you and your peer cannot be equally *rational* in this case; I do however assume that your *judgments* can't both be *correct*. Hence, to the extent that you judge that your peer's judgment is likely to be correct, you are rationally required to reduce confidence in the correctness (though not necessarily the rationality) of your own judgment. (I return to the issue of uniqueness and epistemic peerhood in §5.)

[22] As an anonymous referee notes, the question of *who's a better judge* (in this particular case, or in general) is not the same as the question of *who's more likely to be correct*. It's possible, then, to be agnostic about one of these and not the other (i.e., taking a stand on one doesn't require taking a stand on the other). However, if 'better judge' is not understood in terms of track record, or somehow in terms of the likelihood of being correct (either in general, or in the particular matter at hand), then it seems to me it must be understood in terms of reasoning capability; and in that case, given the understanding of epistemic peerhood I'm relying on, I think it would be straightforwardly inconsistent to regard a disputant as an epistemic peer while withholding judgment on the question of whether s/he's a better (or worse, or equal) judge. So unless there is some other notion of what makes one judge *better* than another, which doesn't rely on considerations of track record or reasoning capability, I claim it's not rational to be agnostic on either question, provided the other is regarded as an epistemic peer.

[23] It might be argued that in many of the cases of disagreement under consideration you may remain steadfast by being deeply agnostic about *whether the other disputant is an epistemic peer*. I don't deny that this may *sometimes* be permissible and that it would allow you to remain steadfast. But I think that the typical sort of case which is of interest to philosophers is one where you have good (perhaps conclusive) reasons to take someone to be an epistemic peer – that is, it's one where your evidence doesn't permit withholding judgment about this. Moreover, Decker never suggests it is permissible to be agnostic about *whether one's peer is an epistemic peer.* He is interested rather in deep agnosticism with respect to certain propositions which presuppose epistemic peerhood, and so understanding of which actually requires you take the other to be an epistemic peer.

inconsistent for you to remain steadfast while adjusting your credences in such propositions.

Let me now present my argument more precisely and clarify my assumptions.[24]

> (P1)   If you (outright) believe that p, then you assign a credence value of 0 to the proposition that there's no chance that p is true (thus you assign credence 1 to the proposition that there's some chance p *is* true). (To put it another way, believing that p requires that it is doxastically necessary (for the believer) that there's some chance that p is true.)[25]

> (P2)   If DA$_d$ is correct, then in cases of peer disagreement regarding p, it is rational for you to withhold judgment on whether there's a zero chance that your peer's judgment is correct. (DA$_d$ implies that it's rational to assign *no value* to the proposition that your peer's judgment is correct, and to the proposition that it's more, or less, likely to be correct than yours is.)

> (P3)   If it's rational for you to withhold judgment on whether there's a zero chance that your peer's judgment is correct, then – since in such cases you take your disputant to have equally good evidence, and to be equally capable of reasoning on the basis of such evidence – it is rational for you to withhold judgment on whether there's a zero chance that your own judgment is

---

[24] I should note that I'm focusing on cases of disagreement that involve *outright* belief, rather than degree of belief (supposing the latter notion makes sense). While this is not the only kind of case of peer disagreement that may be of interest, I think it is typical. My argument doesn't show that deep agnosticism about judgments of an acknowledged epistemic peer is incoherent in cases where one has less than outright belief, I think it strongly suggests that it's incoherent even in such cases. Regardless, I take it that establishing incoherence in cases of outright belief is significant.

[25] Dylan Dodd ("Belief and Certainty," *Synthese* 194 (2017): 4597-4621) defends a credence-doxastic possibility link: If S has a positive credence in p, then p is doxastically possible for S. My claim that if S has positive credence in p sufficient for outright belief that p, then the proposition that there's no chance that p is not doxastically possible for S. This is a kind of credence-doxastic possibility-chance link. (Dodd defines doxastic possibility as follows ("Belief and Certainty," 4604): p is doxastically possible for S iff S's doxastic state leaves it open that p is true—iff S's point of view or perspective is consistent with p.) He argues that if you believe that p then p is doxastically necessary for you. I take no stand on this, since I do not wish to commit to the view that is arguably implied by it (given further assumption that S's having a positive credence in p implies that p is doxastically possible for s), that believing that p requires having credence 1 in p. (Recent proponents of this view include Roger Clarke, "Belief is Credence 1 (in Context)," *Philosophers' Imprint* 13, 11 (2013): 1-18, Dodd, "Belief and Certainty," and Dominik Kauss, "Credence as Doxastic Tendency," *Synthese* (2018), https://doi.org/10.1007/s11229-018-01938-4.) My argument is compatible with the Threshold (Lockean) View that outright believing a proposition p consists in assigning a sufficiently high credence, and with the (less popular) view that outright believing p requires assigning credence 1 to p.

correct.

(P4) If DA$_d$ is correct, then in cases of peer disagreement regarding p, it is rational for you to withhold judgment on whether there's a zero chance that your own judgment is correct.

(C) Since it is not rational for you to withhold judgment on whether there is a zero chance that your own judgment is correct while remaining steadfast (as it is not doxastically possible for you that there is a zero chance that your judgment is correct, in such circumstances), DA$_d$ is incorrect.

Let me offer a few additional remarks in defense of the argument. First, regarding P1, it might seem there are clear counterexamples to the claim that believing that p requires assigning credence 0 to the proposition that there's no chance that p. Suppose, for instance, that I carry out a formal logic proof and conclude on this basis that the argument in question is formally valid. I acknowledge that I might have made a mistake and even that the argument might not in fact be valid. It seems, then, that – assuming that this is a domain of necessary truths – I must acknowledge that there's a possibility that there is in fact *no chance* that the argument is valid. That is, it seems that this is a doxastic possibility for me – something my beliefs leave open. However, I think that as long as I outright believe that the argument in question *is* valid, my belief does in fact rule out the possibility that there's no chance of its being valid.[26] First, note that this claim is much weaker than the claim that I believe that p only if p is doxastically necessary for me. Since the latter claim is plausible,[27] we appear to have good reason to accept the weaker claim. In addition, I maintain that your having credence 1 in the proposition that there's some chance that p is compatible with acknowledgement of your fallibility on this matter. So your having credence 0 in the proposition that there's no chance that the argument is valid is also consistent with your acknowledgement of your fallibility regarding this proposition.

Perhaps the most obvious reason why someone might resist the idea of assigning credence 1 in such cases is that this is seen as incompatible with acknowledgement of fallibility. However, my acknowledgement of my own fallibility regarding p does not require that it is rationally permissible for me to outright believe that p while also believing that it's possible that there's no chance that p. For, as Rosenkranz[28] has convincingly argued, acknowledgement of fallibility

---

[26] Cf. Dodd ("Belief and Certainty," 4609): "[I]t is essential to representations that if a representation assertorically represents the world as being a certain way, it doesn't leave open the possibility that the world is otherwise."

[27] For a defense of this claim, see Dodd, "Belief and Certainty."

[28] Sven Rosenkranz, "Fallibility and Trust," *Noûs* 49, 3 (2015): 616-641.

regarding *p* can be understood in terms of a *lack* of relevant higher-order beliefs about one's immunity to error, together with a readiness to revise one's belief regarding *p* in light of future evidence; and such attitudes are perfectly consistent with assigning credence 1 to *p*.

Finally, on P3, the thought is that you have no relevant evidence which could warrant different attitudes toward your peer's judgment and your own. Hence, if it's permissible for you to withhold judgment about some proposition concerning the likelihood of your peer's judgment, then it must also be permissible for you to withhold judgment about a parallel proposition concerning the likelihood of your own judgment. The only possible exception to this, as I see it, would be that you are required to refrain from withholding in your own case simply because it is yours. But – whether or not this idea is tenable – it is not something that DA$_d$ sanctions. (Recall that Decker explicitly denies that deep agnosticism involves giving such extra weight to your own judgment.)[29]

I conclude that deep agnosticism about your epistemic peer's judgment is not a rational response to your disagreement, as long as you remain steadfast. Deep agnosticism commits you to withholding credence in the proposition that your own judgment has a non-zero chance of being correct – which you cannot do if you stick to your guns.

In the next section, I will turn to a different set of considerations which may seem to support deep agnosticism in cases of peer disagreement. In particular, deep

---

[29] Levy ("The Surprising Truth about Disagreement," *Acta Analytica* (2020), https://doi.org/10.1007/s12136-020-00437-x) argues that different kinds of disagreement carry different kinds of information, and some (viz., "partisan" disagreements on which left and right seem intractably opposed, e.g., over the morality of abortion, or gun control) do not carry any information about likelihood of error; hence, they don't provide any pressure to conciliate. If this is right, then it may seem to undermine P3. While I do not have the space here discuss the details of Levy's account, or to provide an adequate response to it, I'll offer two brief responses. First, supposing Levy is right about this, there are still plenty of other cases of non-partisan disagreement which do seem to carry information about likelihood. Second, I think that provided partisan disagreements are understood as *genuine* disagreements between epistemic peers – i.e., as cases where at most one of the disputants is correct – we must see them as carrying some information about the likelihood of correctness of the judgments involved. On Levy's account, partisan disagreement provides reason for further investigation of the foundations from which one has reasoned en route to the judgment in question. He argues that such disagreements may raise concerns about irrelevant influences on our judgments, rather than concerns about likelihood of error. But I think these two concerns cannot be independent of one another in the way Levy's account requires. Concerns about irrelevant influences involve concerns about correctness of the associated judgments; hence information about likelihood of the former involves information about the likelihood of the latter.

agnosticism may seem to provide a solution to the multiple partitions problem, by explaining why the principle of indifference is false. I'll argue that one important lesson that emerges from this discussion is that deep agnosticism (with respect to relevant propositions) does not adequately reflect the strength of your evidence in these cases. While this alone does not show that such credal states aren't justified (arguably, this conclusion would be warranted only assuming the truth of evidentialism), I argue that the further assumption that credence should be guided by objective chance does provide grounds for rejecting DA$_d$.

4. Mystery Square and Multiple Partitions

The *principle of indifference* states that in cases where one's total evidence no more supports one proposition than another (using '≈' to express this relation, and letting *P(.)* be a *rational* subject's credence function) rational credence in the propositions should be the same:

> (POI) $p \approx q \rightarrow P(p) = P(q)$.

The *multiple partitions problem* for POI is that, in cases where a proposition *p* is a member of two partitions of different size, the principle gives inconsistent answers to the question of what your credence in *p* should be. Take the following example.

> *Mystery Square*. A mystery square is known only to be no more than *two* feet wide. Apart from this constraint, you have no relevant information concerning its dimensions. What is your credence that it is less than *one* foot wide?

It seems you have no more reason to suppose the square is less than 1 foot wide than that it's more than 1 foot wide. Now consider the *area* of the square. Do you have any more reason to suppose it's less than 1 square foot, than that it's between 1 and 2 square feet, or between 2 and 3, or 3 and 4? It seems you don't.

This is a problem for POI, because it seems to entail that we assign credence values evenly over the two partitions,[30] and this leads to contradiction: Since A1≈A2≈A3≈A4,[31] we should assign equal credence value to each one, i.e. 1/4. And

---

[30] White ("Evidential Symmetry and Mushy Credence," in *Oxford Studies in Epistemology,* ed. T.S. Gendler and J. Hawthorne (Oxford University Press, 2010), 161-186) suggests a modified version of POI, which states that in cases of evidential symmetry we are to apportion our credence evenly, i.e.: (*POI**: If {*p1, p2, . . . , pn*} is a partition of your knowledge such that $p1 \approx p2 \approx \ldots \approx pn$, then for all *i* $P(p_l) = 1/n$). White takes POI* to be a corollary of POI; Decker argues however that POI* only follows from POI given a principle he labels "Full Spreading," which states that mutually exclusive and jointly exhaustive propositions *p1, p2, ..., pn*, one's credence function, *C*, should be such that $C(p_1 \lor p_1 \lor \ldots \lor p_n) = 1$.

[31] A1 is the proposition, 1 sq. ft. < *area of mystery square* ≤ 2 sq. ft., A2 is 2 sq. ft. < *area of mystery*

since L1≈L2,[32] we should assign 1/2 to each of these. But L2 iff A2 ∨ A3 ∨ A4; so L2 ≈ (A2 ∨ A3 ∨ A4). White argues however that the problem is really not with POI, since the absurd conclusion follows from other assumptions without the aid of indifference principles. Specifically, he argues that the problem arises from principles of *transitivity* (If $p ≈ q$ and $q ≈ r$, then $p ≈ r$) and *equivalence* (If p and q are known to be equivalent, then $p ≈ q$).[33] Decker criticizes certain of White's assumptions and argues the principles of indifference really are the source of the problem.[34] I won't try to resolve this dispute here; instead I will focus on the question whether deep agnosticism avoids the problem, as Decker argues.

Decker argues that indifference principles, such as POI, are false, hence they cause no trouble for reasonable peer disagreement. (They would if they were true, because they would demand that you update your credences in such cases in a way that eventually brings you into agreement with their peers.[35]) They are false, in his view, because they're incompatible with deep agnosticism.

For instance, Decker argues that in the case of Mystery Square you have good reason to believe A1 v A2 v A3 v A4 (your credence function is such that C(A1 v A2 v A3 v A4) = 1), even though you have *no* reason at all to believe any of the disjuncts (your credence function is simply undefined for each of these). Consequently, it's permissible for you to be deeply agnostic about A1, A2, etc. C(A2 v A3 v A4) and C(A2) are also undefined, in Decker's view. (You have *no reason at all*, according to Decker, to think that A2 v A3 v A4, or A2, is true.)

But if these credence functions are undefined, then two things follow: (i) it's not the case that C(A2 v A3 v A4) > C(A2) and (ii) C(A2 v A3 v A4) and C(A2) are not definitely *non-zero*. Like White, I take (i) to be intuitively false, given (A2 v A3 v A4) is logically weaker than A2.[36] If C(A2 v A3 v A4) and C(A2) are not definitely *non-zero*, then, if deep agnosticism with respect to these propositions is rational, this means it's rational to withhold judgment on the question of whether A1 (A2, etc.) is possible—i.e., on whether there's some objective chance of its being true. But it isn't rational to withhold judgment on that.[37]

---

square ≤ 3 sq. ft., etc.

[32] L1 is the proposition, O ft. < *width of mystery square* ≤ 1 ft.; L2, the proposition, 1 ft. < *width of mystery square* ≤ 2 ft.

[33] See White, "Evidential Symmetry and Mushy Credence," 165-6.

[34] Decker, "Disagreement, Evidence, and Agnosticism," 757-763.

[35] Decker, "Disagreement, Evidence, and Agnosticism," 756-7.

[36] Decker argues the asymmetry here is merely logical and not epistemic, and he argues (761) that the fact that (A2 v A3 v A4) is a "safer bet" than A2 alone can be explained without positing any epistemic asymmetry. I'm not convinced, but for now I'll set this aside and focus on (ii).

[37] Decker compares the undefined credence function to the case of division by zero, noting that it

Now, as Decker correctly points out, the inference from *we must not assign a credence value of zero* to *we must assign a non-zero credence value* is a *non-sequitur*, since we could very well assign *no credence value whatsoever* to the proposition.[38] The argument I have just given does not rely on this inference, however. To see this, notice that the inference from *we need/must not assign a (precise) credence value* to *we need/must not assign any credence value at all* is *also* a *non-sequitur*. In particular, even though I may reasonably refuse to assign any precise credence value to A1, A2, A3, and A4 (in *this* sense, I may be agnostic), this alone does not entitle me to assign *no* non-zero value at all, even one that is highly indeterminate (or, to use Elga's term, "mushy"[39]). In particular, if I know that there is some objective chance that A1 (A2, etc.) is true, then not only should I refrain from judging that it is definitely not true, I should acknowledge that there is a real

---

follows from its being undefined that it's not the case that 7/0 is equal to 0, 7, or Julius Caesar. ("Disagreement, Evidence, and Agnosticism," 766.) Clearly, it would be a mistake to infer from the fact that I ought not identify any of these as the value of 7/0, that I ought to assign some *other* value. Similarly, he suggests, it would be a mistake to infer simply from the fact that I ought not have a credence value of 0 or 1 for *p* that I ought to have some other credence value for *p*. But there is an important difference between these cases. In the case of division by zero, the fact that we ought not to assign 0 or 7 as the value of 7/0 is explained by, and grounded in, the fact that division by zero is undefined. In the case of Mystery Square, the fact that we ought not to assign 0 or 1 as credence values for A1 (etc.), is, it seems, explained by the fact that one's evidence provides some reason to think A1 is true (because one's evidence clearly supports the disjunction A1 ∨ A2….), and also some reason to think it's false (because it provides some reason to think of each of A2-A4 that it is true). (The nature of the grounding relation is of course controversial, but I take it to be a metaphysical, explanatory relation, such that if a truth p is grounded in certain other truths, then these other truths explain (in a sense which is independent of our interests) why p is true. So, for example, the fact that Either Trump is human or he isn't, is grounded in the fact that he's human. And the fact that water has certain macroscopic properties is grounded in its chemical composition.)

I agree, then, that having a high credence value for the proposition that you shouldn't have a credence value of 0 or 1 for p does not by itself require that you have some credence value for p. (Decker, "Disagreement, Evidence, and Agnosticism," 766) But I claim it does require this in cases where high credence value for this higher-order proposition is explained by your evidence for p. Cases of disagreement with an acknowledged peer are, I claim, of this sort.

[38] "Disagreement, Evidence, and Agnosticism," 760.

[39] On the idea of indeterminate credence value, see e.g. Bas van Fraassen, *Laws and Symmetry* (Oxford: Oxford University Press, 1989), Peter Walley, *Statistical Reasoning with Imprecise Probabilities* (London: Chapman and Hall, 1991), James Joyce, "How Probabilities Reflect Evidence," *Philosophical Perspectives* 19 (2005), 153-78, and Scott Sturgeon, "Reason and the Grain of Belief," *Noûs* 32 (2008), 231-46.

possibility that it *is* true. The trouble with deep agnosticism is that it involves withholding judgment as to whether there is such an objective chance.

It is important to ask whether simply refraining from assigning a credence value of zero, without also assigning a positive value, adequately reflects the strength of your evidence in these cases. As we've just seen, it does not. In the horse race case, for instance, if you recognize your peer as an epistemic peer, then you must at least recognize the possibility that she got it right, whereas you got it wrong. So you must give some non-zero credence to the proposition that she got it right. It's not acceptable, then, to give no credence at all to this proposition.[40]

It seems appropriate at this point to raise the question, under what conditions in general is deep agnosticism about a proposition *p* supposed to be rational? On this matter, Decker has only a couple of brief suggestions. First, he suggests that DA is warranted for propositions that one is unable to represent in thought (e.g., a child may for this reason be deeply agnostic about Godel's incompleteness theorem). This is clearly irrelevant to the cases of peer disagreement we are concerned with, however. There's no question of your being able to grasp the relevant thoughts concerning your peer's judgment and its comparative likelihood of being correct. Decker also suggests that deep agnosticism may be warranted in some cases of propositions one hasn't yet considered. But he acknowledges that there are of course many propositions one hasn't considered toward which one nevertheless has implicit attitudes. So lack of explicit belief doesn't imply lack of credence value in a proposition. Moreover, the sorts of unconsidered propositions Decker has in mind are ones for which one seems to have no relevant evidence one way or the other (e.g., the proposition that the current U.S. President is at this very moment in the Oval Office). Again, the cases of peer disagreement that concern us are not like this. In these cases, you have evidence that bears on relevant propositions regarding your peer's judgment; it's not as though the information available to you is simply silent on the matter.

Moreover, even if we suppose that in general failure to consider a proposition is sufficient grounds for deep agnosticism (a position that bears some similarity to certain contextualist views of knowledge), this would not provide an adequate

---

[40] I think there is a danger here of confusing probabilities and credences. As White notes ("Evidential Symmetry," 162-3), indifference principles should be understood as normative constraints on what your credence may be, not as principles for determining what the objective probabilities or chances are. "You can't get probabilities out of ignorance"—this is granted. Principles of indifference don't say otherwise. So, from the fact that you may have no grounds at all for saying what the objective probability of, e.g., Mystery Square's having 1m sides is, this does not mean that your *credence* in the proposition should be undefined, or even that it should be indeterminate. Agnosticism about these objective probabilities doesn't entail deep agnosticism.

defense of the steadfast position. This is because we are clearly interested not only in cases of peer disagreement where the disputants do not consider relevant propositions about their peers, but also cases where they do explicitly consider such propositions. That is, we want to know how you should respond to disagreement with an acknowledged peer, provided you clearly understand all of the relevant facts concerning their peerhood and their handling of the evidence at hand.

The upshot is that although we may have some examples in which deep agnosticism is plausible, this does not support DA_d. The features of the former which seem to explain the rationality of deep agnosticism are absent in the case of peer disagreement.

Rational Credence and the Principal Principle

I have argued that deep agnosticism with respect to A1-A4 and L1-L2 in the case of Mystery Square does not adequately reflect the strength of your evidence. I suggest that we may find further support for this claim in Lewis's Principal Principle:[41]

(PP) $C(A \mid ch(A) = x) = x.$

PP states that a rational agent conforms her credences to known objective chances. It is based on the idea that we should use frequency data to guide our subjective probabilities. For example, if you know that the coin that you're about to toss is a fair one and so has a 50% chance of landing heads, then you should give equal credence to its landing heads and its landing tails. My argument rests on a closely related idea, namely, that rational credence is constrained by objective chance, in the following way: if you know that there's *some* objective chance that p, then you should assign some (perhaps indeterminate) positive credence value to p. (Note that an even stronger principle seems warranted, namely that if, *for all you know*, there's some objective chance that p, then you should assign some positive credence value to p. However, I won't assume that this principle holds.) While this principle is not strictly entailed by PP, I think it is clearly suggested by the underlying normative principle that we ought to be guided by objective chance in forming credences.

If you acknowledge someone as an epistemic peer with respect to some issue about which the two of you disagree, you thereby acknowledge that there is some objective chance that s/he got it right (and thus that you are mistaken). You also acknowledge that there is some objective chance that s/he handled the available evidence better than you did. So, given the principle suggested, you are rationally required to assign some positive credence value to these propositions.

---

[41] David Lewis, "A Subjectivist's Guide to Objective Chance," in *Philosophical Papers, Vol. II* (Oxford: Oxford University Press, 1986).

It might be objected that recognition of epistemic peerhood (and in particular, of the possibility that your peer's judgment is correct) does not require taking a stand on such objective chances, even supposing that credence should be guided by objective chance in the way I've suggested. After all, it may seem that if you have *no reason* at all to think that your evidence (or your handling of it) is superior (inferior) to your peer's, then you should be deeply agnostic about the *likelihood* of your peer's judgment being correct. To see why this is a mistake, let's consider one more example due to Feldman:[42]

> [There are] two suspects in a criminal case, Lefty and Righty. Suppose now that there are two detectives investigating the case, one who has the evidence [against] Lefty and one who has the evidence incriminating Righty. They each justifiably believe in their man's guilt. And then each finds out that the other detective has evidence incriminating the other suspect. If things are on a par, then suspension of judgment is called for. If one detective has no reason at all to think that the other's evidence is inferior to hers, yet she continues to believe that Lefty is guilty, she would be unjustified. She is giving special status to her own evidence with no reason to do so, and this is an epistemic error, a failure to treat like cases alike.

Feldman is claiming that we have reason to accept the equal weight view, as we have no reason to give greater weight to our own evidence in such cases.[43] Decker argues however that it is a mistake to think that if one of the detectives sticks to her judgment she is indeed giving special weight to it:

> This is a *non sequitur*. If she has *no reason at all* to think that her evidence is inferior to the other's, it would be acceptable for her to be deeply agnostic on the matter of who has better evidence. The situation is better described as her simply continuing to give her evidence status *as* evidence. She needn't be giving it privileged status in the sense of assigning a higher credence value to the proposition that her evidence is superior to the other detective's than to the proposition that the other detective's evidence is at least as good as hers. *She need not be placing any credence at all on those propositions.*...[44] (my emphasis)

So Decker's claim is that A has *no reason* to think her evidence is inferior (or superior) to B's, then it is acceptable—perhaps even required—for A to be deeply agnostic on whether *A's evidence is at least as good as B's*. But *this* is a *non sequitur*!

---

[42] Richard Feldman, "Reasonable Religious Disagreements," in *Philosophers Without Gods,* ed. L. Antony (Oxford: Oxford University Press, 2007), 208.

[43] It should be noted that although Feldman is explicitly concerned only with the quality of evidence in this case, I take him to be supposing that neither detective has a reason to think that he *handled the evidence* any better than the other. (I take this to be implicit in his statement that "things are on a par.")

[44] "Disagreement, Evidence, and Agnosticism," 770.

For A's having no such reason is clearly compatible with A's having some reason to think that B's evidence is *equally* good as A's. This, together with the argument given in §3, implies that A's having no reason to think her evidence is inferior to B's is compatible then with A's having *some* reason to think that B's, rather than A's, *judgment* is correct. Moreover, we should note that, supposing A has no reason to think his evidence is *inferior* to B's, this is compatible with his having some reason to think it is at least possible that B's evidence (and, moreover, B's handling of that evidence) is superior, i.e., that there is some objective chance this is in fact the case. And A does have *some* reason to think all of these things *are* indeed the case, provided she recognizes B as an epistemic peer. So, A should give some positive (and perhaps indeterminate) credence to these propositions.

We may grant Decker's claim that A need not assign higher credence to the proposition that her evidence is superior to B's than to the proposition that it is inferior. We may also grant that she need not assign higher credence to the proposition that her evidence is *equal* to B's than to the proposition that they are unequal (although failure to assign such credence does not seem compatible with a clear recognition of epistemic peerhood). It doesn't follow from any of this that it's rational for her to assign *no credence at all* to any of these propositions. In particular, if A recognizes B *as an epistemic peer*, then she should give high credence to the proposition that B's evidence is at least as good as hers—or, at the very least, she should assign a *non-zero* value to this proposition.[45]

I have argued that it is irrational to be deeply agnostic about relevant peer propositions in Feldman's case. That is, it is rationally required that A assign some positive credence to the proposition that B's judgment is correct, and to the proposition that B's evidence is superior to A's, provided A takes B to be an epistemic peer. Does this imply that it would be irrational for A to stick to his original judgment in these circumstances? I don't claim that it does. And I will not take a stand on the issue of conciliationism vs. steadfastness here. My claim is just that it would be irrational for A to be deeply agnostic on such propositions *while at the same time regarding B as an epistemic peer*. Whether A must *conciliate* provided A

---

[45] Decker is clear that he only insists that it is rationally *permissible* to refrain from assigning a non-zero value in this case, not that this is rationally *required*. ("Disagreement, Evidence, and Agnosticism," fn. 23.) Is this satisfactory? Arguably, it is not. Provided you have good reason to take someone as an epistemic peer, this gives you a conclusive reason to assign a non-zero value to the proposition that their evidence is as good as yours. That is, it *rationally requires* you to accept that this is a genuine possibility. Thus you cannot simply withhold credence in this proposition. Compare Mystery Square: Since you assign a value of 1 to L1 v L2 v L3 v L4, and you know that there is some chance that L1 (L2, etc.), rather than L2 v L3 v L4, you shouldn't simply refrain from assigning any value to L1. Rather, you should assign a positive value to L1.

assigns positive credence to these propositions, or whether there is some basis for steadfastness other than DA$_d$, is a question I will leave open.

5. Objection and Reply

It might be objected that my argument depends on a conflation of, or perhaps an equivocation between, two different understandings of epistemic peerhood, *rationality peerhood* and *accuracy peerhood*. These notions may be defined as follows:[46]

> X and Y are *rationality peers* with respect to a certain issue iff X and Y's opinions on that issue are equally likely (given relevant higher-order evidence) to be rational (i.e. X and Y are equally likely to have reasoned correctly about the disputed matter).

> X and Y are *accuracy peers* with respect to a certain issue iff X and Y's opinion on that issue are equally likely (given relevant higher-order evidence) to be accurate.

The worry is that because my criticisms of deep agnosticism seem to depend on the assumption that epistemic peers are *rationality* peers, they also depend on the (controversial) assumption of uniqueness, i.e.:

> UNIQUENESS: Given one's total evidence, there's a unique rational doxastic attitude that one can take to a proposition.

This is because I suggest (in §3) that any reason to think that your peer has reasoned correctly on the basis of the available evidence is a reason to doubt that *you* have done so. However, if uniqueness is false, then this is not so. For in that case, from the fact that your peer's judgment is *rational*, it does not follow that yours is not also (fully) rational. And if it is false that any reason to think your peer has reasoned correctly is a reason to doubt that you have done so, then disagreement with your peer doesn't rationally require you to conciliate or to reduce your confidence in your judgment.

I have two replies to this objection. First, the notion of epistemic peerhood as I understand it is not to be identified with *either* the notion of *rationality* peerhood or the notion of *accuracy* peerhood. It is distinct from both, and consequently, my argument cannot be said to conflate these, or to equivocate between them. This is because epistemic peerhood as I understand it implies equality (i.e., equal quality) of relevant first-order evidence, and equality of evidence in my view is not implied by either accuracy peerhood or rationality peerhood. For instance, suppose that I know that my friend is in "Fake Barn Country." In that case, I shouldn't regard him as an

---

[46] Cf. Christensen, "Conciliation, Uniqueness, and Rational Toxicity."

accuracy peer, though I may still acknowledge that he's a rationality peer (who is equally likely to have reasoned correctly on the basis of the evidence available to him). Should I take him to be an epistemic peer, then? While I acknowledge that some evidential internalists will probably say 'yes,' I maintain that my peer's evidence is inferior to mine. While we both use the same process of forming perceptual beliefs, we base our respective perceptual judgments on different perceived facts: his judgment, *that's a barn*, is based on his perception of a barn facade, whereas mine is based on my perception of a (real) barn.[47] So, although we count as rationality peers, we do not count as epistemic peers. For X to be an epistemic peer of Y requires not just that X be equally likely to have reasoned correctly, but also that X have equally good evidence as Y. Accuracy peerhood is also insufficient for epistemic peerhood, since X and Y may be equally likely to have determined the truth about some subject matter on the basis of unequal or incommensurable evidence.

Because rationality peerhood is not sufficient for epistemic peerhood, even supposing disagreement with rational peers doesn't require conciliation it doesn't follow that disagreement with *epistemic* peers doesn't require conciliation. (If rational peerhood is compatible with permissiveness, it doesn't follow that epistemic peerhood is, too.)

My second response to the objection is more significant. It is that my case against deep agnosticism does not depend on uniqueness even if epistemic peerhood is understood as *rationality* peerhood. This is because I do not need to maintain that in the cases I've considered, any reason for being confident in your peer's judgment is a reason for doubting your own. All I need is the assumption that in these cases regarding the other person as an epistemic peer is clearly warranted and that this requires assigning positive credence value to the proposition that their judgment is correct. The problem with being deeply agnostic in these cases is that it requires not only that you assign no value to the proposition that your peer has reasoned as well as you, but that you assign no value to the proposition that their judgment is correct.[48] While it is hard to see how you could do that while remaining steadfast, my argument does not depend on either conciliationism or the equal weight view.[49]

---

[47] I don't claim that my peer's judgment isn't *justified* on the basis of the evidence he has. But I do claim that the quality of his evidence is less than that of mine.

[48] The permissivist worry that I raised earlier (fn. 19) may arise again here. At this point, I have nothing further to say in response and will simply emphasize how modest the claim in question really is: what is ruled out is that you are rationally permitted to assign no credence value at all, hence what's required is only that you assign *some* value, not that you assign a value that is sufficient for belief, or for belief that it is more likely than not to be true.

[49] It's also worth noting that the motivations for conciliationism and equal weight do not clearly

Granted, I do claim that in Elga's horse race case, for example, confidence in your peer's judgment implies (some degree of) doubt about your own. But here I am concerned only with the correctness of the conflicting judgments, only one of which *can* be correct. The point is that although your recognition of the other as an epistemic peer and of the fact that she's reached a different judgment of who won the race need not diminish your confidence that you've handled the evidence just as well as she has, it *should* diminish your confidence that your judgment of who won is *correct*. In this respect, at least, peer disagreement requires conciliation. Whether or not disagreement in the horse race case casts doubt on the *rationality* of your judgment, it casts doubt on its *accuracy*.[50]

It might be thought that my notion of epistemic peerhood may be understood simply as *accuracy + rationality* peerhood (call this *AR peerhood*). This may seem to be suggested by my talk of epistemic peers being equally likely to "get it right," as well as equally likely to have handled the evidence correctly. But, while I am open to the possibility that AR peerhood is at least extensionally equivalent to my notion of epistemic peerhood—i.e., X is an epistemic peer of Y iff X is an AR peer of Y—I do not think that I am committed to this, and I do not think that AR peerhood and epistemic peerhood are *a priori* equivalent. For epistemic peerhood is, as I've explained, to be understood in terms of *evidence*, and the relationships between evidence and truth, and between evidence and rationality, are a highly contested matter. So I will refrain from claiming that epistemic peerhood is or is not extensionally equivalent to AR peerhood.

To sum up, deep agnosticism is consistent with the plausible view that it's rationally permissible to stick to your judgment in cases of peer disagreement (thus, to be more confident in your judgment than your peer's). Moreover, it implies (again, not without plausibility) that in doing so you need not take a stand—that is, you may remain deeply agnostic—on the comparative likelihood of your judgment's being true versus your peer's. But deep agnosticism about such propositions is

---

depend on uniqueness. Authors who argue that they do not include Christensen, "Conciliation, Uniqueness, and Rational Toxicity," Elga, "Reflection and Disagreement," Kornblith, *Disagreement* (Oxford: Oxford University Press, 2010), Frances, "The Reflective Epistemic Renegade," *Philosophy and Phenomenological Research* 81, 2 (2010): 419-63, and Lee, "Conciliationism Without Uniqueness," *Grazer Philosophische Studien* 88, 1 (2013): 161-88. Thus even if my argument against deep agnosticism were to rely on the assumption that either equal weight or conciliationism is correct, this would arguably not commit me to uniqueness.

[50] Cf. Christensen, "Conciliation, Uniqueness, and Rational Toxicity." Christenson argues that the pressure to conciliate generated by accuracy-based evaluation of peers is less threatened by permissive accounts of rationality than the pressure to conciliate generated by rationality-based evaluation.

incompatible with a recognition of your peer *as an epistemic peer*, in such circumstances. This recognition requires you to take a stand—i.e., to assign some credence value—on the question of comparative likelihood of correctness. Moreover, although deep agnosticism may avoid the multiple partitions problem in cases like Mystery Square, the way in which it does this is unsatisfactory, since it is not consistent with an adequate view of the strength of your evidence in such cases.

LINGUISTIC EVIDENCE AND SUBSTANTIVE EPISTEMIC CONTEXTUALISM

Ron WILBURN

ABSTRACT: Epistemic contextualism (EC) is the thesis that the standards that must be met by a knowledge claimant vary with (especially conversational) contexts of utterance. Thus construed, EC may concern only knowledge claims ("Semantic EC"), or else the knowledge relation itself ("Substantive EC"). Herein, my concern is with "Substantive EC." Let's call the claim that the sorts of linguistic evidence commonly cited in support of Semantic EC also imply or support Substantive EC the "Implication Thesis" (IP). IP is a view about which some epistemologists have equivocated. Keith DeRose is a case in point. Herein I argue that IP is false, and that it is false for interesting reasons. To this end, I consider two other terms which DeRose investigates, "free will" and "potency" in his efforts to demonstrate the alleged inability of distinctly philosophical or skeptical doubts to infect ordinary epistemic discourse. I describe how and why these two examples speak against, rather than for, DeRose's recommendation of Substantive EC.

KEYWORDS: skepticism, contextualism, knowledge, DeRose, Rieber

1. Introduction

Ignoring minor endogenous disagreements, we can take epistemic contextualism (EC) to be the thesis that the standards that must be met by a knowledge claimant vary with (especially conversational) contexts of utterance. Thus, even though knowledge claims must satisfy relatively low epistemic standards in some contexts, they must satisfy higher standards in other contexts, where more remote sources of possible disinformation and error (ultimately generating skeptical scenarios) arise for consideration. Using precedent diction, we can say that contexts are formal structures that provide values for what counts as proof, thus determining the truth values of epistemic claims. They are distinct from situations, i.e., concrete arrangements of items within which sentential utterances occur. Consequently, situations include utterances and determine contexts which, in turn, generate variable sentential truth values. A single sentence can have different truth values at different times as a function of different contexts, which is to say, different situations in which it is uttered, causing it to vary in meaning.

Thus construed, contextualism is a semantic thesis that may or may not have epistemological consequences: it can concern only knowledge *claims,* or it can

concern the knowledge relation itself. Let's call the view that what "knowledge" *means* depends on contextual factors "Semantic EC." Let's call the claim that what knowledge *is* depends on contextual factors "Substantive EC."[1] Let's call the claim that Semantic EC presupposes, *and thus implies,* Substantive EC the "Presupposition Thesis," and the denial of this position "(Contextualist) Separatism." More specifically, Semantic EC is the view that "knowledge" discourse has an indexical status that causes the meaning and truth conditions of sentences containing "know" to vary with contextually determined standards of appropriate rigor (concerning stakes, interests, etc.) Substantive EC is the view that the knowledge relation itself varies with differences in contextually determined standards of appropriate rigor (concerning stakes, interests, etc.) Finally, the Presupposition Thesis is the assertion that Semantic EC is only plausible on the assumption of Substantive EC. It is the view that, if true, Semantic EC provides *grounds* for Substantive EC because the contextual character of "knows" implies the contextual character of the knowledge relation as a result of presupposing it. Finally, the Presupposition Thesis is itself to be distinguished from the "Implication Thesis," which is the claim that the sorts of linguistic evidence that contextualists cite in favor of Semantic EC imply Substantive EC. Note the difference between the two doctrines. The Presupposition Thesis claims that Substantive EC is a necessary condition for Semantic EC. The Implication Thesis maintains that the facts about use commonly invoked in defense of Semantic EC ensure the truth of Substantive EC.

For illustration, consider the bearing of each of these views on worldly skepticism. Semantic EC maintains that Moore's assertions and those of the skeptic don't conflict. Substantive EC holds that the skeptic can gain no critical traction against ordinary knowledge claims because there is no knowledge relation with a singular determinate nature at issue. The Presupposition Thesis implies that there is no acontextual, univocal meaning of "knowledge" that the skeptic can critically invoke *because* knowledge has no singular, determinate nature. The Implication Thesis is the claim that the observed variability of standards in our everyday use of "know" implies that knowledge lacks a singular, determinate nature.

I (apologetically) present this tiresome topography of positions only to make my limited aspirations in this paper clear. I argue elsewhere *for* the Presupposition Thesis, that is, against Contextualist Separatism. I argue here only *against* the Implication Thesis. I ask if features of linguistic employment that lead contextualists to accept Semantic EC imply Substantive EC. That is, suppose we concede that

---

[1] Patrick Rysiew, "Epistemic Contextualism," in *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), eds. Edward N. Zalta (Stanford: Stanford, CA, 2016): URL = https://plato.stanford.edu/archives/win2016/entries/ contextualism-epistemology/.

various facts about the use of "know" that the Semantic Contextualist cites (regarding variable standards for "knowledge" discourse in different contexts of inquiry) obtain. Does this linguistic evidence support the truth of Substantive EC?

Few philosophers explicitly embrace the Implication Thesis (e.g., David Innis, David Lewis, perhaps Michael Williams).[2] Most consistently deny it or at least remain agnostic as to its truth. Keith DeRose, I submit, is an interesting case in that he has tried to have it both ways. In some of his work, he is adamant that his is merely a Semantic EC, and that the evidence from use he presents is offered in support of nothing more.[3] But at other points, he seems to describe his contextualism as a Substantive one that receives support from observable facts about the use of "knows" in varying conversational contexts. He explicitly repudiates the claim that contextualist reasoning and conclusions pertain merely to assertions *about* knowledge and not to knowledge itself. He objects to the contention that contextualism has no bearing on epistemology and skeptical concerns about the scope and limits of knowledge.

DeRose derides this critical contention and maintains that it is easily refuted. To illustrate how, he considers the case of free-will attribution, maintaining that its potential amenability to contextualist analysis provides a sanguine parallel for the example at hand. Let's examine the Implication Thesis largely through the lens provided by this suggested parallel. Once we make this analogy out in more detail, I suggest, we can see, contra DeRose's assertions, that it thwarts rather than supports the Implication Thesis. The reason: contextualist accounts of both "knowledge" and "freedom" commit us to analyses of their respective concepts upon which absolutist criteria governing these concepts' application to limiting cases (identifying "knowledge" and "freedom" in their "strictest" senses) are no more deeply motivated than any other potential set of criteria. But these analyses, I submit, are mistaken. Even though such absolutist criteria issue from highly distinctive and idiosyncratic reflections, they still exercise broad critical authority over questions we should ask about knowledge and freedom in other, more prosaic, contexts. In other words, even if the use of "know" does vary with context, the skeptic's sense of the term is privileged in a way that should give us special pause when doing epistemology.

In section 1, I sketch this analogy between Semantic EC and Semantic Free Will Contextualism. In section 2, I contrast these two positions with simple

---

[2] David Annis, "A Contextualist Theory of Epistemic Justification," *American Philosophical Quarterly* 15 (1978): 213-219, David Lewis, "Elusive Knowledge," *Australasian Journal of Philosophy* 74 (1996): 549–567, Michael Williams, *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism* (Princeton: Princeton University Press, 1996).

[3] Keith DeRose, "Assertion, Knowledge, and Context," *Philosophical Review* 111 (2002): 167-203.

invariantist alternatives. In section 3, I turn from expository to critical aims and address four questions: what is the alleged semantic evidence for substantive "knowledge" and "free will" contextualism; how does this evidence allegedly provide these positions support; and what must we assume (regarding language, the world, and the relation between the two) to view semantic phenomena as having substantive consequences for our understanding of the phenomenon of free will and knowledge, respectively. In section 4, I ask if the assumptions identified in section 3 are supported by DeRose's linguistic evidence. I contend that they are not. That is, I argue that DeRose's linguistic evidence doesn't support an account of knowledge that obliges invariantist epistemologists to rethink their views about knowledge and skeptical threats. Finally, I offer morals and conclusions concerning the status of distinctly philosophical inquiries about knowledge and agency. Note that my consideration of agency talk is purely heuristic. I have no immediate interest in this paper with issues of free will. I discuss them merely to shed light on epistemic concerns. My point is that DeRose's parallel between epistemological concerns and concerns regarding free will speak *against*, rather than *for*, the conclusions that DeRose recommends.

2. DeRose's Analogy and Rieber's Account

Let's begin with exposition. What are the details of DeRose's analogy between the attribution of knowledge and the attribution of free will, and what form must a contextualist account of free-will attribution take? Consider the following passage from one of DeRose's most extended expositions of the alleged parallel between knowledge and free-will attribution.

> Though contextualism/invariantism is an issue in the philosophy of language, it's a piece of philosophy of language that certainly has the potential to be of profound importance to epistemology. How we should proceed in studying knowledge will be greatly affected by how we come down on the contextualism/invariantism issue. For contextualism opens up possibilities for dealing with issues and puzzles in epistemology which, of course, must be rejected if invariantism is instead correct. And how could it be otherwise? Those who work on the problem of free will and determinism, for instance, should of course be very interested in the issue of what it means to call an action 'free.' If that could mean different things in different contexts, then all sorts of problems could arise from a failure to recognize this shift in meaning. If there is no such shift, then that too will be vital information. In either case, one will want to know what such claims mean. Likewise, it's important in studying knowledge to discern what it means to say someone knows something. If that can mean different things in different contexts, all sorts of problems and mistakes in epistemology, and not just in philosophy of language, will arise from a failure to recognize such shifts in meaning. If, on the other hand, there is no such

shift, then we're bound to fall into all sorts of error about knowledge, as well as about 'know(s),' if we think such shifts occur. It's essential to a credible epistemology, as well as to a responsible account of the semantics of the relevant epistemologically important sentences, that what's proposed about knowledge and one's claims about the semantics of 'know(s)' work plausibly together across the rather inconsequential boundary between these two subfields of philosophy.[4]

Here DeRose imagines how the dialectical precedent established by the free will debate might support his own strategic goal of deriving substantive epistemological morals from facts about the use of the term "knowledge."

The above passage calls for a bit of rational reconstruction. What would have to be the case for DeRose's observations about *use* to have consequences for the *phenomena* of knowledge and freedom, respectively, particularly insofar as skepticism about knowledge and freedom are concerned? Just as a speaker can truthfully say 'S freely performed A,' in low-standard contexts and 'S doesn't freely perform A' in high standards contexts, even though S and A remain constant, a speaker can truthfully say 'S knows that p' in low standards contexts and 'S doesn't know that p' in high standards contexts, where, again, S and p remain constant. In the case of each topic, the use of a critical term ("freedom," "knowledge") directly determines that term's range of application. Thus, by showing that philosophical uses of "free will" and "knowledge" differ from more ordinary uses of these terms, we automatically show that philosophical claims and queries concerning the phenomena of freedom and knowledge do not overrule our more ordinary claims and queries about these topics.

In the epistemic case, the skeptic aims to derive a negative philosophical conclusion about knowledge and then generalize this conclusion across all other contexts of epistemic claim-making. The epistemic contextualist then seeks to undermine this strategy by invoking the alleged indexicality of epistemic standards to show that the skeptic's use of "know" differs from more prosaic senses. Finally, the opponent of the Implication Thesis contends that this contextualist rejoinder is illegitimate, as it concerns the use of the word "knowledge" rather than knowledge itself. DeRose then responds that facts about the use of "knowledge" do indeed yield substantive truths about knowledge. Given the nature of the disagreement, it is clear that DeRose must concede this, as it is otherwise difficult to ascribe consequence to his claim that if "knowledge means different things in different contexts, then all sorts of problems could arise" for epistemology. What problems could arise for epistemology if the issue is merely semantic? It is hardly a difficulty for our

---

[4] Keith DeRose, *The Case for Contextualism: Knowledge, Skepticism, and Context*, Vol. 1 (Oxford: OUP, 2009): 18.

understanding of the skeptic that standards of knowing might shift. This is a mere equivocation which, once pointed out, dissolves in a flash as the skeptic, once she becomes cognizant of it, restates her original challenge with perfect clarity. Problems do not remain for epistemology because of equivocation once this equivocation is resolved. And such resolution requires nothing more than a moment's conversation and reflection.

As for his parallel with issues of free will, DeRose must reason in similar terms. He must believe that the agency skeptic who imposes libertarian demands on freedom but denies that they can be satisfied, much like the epistemological skeptic, aims to derive a negative philosophical conclusion and then generalize this conclusion across all contexts of personal agency ascription. Second, he must have us imagine a contextualist rejoinder to this reasoning. Suppose we found out that said agency skeptic's use of freedom" is distinctly idiosyncratic to philosophical inquiry. Surely, DeRose maintains, we would not suspect the free-will contextualist of illegitimately drawing substantive conclusions from "merely" semantic premises. By parity of reasoning, we should not suspect the epistemic contextualist of doing this either. Thus, DeRose suggests, the precedent presented by the free will debate helps us see why facts about the use of "knowledge" have a bearing on knowledge itself.

Steven Rieber elaborates on these matters by offering us a more detailed contextualist account of free will ascription. Significantly, this account is closely modeled on DeRose's contextualist account of knowledge attribution.[5] Rieber asks us to consider a statement triple (1-3) that closely mirrors the statement triple (4-6) that DeRose uses to articulate both his skeptical puzzle and his response to it.

> (1) Emma raised her hand freely.
>
> (2) If Emma's raising her hand is the product of a causal chain going back to something other than Emma, then her raising her hand was not free.
>
> (3) Emma's raising her hand is the product of a causal chain going back to something other than Emma
>
> (4) Tom knows that this animal is a zebra.
>
> (5) If Tom does not know that this animal is not a cleverly disguised mule, then he does not know that it is a zebra.
>
> (6) Tom does not know that this animal is not a cleverly disguised mule.

(1) and (4) articulate ordinary claims concerning the phenomenon (e.g., free will, knowledge) at issue, each employing a perfectly ordinary instance of the relevant

---

[5] Steven Rieber, "Free Will and Contextualism," *Philosophical Studies* 129 (2006): 223–252.

concept. (2) and (5) articulate conditionals asserting "very reasonable necessary condition[s] for the application[s] of the[se] concept[s]." (3) and (6) articulate negative conclusions regarding the possibility that these concepts might ever apply.[6]

In large part, it is because these two puzzles are so similar in form that Rieber derives the strategic moral that he does: contextualist analysis is no less appropriate to the analysis of "free will" attribution than it is to "knowledge" attribution. Both notions are indexical ones, characterized by meanings that vary with speakers' circumstances of use. Paralleling DeRose, Rieber takes his analysis of "freedom" to bear upon a proper understanding of freedom itself. Just as epistemic contextualists maintain that applicable constraints on knowledge automatically increase when we consider the likes of (4) and (5), freedom contextualists should maintain that applicable constraints on free agency automatically increase whenever we consider the likes of (2) and (3). When we consider (4) and (5), the number of counter possibilities we are obliged to rule out before issuing knowledge claims increases automatically.

Similarly, when we consider (2) and (3), the remoteness of the prompting causes that we are obliged to consider before judging an action free increases to a similar degree. In consequence, just as (4) is true in ordinary contexts in which we ignore skeptical scenarios and false in contexts in which we contemplate them, (1) is true in ordinary contexts in which we ignore distant prompting causes of action and false in contexts in which we consider them. To parallel epistemic contextualist strategy, Rieber phrases his solution to the free will puzzle in terms of shifting standards governing the ascription of agency. In short, we need reject none of hard determinism, compatibilism, or libertarianism. Each of these accounts is legitimized by different criteria of use appropriate to distinct situational and conversational contexts.

To Rieber's credit, his account does not stop with the above story. He does not merely describe the form that a contextualist treatment of freedom should take. He also endeavors to justify this description with a background account that receives support from something other than its mere ability to solve the puzzle posed by (1) - (3). That is, Rieber aims to show that the consequence described above is but one application of a more general feature of our causation discourse, a hallmark that manifests itself in ordinary conversation about agency. This feature regards the meaning of the phrase "did an action freely." "To say that an agent did F freely is to say that the agent caused F and in so doing was the original cause of F."[7] Here the expression "in so doing" ensures identity between the agent's performing the action

---

and the agent's being the original cause of the action, and the expression "original cause" fuels contextualist meaning shifts by being itself deeply contextual in character.[8]

Puzzle cases and questions of agency aside, Rieber insists, the contextuality of "original cause" also elucidates our common practices of causal ascription. It explains, for instance, why "if the burning of the house is the product of a causal chain going back to something that is not lightning, then the original cause of the house's burning was not lightning" is true in any context in which we utter or entertain it.[9] Rieber makes much of this fact because he takes it to show why his puzzle solution is not *ad hoc*. It is a special application of a more general rule applying to a wide swath of ordinary language.

3. Classic Absolutist Alternatives

We've sketched Rieber's contextualist analysis of the concept "free will" as an aid to understanding freedom itself. But is this account convincing? Rieber defends its cogency through an argument that is virtually identical in form to DeRose's argument for Semantic EC and the Implication Thesis. Both arguments are essentially comparative, contrasting Semantic contextualist accounts with their invariantist alternatives (according to which sentences containing relevant terms retain single sets of truth-conditions across all contexts of inquiry), and then urging the greater relative plausibility of the former. Let's briefly review these invariantist notions of "knowledge" and "free will," respectively, as well as the comparative arguments for contextualism that they are used (as critical targets) to fuel.

In the case of knowledge, we can do no better than focus on Peter Unger's original, classic, and quintessential classical skeptical invariantist account in *Ignorance*.[10] On Unger's account, "know" is a verb that conversational conventions allow us to employ even though its hyper-stringent conditions of application seldom,

---

[8] This account selectively obviously invokes an originationist, rather than a so-called "consequence" account of what free will involves (à la Peter Van Inwagen, *An Essay on Free Will.* (Oxford: Clarendon Press, 1983). I think that this focus is justified, however, for two reasons. For one thing, the originationist account seems more clearly to be what Rieber has in mind, given the continuum analysis. For another thing, I take these two sorts of accounts to be very closely related: the ability to cause personal actions is effectively identical to possessing the ability to have done otherwise than one has (à la Robert Kane, *A Contemporary Introduction to Free Will* (Oxford: Oxford University Press, 2005). It's being the case that one could have one otherwise than one did is a best understood as a *symptom* of the fact that one is the source or origin of the action in question.

[9] Rieber, "Free Will and Contextualism," 235.

[10] Peter Unger, *Ignorance: A Case for Scepticism* (Oxford: Oxford University Press, 1975).

if ever, apply. Consequently, on Unger's account, the skeptic's professions of ignorance as to whether she is, e.g., possessed of hands, though true, carry numerous false implicatures concerning what we are ordinarily allowed to assert and infer. Moreover, this fact about "know" is not an isolated one. Unger is as concerned as Rieber to offer a solution that isn't *ad hoc*, and to offer one that appeals to even more general features of language than does Rieber's. On Unger's account, "know" is a member of a broad class of "absolute" terms (e.g., "flat," "straight," "empty") whose conditions of conversational appropriateness vary similarly from their truth conditions.

Let's call the principle governing this variance the "Absolute Term Rule" (ATR). According to ATR, one may use absolute terms in circumstances that approximate literal satisfaction conditions well enough in ordinary contexts to serve our purposes, even when this use is, strictly speaking, incorrect. It is incorrect because absolute terms refer to the logical upper limits of their target properties. To say that an object is flat is to say that it "could not possibly get any flatter."[11] Similarly, to say that one knows a proposition is to say that one could not be more certain of it.[12] DeRose argues against this account because it displays a fatal attraction to error theory according to which most of our claims about knowledge, though meaningful, are false. It is an account according to which we seldom, if ever, know anything, and few, if any, of our ordinary knowledge claims are correct.

Using Unger's skeptical invariantism as a model, let's ask what a parallel free will invariantism must be like. Its defining feature must be a construal of freedom as a limiting target or regulative ideal. An action is free, on such an account, only if it "could not possibly get any freer." Thus, even though attributions of "freedom" might colloquially vary in degree along a single scale of measurement, they can only strictly and correctly apply at this scale's limiting upper boundary. For Rieber, the central requirement of invariantist free will is apparent in the second premise of his puzzle, which stipulates that a person's actions are free only if they eventually "go back" to oneself. Even more fundamentally, this requirement is apparent in Rieber's decision to treat "original cause" as more primitive than "free will" itself, as when he writes, "the claim is that statements about free will are, upon analysis, statements about the original cause of action."[13] Robert Kane effectively describes the libertarian invariantist notion of freedom at play here with his account of "ultimate origination."[14] On Kane's telling, different senses of "freedom" are best compared

---

[11] Unger. *Ignorance*, 64.
[12] Unger. *Ignorance,* 61.
[13] Rieber, "Free Will and Contextualism," 235.
[14] Robert Kane, *A Contemporary Introduction to Free Will* (Oxford: Oxford University Press,

through their relative placement along a single serial scale of increasing autonomy. At the base of this scale lies the freedom of "self-realization" (exploited by simple compatibilists, for instance, Hume on a standard interpretation) to do what we want in the absence of external constraint. Above this lies the freedom of "self-control" exploited by hierarchical compatibilists (e.g., Frankfurt and Fischer) to be moved by desires that are themselves regulated by higher-order wants and values.[15][16] At the upper limit of this scale lies the fully incompatibilist libertarian freedom of "self-determination," according to which we somehow autonomously source our higher-order desires and values, as it were, from nothing (or ourselves). Self-determination, if it existed, would consist in ultimate originating control," exercised by agents when it is "up to them which of a set of possible choices or actions will now occur, and up to no one and nothing else over which the agents themselves do not also have control."[17]

Paralleling DeRose, Rieber argues against this invariantist account of freedom because it displays a fatal attraction to error theory. That is, on this absolutist account, to say that A is the original cause of B is to say that A is absolutely the first cause in the causal chain leading to B. Still, given our actual use of the phrase "original cause," most, if not all, of our claims, not just about free will but about original causation itself, are rendered false on such an absolutist telling.

4. Basic Questions

Our aims so far have been expository. Through DeRose, we have articulated the idea of using the free will debate to clarify the relevance of Semantic EC to Substantive EC. Through Rieber, we have put ourselves in a better position to explore this strategy. Since Rieber's "free will" contextualism is closely modeled on DeRose's "knowledge" contextualism, it affords us a useful instrument with which to identify basic presuppositions present in both accounts. To provide this explanation, let's pursue more critical aims in this section by addressing the following questions about both epistemic and free will contextualism. First, in the case of both "knowledge" and "free will," what is the alleged evidence for semantic contextualism (about "knowledge" and "free will")? Second, how might this evidence provide support for

---

2005).

[15] Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (1971): 5–20.

[16] John Martin Fischer, *The Metaphysics of Free Will: An Essay on Control* (Oxford: Blackwell Publishers, 2005).

[17] Robert Kane, "Free Will: New Directions for an Ancient Problem," in *Free Will*, ed. Robert Kane (Oxford: Blackwell (2003): 232.

substantive contextualism (about knowledge and free will)? Third, what must we assume (regarding language, the world, and the relation between the two) to view the semantic evidence used to adduce semantic contextualism as having substantive consequences for our understanding of the phenomena of free will and knowledge? Fourth, does it provide such support?

The answer to the first question is clear. It is apparent in the common form of the puzzles described above. The evidence for Semantic EC consists of observed facts about our use of the relevant terms. In the puzzle cases, these facts regard conflicts that arise between attributions or denials (of knowledge and free will) that we assert juxtapositionally in the same breaths. Because all the claims (1-6) seem true to us, we are loath to deny any of them. Semantic EC purports to save the day by making such denials unnecessary. By allowing all of our seemingly true claims to remain true, albeit within their separate and distinctive contexts, it saves us from having to judge any of them false. The alleged advantage of Semantic EC, then, is that it allows us to retain our original convictions concerning the truth and falsity of relevant claims. We need not choose between some of these claims and others in ways that court fatal error theoretic consequences by making intuitively true claims false. To attribute an absolute invariant character to "know" is to deviate so wildly from our ordinary attributive knowledge talk as to render it unrecognizable.

The answer to the second question is still not obvious. We have seen DeRose assert that if either "know" or "free" means different things in different contexts, then "all sorts of problems could arise from a failure to recognize this shift in meaning." But this is merely to assert that a connection obtains between the investigations of "knowledge" and knowledge, respectively. It is not to explain what this connection is. As noted above, the "problems" in question cannot only be that "know," and "free" are equivocal, irrespective of whether it is a contextualist mechanism or some other that spins the wheel of meaning variation. Again, what problems could arise for epistemology if the issue is of this nature? It is hardly a difficulty for our understanding of the skeptic that standards of knowing might shift. This is a mere equivocation which, once pointed out, dissolves in a flash as the skeptic, once she becomes cognizant of it, restates her original challenge with perfect clarity. Serious problems do not remain for epistemology due to trivial equivocations whose resolutions require nothing more than a moment's conversation and reflection. We will return to this issue shortly.

The answer to the third question obviously depends upon assumptions we make regarding the evidential value of our normal patterns of assent and dissent to attributions of knowledge and freedom, respectively. Must we assume that our initial pre-theoretical patterns of knowledge and agency attribution should be taken at face

value? Why should we assume that an adequate account of such attributions must accord with our initial usage? Isn't it enough to explain why we *attribute* truth-values in the ways we do? Must we also show that these attributions are correct? On the face of it, there seems to be no reason to assume that our initial convictions about knowledge and agency attribution are truths to be grounded rather than predictable illusions to be explained away.

What makes this fact pressing, of course, is that the Semantic invariantist accounts of knowledge and agency attribution described above explain, no less effectively than their contextualist alternatives, why we *initially* distribute truth-values to the relevant attributions in the ways that we do. And they do so in a manner that is not *ad hoc*. Invariantist positions explain the behavior of "knowledge" in terms of a general ATR that applies across broad swaths of language (e.g., "flat," "vacuum," and the like). Similarly, invariantist positions explain the behavior of "freedom" in terms of a general notion of agency that is itself grounded in a broader concept of "ultimate origination." Importantly, they provide this explanation without assuming the ultimate correctness of the attributions at issue: indeed, they are engineered to avoid precisely these attributions. Skeptical invariantism typically tells us that we ordinarily ascribe knowledge and agency as we do because our purposes in engaging in such ordinary ascription grant us practical license to speak loosely.[18]

On this telling, we can employ "knowledge" and "freedom" in circumstances that approximate the satisfaction conditions of these terms well enough for conversational and situational purposes, even though, strictly speaking, these terms hardly ever or never literally apply (such being the nature of absolute terms). Bear in mind that I have no concern *here* to argue that Semantic EC is false (though I believe it is unmotivated if Substantive EC is false, and I believe that Substantive EC is false). My concern is with the question of what, if anything, the contextualist's account of knowledge attribution tells us about the phenomenon of knowledge itself. One might argue that a construal of knowledge as a singular, determinate relation is distasteful because it fails to accord with the surface grammar of

---

[18] Note that DeRose would presumably take exception to the claim that ATR "applies to broad swaths of language." He explicitly contrasts Unger's account of absolute terms with the likes of "Assert the Stronger" on the grounds that the latter, but not the former, functions as a *general* conversational rule. I don't understand the motivation behind this criticism, however, given that we have no reason to suppose that a linguistic rule must be *completely* general to be significantly general at all. "I before e," "avoid contractions in formal academic writing," and so on, come immediately to mind, as does Grice's "assert the stronger" rule itself. To describe Unger-style warranted assertability analyses of "knowledge " as "bare" or "lame" is to practice name-calling, not philosophy.

knowledge attribution. However, I suggest that this fact gives us no more reason to accept Substantive EC than it gives us reason to take surface grammar less seriously.

In consequence, substantive contextualist accounts of knowledge and agency are not recommended by the linguistic evidence *per se*. Rather, they are supported by a specific construal of the prosaic linguistic evidence as a repository of criteria for identifying instances of genuine knowledge. What, if anything, favors such a construal? My focus here is on the work of DeRose. Much of DeRose's writing in this connection is negative; he argues *against* alternative positions according to which truth conditions and warranted assertability conditions (which I here take to be conditions that render knowledge assertions "appropriate" even when said conditions need not track truth) come apart. Since these arguments take us too far afield and constitute a minor paper on their own, however, I must leave them for another occasion. So, let's concern ourselves here only with possible positive arguments, that is, direct arguments for the position that prosaic linguistic evidence is a repository of evidence for identifying instances of genuine knowledge.

5. Is Prosaic Use Relevant to Knowledge and Freedom?

What must such arguments presuppose? To answer this question, we must ask the following: what must we assume (regarding language, the world, and the relation between the two) to view the linguistic evidence for contextualism as having substantive consequences for our understanding of free will and knowledge? I submit that the assumptions we would have to make to do this are implausible. Thus, I deny the Implication Thesis.

Peter Ludlow writes, "if someone claimed that to know that "Snow is white" is to bake a cake and write "Snow is white" in icing on the cake, the first and most obvious objection is that they do not know what "knows" means."[19] This is true enough, but hardly helpful or telling. It is not enough to note that no "investigation into the nature of knowledge that does not conform to some significant degree with the semantics of the term "know" would simply miss the point." Even if we take meaning to be a function of use, the fact that "know" is used in different ways in different contexts tells us nothing definitive about how the various senses of "know" stand in relation *to each other*. This is a question that we must answer before the issue of the context variability of knowledge standards to Substantive EC can be addressed.

---

[19] Peter Ludlow, "Contextualism and the New Linguistic Turn in Epistemology," in *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, eds. Gerhard Preyer and Georg Peter (Oxford: Oxford University Press, 2005): 13.

Ron Wilburn

One implicit critical commonality between DeRose's and Rieber's accounts is what we may call the "Continuum Account." The central idea behind the continuum account is that stringent, absolutist notions of "knowledge" and "freedom" are nothing special. They are clearly extreme: in the case of knowledge, the absolutist's criteria, à la Unger, is that of certainty; in the case of freedom, the absolutist's criteria, à la Kane, is that of ultimate origination. However, this extremism, according to DeRose and Rieber, does nothing to render them privileged. They are mere points on continua of possible sets of criteria, applicable only within their own restricted contexts of usage. Both the skeptic and the libertarian go wrong because they force unusual features of context onto conversation as though they were general features of knowledge and freedom themselves. On this reading, skeptics and libertarians leave ordinary knowledge claims untouched because their use of "knowledge" and "freedom" fail to accord with those of ordinary claim-making practices and because there is nothing privileged about their use. Indeed, the very extremity of these uses highlights the fact that there are no univocal senses of "knowledge" and "freedom," but only different senses of "know" and "free" appropriate to different contexts of use. In the case of knowledge, this leads to DeRose's own positive argument for the relevance of his linguistic evidence concerning "knows" to his epistemological conclusion regarding knowledge. Based on the Continuum Account, he suggests, we can conclude that there is nothing privileged or authoritative about the skeptic's standards of knowledge. Thus, the skeptic's conclusions about knowledge carry no more weight than findings derived from any other investigative stance. Therefore, he writes.

> for my part, once the skeptical strategy is seen to have no tendency to show that any of my claims to know—except those very rare ones made in settings governed by 'absolute' standards—are in any way wrong, and once I start to get a clear look at what it would take to 'know' according to the skeptic's absolute standards, I find the distress caused by my failure to meet those standards to be minimal at best—perhaps to be compared with the 'distress' produced by the realization that I'm not omnipotent.[20]

I submit that this is mistaken. The reason is that questions about knowledge and questions about freedom, as opposed to questions about personal potency, understood strictly, both arise from highly distinctive considerations, but considerations which, by virtue of this distinctiveness, manage to exercise broad critical authority over questions we ask about these relations in other, more prosaic, contexts. We can eschew invariantist semantic analyses of knowledge attribution in

---

[20] Keith DeRose, "Sosa, Safety, Skeptical Hypothese," in *Ernest Sosa and His Critics.*, ed. J. Greco (Oxford: Blackwell, 2004): 38.

specific contexts if we like. However we should view this eschewal as having no effect against the regulative authority of stringent, invariantist senses of "know." Similarly, we can eschew libertarian standards of agency in specific contexts. Still, we should view this eschewal as having no effect against the regulative authority of stringent, invariantist senses of "free."[21] To realize that we fail to know in the skeptic's sense of freedom in the libertarian's sense is not like realizing that we lack omnipotence. To lack omnipotence is to possess a set, quantifiable, but limited amount of power, which is a clearly scalable commodity. However, to lack epistemic certainty is to lack the grounds upon which we may reason that *any* of our worldly beliefs track the truth at all. Similarly, to lack libertarian agency is to lack the grounds upon which we may reason that any of "our" actions flow from us in any significant sense. Even if we accept context principles that tell us that the legitimacy of our judgments concerning what we know or how free we are depends somehow on context, the task of identifying our contexts remains.

These claims call out for elaboration. First, consider knowledge. As I have repeatedly noted, I constrain my account of knowledge's nature to subject-oriented internalist accounts. The reason is for this, again, is that my overriding concern in this paper is with first-personal skeptical challenges, and there is something distinctly suspicious about externalist responses to skepticism that leave the subject thinking: "If I stand in the right causal (veridical belief-forming) relations with reality then I have true beliefs about it; if I don't stand in such relations then I don't have true beliefs about it. Now, I wonder whether I do stand in such relations."[22] Because questions about which sorts of causal relations obtain between us and the

---

[21] Thus, James writes, "Nowadays, we have a soft determinism which abhors harsh words, and, repudiating fatality, necessity, and even predetermination, says that its real name is freedom; for freedom is only necessity understood, and bondage to the highest is identical with true freedom." William James, *The Will to Believe* (New York: Dover, 1956): 149.

[22] This, in essence, is Stroud's (Barry Stroud, *Understanding Human Knowledge: Philosophical Essays* (Oxford: Oxford University Press, 2000) response to Sosa's (Ernest Sosa, "Philosophical Scepticism and Epistemic Circularity," *Proceedings of the Aristotelian Society: Supplementary Volume* (1994)) contention that a 'metaepistemic' bias which privileges the centrality of justification of beliefs by other beliefs underlies the skeptic's reasoning. Stroud's response is an ingenious one. Suppose we imaging an externalist Descartes who takes the sign of a belief's veracity to consist in its external Divine causal origin, as indicated by its clarity and distinctness. Certainly, we wouldn't accept this account on the grounds that the conviction that this alleged belief forming method is accurate is itself veridical *on its own lights*. Granted, we may (and hopefully do) take the modern externalist's story about his belief-forming mechanisms to be true and the externalist Descartes' story to be false. But this is not to the point. What is relevant is that the modern externalist is in no better a position than the Cartesian externalist to provide a complete account of human knowledge by explaining how he knows that his causal story is accurate.

world are no less problematic for the internalist than are questions about any other aspect of mind-independent reality, presuppositions about these relations, just like presuppositions about context itself, can be treated as nothing more than mere notional content that we contribute to our own processes of belief formation.

On any recognizable internalist understanding, knowledge is a relation between minds and the world mediated by accessible experiential evidence. Suppose we ask the following obvious question: "how could concerns about conversational context ever affect the knowledge relation, so construed, in any relevant way?" The skeptic, of course, has always offered an answer to this question: such conversational concerns couldn't possibly have such effects because, whereas both the world and the evidence are constitutive of mind-independent reality, our presuppositions about conversational context are not. Such presuppositions can count as nothing more than presuppositional contributions that come, not from the world, but from us. Certainly, conversational context dictated by practical constraints on counter possibility salience has no bearing on the truth or falsity of worldly propositions that we claim to know. Neither does it have bearing upon the available evidence for these propositions, even if it does have bearing upon our subjective willingness to take risks. It is this very feature of our epistemic condition that lands us in ignorance on skeptical accounts, as it is this very feature of our claim-making practices that motivates concerns over whether our modes of belief modulation based on evidence lead us closer to or farther away from true belief. Skepticism arises because we are unable to determine the extent to which our own conjectures about the origins of our evidence inform our subsequent beliefs about the world, the character of which is underdetermined by such evidence. To anthropomorphize a bit, neither the world nor the evidence care about either our conversational context or our presuppositions about said context. We may care, but this only highlights the fact that whatever conclusions we draw as a result of such care reflect the ways of humanity rather than the way of the world. This criticism of contextualism, variously stated, is a common one.

Common or not, however, this criticism is a powerful one. What it points to is the fact that the skeptic's criterion of knowledge is not merely one amongst many. It is not some super-stringent variation of ordinary epistemic standards. Instead, it arises from distinctive and fundamental deliberations upon the question of how our mere assumptions about context could ever be relevant to the question of whether we know, that is, the question of how the practice of defining parochial "contexts of inquiry" as background suppositions about what we may take for granted could ever acquire epistemological traction. We cannot claim that knowledge is made possible, in contextualist fashion, by our adoption of various contexts of presupposition

without saying something about what makes these contexts of presupposition presumable. To say that context-defining background assumptions are chosen for interest-relative pragmatic reasons does not address this concern. For, if we take the target of knowledge to be justified true belief (at minimum), then we must be concerned with the question of how assumptions about context (conversational and otherwise) could ever serve as indices of either justification or truth. *It is through the invocation of the skeptical invariant's absolutist, limiting conception of truth that we naturally express this concern.*

Consider the strongest criterion of knowledge for which the skeptic is infamous, i.e., "absolute certainty." To avoid confusion, we must emphasize that such certainty is not some mere psychological state of "feeling quite sure," except with more oomph; rather, it is the (presumably unrealizable) state of fully satisfying the demands of strong epistemic deductive closure. The critical point to note here is the following. This conception of knowledge should not be viewed as the result of some unmotivated and irrational decision to impose arbitrarily high standards upon ordinary epistemic practice. Instead, it should be viewed as a result of the skeptic's attempts to question how assumptions about context, *qua* assumptions, could ever bear upon our knowledge claims, of how the evidence we cite for our worldly claims could ever count as evidence. A "context of inquiry" is best understood as the set of defining presuppositions that determine what we may take as given when we pursue further questions about knowledge. To skeptically question how our common assumptions about context can take us from evidence to true conclusions is to engage in an inquiry the generality of which robs the notion of "degrees of justification" of functional purchase because it requires that we entertain comprehensive skeptical scenarios. Comprehensive Cartesian-style skeptical scenarios describe maximal possible worlds that jointly exhaust the whole of logical space, leaving little or no presuppositional material behind with which we might judge common sense and refined scientific realism to be more likely than its various skeptical alternatives. From within such scenarios, we can agree that known facts about conversational context determine the range of counterpossibilities meriting consideration while having no access to "conversational facts" with which to delimit this range. Thus, to ask if such realism rather than, say, Berkeleyan idealism, obtains is to put ourselves in an untenable epistemological situation. This is not merely because both scenarios account for all available empirical evidence. It is also because each scenario offers its own take on the nature and authority of such non-empirical criteria as theoretical comprehensiveness and simplicity.

Knowledge for the skeptic thus becomes all or nothing; the range of epistemic states intermediate between absolute certainty and abject ignorance collapses to a

single point like the converging ends of a broken accordion. This explains why Bayesianism, for example, must always remain inadequate before the skeptical challenge. The skeptic's investigative stance is general in a way that puts all presupposition out of bounds, keeping the Bayesian from plugging initial and revised probability values into her likelihood revision machinery. In a skeptical context, Bayes' theorem would have to serve as a tool for moving from an initial pre-evidential conviction in common sense realism's truth to a revised degree of confidence formed in light of evidence of true and false positives provided by observation. But, of course, it could never do this. No matter what our prior probability ascription might be, we could never update this ascription because doing so would require observational tests able to distinguish between the truth of common-sense realism and its skeptical alternatives.

The problem with contextualist construals of skepticism is that they automatically address parochial assessments of knowledge attributions made under limited and particular circumstances. Skepticism, however, evaluates all our putative worldly knowledge at once by asking how knowledge could ever emerge in one fell swoop from something that isn't knowledge.[23] The skeptic's challenge doesn't arise from within one context amongst others. The skeptic challenges us to explain our ability to adopt particular contexts of presupposition. The skeptic's stance is one from which no presuppositions about worldly reality are available. However, this stance doesn't derive from merely one undistinguished set of intuitions we have about truth or evidence. Instead, it derives from fundamental considerations concerning the underdetermination of belief by experience. DeRose is thus off-base when he dismisses the skeptic's scenarios as no more motivated than his own "deeply felt conviction" that he knows such scenarios don't obtain. He misdiagnoses the skeptic when he accuses him of merely "playing king of the mountain."[24] This depiction represents the skeptical challenge as a mere formal exercise rather than what it is, the result of consistently pressing the demand for justification to its uniquely consistent end result.

Similar considerations apply in the case of "agency." In fact, the commonalities that the agency case has with the "knowledge" case above help us discern a problem that contextualist accounts often have when dealing with absolute terms. As in the case of knowledge, skepticism about agency is not best viewed as the result of imposing arbitrarily high standards upon the notion of "freedom." Rather, it results from deep considerations about the issue of what would have to

---

[23] Barry Stroud, *The Significance of Philosophical Scepticism* (Oxford: Oxford University Press, 1984).

[24] Keith DeRose, "Solving the Skeptical Problem," *The Philosophical Review* 104, 1 (1995): 50.

obtain for assumptions about context even to be relevant to agency. Remember that on the account we have been considering, the defining feature of libertarian invariantist agency is "ultimate origination." An action is free, on such an account, only if an agent is its uncaused cause. Other criteria of agency vary from this in terms of the degree to which they dictate that events merely flow through, rather than from, an agent, in the sense of being caused by personal desires of which one is not the cause, or else modulated by higher-order desires and values of which one is not the cause.

However, this is hardly the end of the story. We also talk as though people's values and first and second-order desires as free to varying degrees, and we measure this level of freedom in terms of prior conditioning and influence within and outside the agent's original control. In this manner, there is a privileged sense of "freedom," just as there is a privileged sense of "knowledge," that designates a regulative ideal to which we naturally refer in our attempts to determine how assumptions about context could possibly be relevant to our ascriptions of agency. On simple and complex compatibilist theories, we endeavor to define free actions as those that flow from our desires or are at least modulated by our second-order desires and values. However, libertarian theories aim to provide a more fundamental account of what it is for these desires and values *to be ours*, of why the fact that these actions follow from these desires and values makes these actions ours. In this way, they aim to explain how the fact that an action is caused by our wants or regulated by our higher-order desires and values could ever bear upon the question of whether said action is free. Epistemic contextualism leaves us with the question of how assumptions about the likes of conversational context could have bearing on such philosophical concerns about the possibility of knowledge? Compatibilist theories of free will leave us with a formally identical question: how could facts about the likes of second-order desires address philosophical concerns regarding ultimate origination? These are perfectly well-motivated concerns irrespective of any hopes we might have to define them away. We can choose to define "knowledge" and "free will" differently, but in doing so we cannot claim to be doing anything other than changing the subject to address entirely different, and arguably less fundamental, concerns.

In short, the manner in which Substantive EC misconstrues the challenge posed by epistemic skepticism about knowledge closely parallels the way in which free will contextualism misconstrues the challenge posed by libertarianism or incompatibilist skepticism about agency. We commonly speak of both knowledge and freedom as coming in degrees: knowledge because different background assumptions affect our probability assessments, freedom because different degrees of background conditioning affect our liberty of action. What both epistemological and

incompatibilist skeptics note, however, is that these background assumptions cannot be based on considerations of sakes, interests, needs, or practical purpose if we are to avoid changing the subjects at hand. In both cases, extreme limiting demands on the phenomenon at issue (knowledge and agency, respectively) cannot be legitimately discarded as mere fetishes. The regard for certainty does not arise from the imposition of arbitrarily extreme demands on "knowledge." It stems from fundamental concerns about our ability to delineate contexts in which less restrictive, more specialized demands on knowledge are appropriate. The regard for unconditioned agency doesn't arise from the arbitrary imposition of extreme demands on "free action." It arises from concerns about our ability to delineate contexts in which less restrictive, more specialized demands on liberty are appropriate. Concerns about both certainty and unconditioned agency arise from distinctive and fundamental deliberations over the question of how our assumptions about context could ever be relevant to questions of whether we know or if our actions are free. The demands of both the skeptic and the libertarian express concerns of general interest. Their criteria of knowledge and freedom, respectively, are not mere parochial sets of criteria amongst others. Substantive contextualism fails because it diagnoses skeptical and libertarian concerns as confined to contexts of inquiry no more interesting than others.

6. Conclusions

That matters concerning use have little or no direct interest for epistemology is hardly surprising. It reflects an obvious fact about the so-called "new linguistic turn" in philosophy. In the heyday of positivism, semantic analysis was taken to have epistemological and (anti)metaphysical consequence only because a background account existed that purported to explain the pertinent connections. This background account constituted a rationale for construing philosophical issues in linguistic terms via a verificationist theory of meaning that aimed to describe the limits of our knowledge by reference to language's sensory provenance, thereby throwing a wet blanket over our metaphysical aspirations by segregating meaningful from non-meaningful talk. The problem with DeRose-type Semantic EC is that it comes equipped with no such doctrine. It is, in a very real sense, semantic analysis devoid of the kind of accompanying doctrine required to give it philosophical consequence. This is not to say that the strategy of logical positivism was successful. It is only to say that it was a strategy.

On the other hand, the various strategies of ordinary language philosophy, on which we might take defenders of the Implication Thesis to rely, have never been convincingly articulated as constituting a systematic program. If the point of

Semantic EC is to highlight the richness of language and remind us that "know" enjoys different uses in different contexts, then this point has no force against the above argument, which does nothing to deny such richness.[25] If the point of Semantic EC is that we must take language at face value because inquiries into natures (in this case, the nature of knowledge) are inseparable from their associated linguistic inquiries (because we have no language-independent methods for studying such natures), then it is false.[26, 27, 28] The fact that reality can only be described in language (i.e., that we can only talk *in* language) does nothing to imply that reality lacks a language-independent character, which some languages are better at capturing than others. If the point of Semantic EC is that the skeptic's sense of "know" is offered without an explanation of how it is to be used, then it is false again: the use of skeptical "knowledge" is clearly analyzed in terms of the demands imposed by epistemic deductive closure. If the point of Semantic EC is that skeptics offer conclusions about knowledge which are, in fact, merely disguised prescriptions of certain forms of speech, then it is false again: the skeptic poses a specific factual question about whether any effective evidence at all is available to those operating from a subject-regarding, internalist stance.[29] If the point of Semantic EC is to proffer a paradigm case argument that ordinary knowledge assertions are meaningful to the extent that they mark significant distinctions in linguistic use, then we need take no exception to it.[30] Nothing said in this paper flies in the face of the idea that term meanings are partly identified by reference to their common instances of use. We can grant that some contextually defined relation (or set of relations) is picked out by the term "know," but still deny that it is the relation which distinctively philosophical deliberations present to us for consideration and review. Finally, if the point of Semantic EC is that skeptical arguments play on attempts to pass off specialized senses of "know" as ordinary senses of "know," then its point, once again, is implausible. This is the contention against which I have argued at length in this paper.

It may be this last "ordinary language" strategy that many Semantic Contextualists have in mind. DeRose, for instance, claims that he "find[s] skepticism

---

[25] J. L. Austin, *How to Do Things with Words* (Oxford: Oxford University Press, 1962).

[26] J. L. Austin, "A Plea for Excuses," in *Ordinary Language*, ed. V. C. Chappell (Englewood Cliffs: Prentice-Hall, 1964): 41-63.

[27] Stanley Cavell, "Must We Mean What We Say?," in *Ordinary Language*, ed. V. C. Chappell (Englewood Cliffs: Prentice-Hall, 1964): 75-112.

[28] John McDowell, *Mind and World* (Cambridge: Harvard University Press, 1994).

[29] Stanley Cavell, "Must We Mean What We Say," 75-112.

[30] Norman Malcolm, "Certainty and Empirical Statements," *Mind* 52 (1942): 18-36.

persuasive and [merely] wants to "explain the persuasiveness of the skeptic's attack."[31] Cohen similarly maintains that contextualism "preserves our belief that we know things" while "explaining the undeniable appeal of skeptical argument."[32] On their simplest construal, such remarks amount to the suggestion that skeptical arguments turn on equivocations that skeptics fail to recognize and which Contextualism describes and explains. It is difficult to take this critique seriously, however. It strains credulity to suggest that the skeptical Cartesian tradition arose from nothing more than a failure to recognize a simple equivocation that is easily recognized and conveyed after a moment's reflection and conversation. Hence, I offer my alternative account: skeptical worries are motivated by the realization that parochial criteria of knowledge fail to address fundamental questions about how contexts are to be identified.

Beyond displaying the inability of semantic facts to provide evidence for Substantive EC, the primary contention of this paper is modest. Thus, it is important to note what I have *not* endeavored to show. First, I obviously do not deny that we invoke different criteria for "knowledge" and "freedom" when we use these terms in different conversational contexts. Of course, we do. It is only by doing so that we render them useful in everyday life.

Nor have I denied that there might be *non-semantic* grounds for accepting Substantive EC. In the epistemic case, there might be non-semantic evidence that Substantive EC, now taken to include the contention of the Continuum Account – that all senses of knowledge are on a par – is true. One might argue, à la Michael Williams, that the notion of "evidence" I invoke is crucially equivocal because sensory knowledge is not intrinsically more secure than other kinds of knowledge (i.e., that there are no "epistemic natural kinds").[33] One might do this by arguing that the causal/representational account of the object/evidence/mind relation I invoke to justify a phenomenal basis for evidence (and the skeptical motivations that flow from them) are suspect. In the free will case, they might argue that the very notion of libertarian agency is unintelligible, perhaps because agent causation proves epiphenomenal with respect to event causation, thus leaving libertarian "freedom" with no more content than the idea that human action is mysteriously dredged from the existential abyss.

One *inadequate* line of response, however, is that of asserting the Implication Thesis through the simple reiteration of unargued-for contextualist presuppositions themselves. That is, it does not do to maintain that the arguments I have offered are

---

[31] DeRose, "Assertion, Knowledge, and Context": 168.

[32] Stewart Cohen, "Contextualism and Skepticism," *Philosophical Issues 10* (2000): 100.

[33] Williams, *Unnatural Doubts.*

themselves of merely contextualized pertinence on the alleged grounds that they stem from philosophical methods or principles that enjoy application only from within a limited context of philosophical inquiry. For instance, one cannot reject the causal/representational account of perception by claiming that in ordinary life, we presuppose something more like direct realism and that it is this doctrine that correctly directs our talk and assumptions about context. I have described a causal/representational model underlying skepticism (i.e., a model on which our representations remain distinct from the world we take them to represent) at a very high level of generality, and thus in a way that makes it dependent upon few if any particularizing problematic assumptions. To take exception to it, therefore, requires not its mere eschewal, but an intelligible account of how direct realism or some alternative account of perception, might work. Nor can one object to my argument against the Implication Thesis by reasserting the "continuum account" without responding to my principled reasons for thinking that the continuum account is false (given its problematic assumptions concerning how assumptions about context could ever be relevant to the question of whether knowledge obtains). There is something deeply illegitimate about simply specifying from the outset that the distinctly general questions about the possibility of knowledge presented by traditional epistemology critique of contextualism are too parochial to enjoy general application.

A second *inadequate* line of response bears one last repetition. DeRose offers it himself. This is his claim that the sort of reasoning I present rests on a "levels confusion" between subject's and contributor's knowledge. His contention, he maintains, is *not* that whether a subject can know depends on non-truth-relevant factors. Rather, it is that whether an attributer can truthfully describe a subject as 'knowing' depends on such factors. The reason: such factors can affect the precise content of a third-person attributer's claim without changing the subject's own epistemic state. Because I have identified skeptical concerns as those that raise the issue of how context could ever be relevant to knowledge and located such skeptical concerns within the realm of first-person phenomenally based epistemic claim-making, however, this meta-semantic response is of little help. In the cases I am considering, attributers of knowledge to subjects ("speakers") and subjects of knowledge attributions ("subjects") are one and the same. As I have explained, they must be if we are to take contextualism, either semantic or substantive, to have any relevance to skepticism at all. Thus, it can hardly be adequate to say that one knows even when one cannot truthfully claim to know. This claim is not merely mysterious; it is unintelligible.

Ron Wilburn

Contextualism comes in many flavors and fragrances, equipped with many bells and whistles. However, I submit that whichever subtleties may characterize whatever contextualism *du jour* one may have in mind, these subtleties are simply irrelevant to the most interesting issues at hand. The reason is that questions about knowledge and questions about freedom both arise from highly distinctive considerations, but considerations which, by virtue of this distinctiveness, manage to exercise broad critical authority over questions we ask about these relations in other, more prosaic, contexts.

What I have argued is that the Implication Thesis is false: facts about usage alone concerning our shifting indexical criteria for applying "knowledge" and "freedom" imply nothing concerning the phenomena of knowledge and freedom themselves. To many, this must seem like a thesis barely worth mentioning. However, the reason why this is the case is an interesting one: limiting, invariantist conceptions of "knowledge" and "freedom" express special concerns which grant them broad critical authority over questions we should ask about two these relations (i.e., knowledge and freedom) in other, more prosaic, contexts. The presuppositions that all senses of "knowledge" and "freedom," respectively, are on a par presuppose continua analyses that under-appreciate the unique authority of "knowledge" and "freedom" understood in absolute terms. What these understandings do is help us articulate our most fundamental concerns over the question of how invocations of context could ever be of epistemic or metaphysical relevance – of how our assumptions about context could be evidence and how our wants and desires could be our own. What I claim to have shown, in other words, is the following: even if there are numerous different uses of "knowledge," this fact does nothing to show that the meanings we ascribe to "knowledge" are of equal philosophical interest, or that the consequences of these senses are equally restricted to their own limited contexts of use. I thus take issue with DeRose's earlier quoted contention that contextualist reasoning "opens up possibilities for dealing with issues and puzzles in epistemology." On the contrary, I suggest that it closes more possibilities than it opens. Simple observations about use in themselves imply nothing of deeper philosophical significance. Some might regard this claim as atavistic; I obviously disagree. I maintain that it should hardly come as a surprise. To again quote DeRose, albeit to contrary argumentative ends, "how could it be otherwise?"[34]

---

[34] DeRose, *The Case for Contextualism*, 18.

INFERENCE AND SELF-KNOWLEDGE

Benjamin WINOKUR

ABSTRACT: A growing cohort of philosophers argue that inference, understood as an agent-level psychological process or event, is subject to a "Taking Condition." The Taking Condition states, roughly, that drawing an inference requires one to take one's premise(s) to epistemically support one's conclusion, where "takings" are some sort of higher-order attitude, thought, intuition, or act. My question is not about the nature of takings, but about their contents. I examine the prospects for "minimal" and "robust" views of the contents of takings. On the minimal view, taking one's premise(s) to support one's conclusion only requires focusing on propositional contents and putative epistemic support relations between them. On the robust view, taking one's premise(s) to support one's conclusion also requires knowledge (or being in a position to have knowledge) of the attitudes one holds toward those contents. I argue that arguments for the Taking Condition do not entail or sufficiently motivate the robust view. Accordingly, contra several philosophers, the Taking Condition does not illuminate a deep relationship between inference and self-knowledge.

KEYWORDS: inference, Taking Condition, self-knowledge, second-tier thought, cognitive agency

1. Introduction

Plausibly, drawing an inference can be an agential phenomenon: it can be a psychological process or event that is predicable of the agent herself rather than of her sub-agential cognitive mechanisms.[1] But what makes inference an agential phenomenon when it is one? A growing cohort of philosophers argues that inference involves agency because drawing an inference requires the agent to (1) have a "take" on how her premise(s) confer epistemic support on her conclusion, and (2) to draw her conclusion on the basis of this take.

Paul Boghossian offers the following formulation of this suggestion:

---

[1] Philosophers often distinguish agent-level and sub-agential inference via the distinction between system 2 and system 1 processing. See Keith Stanovich and Richard West, "Individual Differences in Reasoning: Implications for the Rationality Debate," *Behavioral and Brain Sciences* 23, 5 (2000): 645-665; See also Daniel Kahneman, *Thinking Fast and Slow* (Macmillan Press, 2011). In this paper all talk of inference/reasoning refers roughly to system 2 (agent-level) inference/reasoning. I say 'roughly' because these systems may have blurry lines. See, for this reason, Paul Boghossian's talk of "system 1.5 and up" inference/reasoning in his "What is Inference?" *Philosophical Studies* 169 (2014): 1-18.

Benjamin Winokur

> (**Taking Condition**): Inferring from *p* to *q* necessarily involves the thinker *taking p* to support *q* and drawing *q because* of that fact.[2]

If the Taking Condition (TC) is true, then agent-level inference (hereafter just *inference* or *reasoning*[3]) is not a wholly first-order process. Rather, the agent must have an intermediating conception of the quality of epistemic support between her premise(s) and conclusion.

Besides arguing for TC itself, its proponents must also clarify exactly what taking one's premise(s) to support one's conclusion amounts to. On this score, philosophers have variously argued that "takings" are beliefs,[4] intuitions,[5] mental actions,[6] and *sui generis* mental states.[7] I will not be adding to this particular debate here, nor will I be questioning whether TC is true.[8] Instead, I will ask a different question about TC, one that concerns the *contents* of takings. Specifically, I will ask:

---

[2] See Paul Boghossian, "Inference, Agency, and Responsibility," in *Reasoning: New Essays on Theoretical and Practical Thinking*, eds. Brendan Balcerak Jackson and Magdalena Balcerak Jackson (Oxford University Press, 2019), 101-124, 110. He adapts this proposal from Gottlob Frege's *Posthumous Writings*, edited by Hermes Hans, Friedrich Kambartel, and Friedrich Kaulbach (Oxford: Basil Blackwell, 1979). In "Inferring By Attaching Force," *Australasian Journal of Philosophy* 97, 4 (2019): 701-714, Ulf Hlobil notes that a version of this idea goes as far back as Bertrand Russell, "The Nature of Inference," *The Athenæum* 4694 (1920): 514-15.

[3] Some distinguish reasoning from inferring, such as Nicholas Koziolek, "Inferring as a Way of Knowing," *Synthese* (2017), https://doi.org/10.1007/s11229-017-1632-4. Like most others, however, I use these terms interchangeably.

[4] See Christian Kietzmann, "Inference and the Taking Condition," *Ratio* 31, 3 (2018): 294-302; Nicholas Koziolek, "Inferring as a Way of Knowing;" Andreas Müller, "Reasoning and Normative Beliefs: Not Too Sophisticated," *Philosophical Explorations* 22, 1 (2019): 2-15.

[5] See Elijah Chudnoff, *Intuition* (Oxford University Press, 2013); John Broome, *Rationality Through Reasoning* (Wiley Blackwell, 2013).

[6] Ulf Hlobil, "Inferring By Attaching Force."

[7] Boghossian, "What Is Inference?"

[8] Skeptics include Kieran Setiya, "Epistemic Agency: Some Doubts," *Philosophical Issues: A Supplement to Noûs*, 23 (2013): 179-198; Crispin Wright, "Comments on Paul Boghossian, 'What Is Inference'," *Philosophical Studies* 169 (2014): 27-37; Conor McHugh and Jonathan Way, "Against the Taking Condition," *Philosophical Issues* 26, *Knowledge and Mind*, (2016): 314-331; Alex Kiefer, "Literal Perceptual Inference," in *Philosophy and Predictive Processing: 17*, eds. Thomas Metzinger and Wanja Wiese (MIND Group, 2017), 1-19; Luis Rosa, "Reasoning Without Regress," *Synthese* 196 (2017): 2263-2278; Ladislav Koreň, "Have Mercier and Sperber Untied the Knot of Human Reasoning?," *Inquiry* (2019), https://doi.org/10.1080/0020174X.2019.1684988; Susanna Siegel, "Reasoning Without Reckoning," in *Reasoning: New Essays on Theoretical and Practical Thinking*, eds. Brendan Balcerak Jackson and Magdalena Balcerak Jackson (Oxford University Press, 2019), 15-31; Mark Richard, "Is Reasoning Something the Reasoner Does?," in *Reasoning: New Essays on Theoretical and Practical Thinking*, eds. Brendan Balcerak Jackson and Magdalena Balcerak Jackson (Oxford University Press, 2019), 91-100.

does an agent's taking her premise(s) to support her conclusion require that she have self-knowledge of the attitudes she bears toward her premise(s) and conclusion? Boghossian appears to think that it does. Focusing on theoretical inferences, he espouses:

> **(Self-Awareness Condition)**: 'Person-level reasoning [is] mental action that a person performs, in which he is either aware, or can become aware, of why he is moving from some beliefs to others.'[9]

His view seems to be that this Self-Awareness Condition (SAC) is either a direct upshot of TC or is a different way of articulating it.

On the assumption that talk of self-awareness is interchangeable with talk of self-knowledge (ditto for talk of self-consciousness), and on the assumption that TC indeed leads to or amounts to SAC, TC may underpin an interesting "agentialist" account of self-knowledge.[10] Agentialist accounts of self-knowledge argue that, due to the nature of our agency, self-knowledge cannot be acquired by ordinary empirical methods (though they can sometimes leave open exactly how self-knowledge is acquired instead). Here, an agentialist might claim that, because inference presupposes self-knowledge (as per SAC), at least one ordinary empirical route to self-knowledge—the inferential route—is closed off. At least, it follows that not all self-knowledge can be acquired inferentially.[11] Some agentialists may even want to argue for something stronger here. After all, if self-knowledge or something close to it is required in order to infer, and if a capacity for inference is basic to our rational agency, then self-knowledge of at least some of our mental states may seem

---

[9] Boghossian, "What Is Inference?," 16. The label comes from Siegel, "Inference Without Reckoning." Others who appear to endorse SAC include Sebastian Rödl, *Self-Consciousness* (Harvard University Press, 2007), Matthew Boyle, "Transparent Self-Knowledge," *Proceedings of the Aristotelian Society*, *Supplementary Volumes,* 85 (2011): 223-241, and Kietzmann, "Inference and the Taking Condition."

[10] I take this label from Brie Gertler, "Self-Knowledge and Rational Agency: A Defense of Empiricism," *Philosophy and Phenomenological Research* 96, 1 (2016): 91-109, and Ben Sorgiovanni, "The Agential Point of View," *Pacific Philosophical Quarterly* 100, 2 (2018): 549-572. Sorgiovanni further distinguishes between non-substantive and substantive agentialist accounts. Non-substantive agentialist accounts (my focus) account for the functions of self-knowledge while potentially leaving open its exact source(s). Substantive agentialist accounts, in contrast, argue that rational agency is a source of self-knowledge.

[11] *Pace* Gilbert Ryle, *The Concept of the Mind* (University of Chicago Press 1949; reprinted in 2009 by Routledge). Some epistemologists argue that, while self-knowledge is inferentially acquired, the relevant inferences are sub-agential. See, for example, Quassim Cassam, *Self-Knowledge for Humans* (Oxford University Press, 2015). These accounts are harder to dismiss by way of the present agentialist suggestion.

like a necessary rather than contingent feature of our psychological lives. In that event, any account of self-knowledge that treats it as a contingent cognitive achievement will be objectionable.

To be sure, this agentialist account may have its limits. First, note that SAC includes a crucial "or can become aware" qualifier. As such, SAC does not entail that we actually have self-knowledge in inferring. Now, as we will see below, this qualifier is dropped by several arguments that might be offered for SAC. But even in its unqualified version this agentialist argument has its limits. For one thing, it *only* proposes a necessary connection between rational agency and self-knowledge of those mental states that figure into one's inferences. So pains, tickles, various emotions, and more will not fall under the purview of this agentialist account. Moreover, this account might only deliver a necessary connection between rational agency and self-knowledge *during the inferential process*, such that it says nothing about self-knowledge of standing attitudes that are not occurrently deployed in one's inferences.[12]

Despite the limits of this agentialist account, it is surely interesting if SAC is true. In what follows, however, I will argue that extant arguments for TC do not establish SAC. In other words, I will argue that SAC is *not* equivalent to (or an upshot of) TC. This means that, even if TC is true, no agentialist conclusions like the above follow. To reach this conclusion, I will evaluate many arguments for TC. These will be arguments that appeal to TC in order to illuminate: (1) the inference/association distinction, (2) the good/bad inference distinction, (3) a Moore-paradoxical phenomenon associated with inference, (4) inference as a mental act, (5) inference as involving cognitive agency, and (6) rational responsibility for inferences. I will argue that none of these arguments lead us from TC to SAC. Before I evaluate these arguments, however, I want to articulate a conception of inference that accepts TC but rejects SAC. Once we have this conception of inference in view, we can see whether the arguments for TC encourage us to accept this conception instead of one that also accepts SAC.

2. Robust and Minimal Inference

Answering the question of whether taking your premise(s) to support your conclusion requires self-knowledge depends on whether there can seem to be any

---

[12] On this point, one might respond (with some further machinery in place, no doubt) that SAC grounds a constitutive connection between rational agency and those mental states that are "available" for inferential application, whether or not they are occurrently embedded in an inference at any given time. To see how this might go, see Sydney Shoemaker's discussion of available belief in "Self-Intimation and Second-Order Belief," *Erkenntnis* 71, 1 (2009): 35-51.

alternative conception of takings. Consider, then, the following toy inference schemas:

(Minimal): p. p provides sufficient epistemic support for q. Therefore, q.

(Robust): I believe that p. p provides sufficient epistemic support for q. Therefore, I now believe that q.

At least on the surface, the difference between (Minimal) and (Robust) is that the former does not involve thoughts about mental states. Nevertheless, (Minimal) seems to involve some sort of appreciation (taking) of epistemic support: it seems to involve what we might call a 'meta-propositional' as opposed to a 'meta-attitudinal' taking-attitude.[13] Thus, on what I will call *the minimal view*, there can be instances of agent-level inference that involve nothing over and above what Christopher Peacocke calls "second-tier" thought: thought that is about "relations of support, evidence, or consequence between contents."[14] Such thought is 'second' rather than 'first' tier because it represents epistemic relations between propositional contents rather than *just* representing said contents. But it is not (we might say) 'third-tier' because it does not also include content about the subject's own attitudes toward these contents. *Ex hypothesi*, then, a putatively second-tier inference like (Minimal) cannot require the agent to conceptualize p and q *as believed*. It also seems dangerous to allow that p and q are conceptualized in such an inference as *reasons for belief*. Instead, the epistemic concepts deployed in appreciating an epistemic support relation between p and q should squarely concern indicators of what is or is (probably) *true*. Thus, concepts like EVIDENCE and CONSEQUENCE may be involved, as long as these capture relations between propositions only.

Of course, one can harbor different attitudes toward p, and it is true that whether one hopes, believes, or doubts p can have drastic effects on what sorts of inferences one can and should draw from p. Because of this, it may seem that inferring something from p obviously requires self-knowledge, since it might seem to require one to know the specific attitude one has toward it. However, John Broome has articulated an alternative possibility, according to which a second-tier reasoner need only view p "in a believing way" in theoretical inference,[15] or to view

---

[13] Cf. Philip Pettit, "Broome on Reasoning and Rule-Following," *Philosophical Studies* 173 (2016): 3373-3384.

[14] Christopher Peacocke, "Our Entitlement to Self-Knowledge: Entitlement, Self-Knowledge and Conceptual Redeployment," *Proceedings of the Arisototelian Society* 96, 1 (1996): 117-58. See especially pp. 129-130.

[15] "A Linking Belief is not Necessary for Reasoning," in *Reasoning: New Essays on Theoretical and Practical Thinking*, eds. Brendan Balcerak Jackson and Magdalena Balcerak Jackson (Oxford University Press, 2019), 32-43.

it in an intending or desiring way in the practical case. The idea is that viewing a content in a believing, desiring, or intending way is not equivalent to having a higher-order belief about that attitude qua attitude. This may sound strange, but it need not. After all, first-order propositional attitudes are just the same: one views *p* in a believing way *by believing it*, for instance. Similarly, second-tier inferences might involve taking-attitudes that involve viewing (merely perceived or actual) epistemic support relations between propositions in a believing way.

In fact, what I am calling the minimal view has been repeatedly presented as a major problem for a different and highly influential agentialist account of self-knowledge, one that argues for a constitutive connection between self-knowledge and our capacity for a specifically "critical" form of reasoning.[16] Tyler Burge describes critical reasoning as reasoning undertaken with the aim of evaluating and adjusting one's attitudes so that they better conform to norms of rationality, and he argues that such reasoning requires a non-empirical form of access to one's mental states. The objection from the minimal view is that agents can adjust their attitudes just fine, and in a sufficiently critical way, even if they never conceptualize their attitudes as such and only focus, instead, on propositional contents and relations of epistemic support between them.[17] Now, if SAC is true, then the minimal view is in danger. However, the significance of Burge's agentialist account will also be threatened, since we won't need to focus on *critical* reasoning to establish an important connection between inference and self-knowledge. Because of this, making the case for or against the minimal view is also significant for debates between Burge-style agentialists and their opponents.

Opposed to the minimal view is what I will call *the robust view*. This is the view that I take authors like Boghossian to accept, given their acceptance of SAC. On this view, agent-level inference necessarily requires at least being in a position to appreciate epistemic support relations between one's own inferential mental states themselves. Crucially, this qualifier allows proponents of the robust view to grant that agents do occasionally second-tier infer. What they must argue is only that any agent that can second-tier infer is also in a position to know her inferentially embedded mental states. Nevertheless, as aforementioned, we will

---

[16] Burge, "Our Entitlement to Self-Knowledge," 98.

[17] See David Owens, *Rationality Without Freedom* (London: Routledge, 2000) and "Deliberation and the First-Person," in *Self-Knowledge*, ed. Anthony Hatzimoysis (Oxford University Press 2011), 261-278; Cassam, *Self-Knowledge for Humans*; Annalisa Coliva, *The Varieties of Self-Knowledge* (Palgrave Macmillan: London, England 2016); Anna-Sara Malmgren, "On Fundamental Responsibility," *Philosophical Issues: A Supplement to Noûs* 29, 1 (2019): 198-213; Broome, "A Linking Belief is Not Necessary for Reasoning."

examine many arguments for the robust view that drop this qualifier and so aim to establish a stronger connection between inference and self-knowledge.

We can now clarify the question of this paper as the question of whether the minimal view or the robust view is true. Before we turn to evaluating arguments for TC in order to see whether SAC follows, such that the robust view prevails over the minimal view, four final preliminary points are in order.

First, as one might have guessed, I will be focusing (with Boghossian) on *theoretical* inference in what follows and, even more narrowly, on *non-suppositional* inferences where one's premises are one's actual beliefs. These decisions are intended to simplify and shorten discussion, though I cannot rule out the possibility that they might cause me to unwittingly bypass important complications.[18]

Second, let me pre-empt a possible concern: if second-tier taking-attitudes are so much as *intelligible*, doesn't it follow that second-tier inference is possible and, hence, that we can second-tier infer? And if so, isn't the robust view obviously false? The answer to this second question is *no*. For it is open to a proponent of the robust view to say that, while second-tier taking-attitudes might be *components* of agent-level inference, they must be supplemented with self-knowledge, because of further desiderata for any good account of (agent-level) inference. Alternatively, as aforementioned, some proponents of the robust view can grant that we can second-tier infer, while arguing that agents who can second-tier infer must still be in a position to know their inferentially implicated mental states.

Third, I do not deny that *some* inferring might require self-knowledge, as when we reason explicitly about whether we ought to have the beliefs that we do. But now one could ask: why can't an agentialist simply argue that some inference requires self-knowledge? I suppose one can argue this, but it should disturb us that such a view trivializes the desired connection between self-knowledge and inference, for it amounts to saying no more than that inferences that are explicitly about one's own mental states require us to have knowledge of them. Importantly, philosophers like Boghossian make a stronger claim: they claim even inferences about matters besides one's own mental states require one to (be in a position to) know how one's mental states epistemically support other mental states.

---

[18] One possible complication comes from John Broome, who sees taking-attitudes in practical inferences as unintelligible. See his "Comments on Boghossian," *Philosophical Studies* 169 (2014): 19-25, and "A Linking Belief is not Necessary for Reasoning." For a rejoinder, however, see Markos Valaris, "What Reasoning Might Be," *Synthese* 194 (2017): 2007-2024. Another possible complication is that practical inference may require self-knowledge for reasons having nothing to do with TC—see Owens, "Deliberation and the First-Person."

Benjamin Winokur

A fourth and final point concerns the question of how "explicit" one's self-knowledge must be according to the robust view. I have been writing as if self-ascriptive, higher-order beliefs involving mental state concepts must be involved, but this does not tell us whether such beliefs can amount to "implicit," or "tacit," or "backgrounded" instances of self-knowledge.[19] I will typically characterize the robust view as the view that inferring requires taking-attitudes with a conceptualized, psychological self-ascriptive structure, consciously foregrounded or otherwise. This is partially to avoid repeatedly appending cumbersome qualifiers about the possibilities of tacit or backgrounded taking-attitudes to claims made in the ensuing discussion. But it is also because I am somewhat skeptical of the idea of tacit self-knowledge,[20] and so take as my target the clearest (by my lights) version of the robust view.[21]

3. Arguments for the Taking Condition

It is now time to look at the arguments for TC, and to see whether they entail or sufficiently motivate SAC, thus motivating the robust view over the minimal view. I will consider six such arguments in total (I run two together in §3.1) and show that they do not motivate the robust view.

---

[19] Müller, "Reasoning and Normative Beliefs," 8.

[20] See, e.g., Coliva, *The Varieties of Self-Knowledge*, chapter seven, for her critical discussion of Sydney Shoemaker's appeal to tacit self-knowledge in his *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press 1996).

[21] A final preliminary is not a caveat for what follows but, rather, a comment about a different dialectical context in which the robust view might be motivated. I have in mind the possibility that the robust view can solve a problem for TC even if it does not motivate it. Many philosophers worry that TC gives rise to an infinite regress inspired by Lewis Carroll, "What the Tortoise Said to Achilles," *Mind* 4, 14 (1895): 278-280. The regress worry is that, if taking-attitudes are themselves premises of inferences, they must also be taken by further-taking attitudes to support one's other premises, *ad infinitum*. There are various TC-friendly strategies for avoiding this regress—see, e.g., Kietzmann, "Inference and the Taking Condition;" Müller, "Reasoning and Normative Beliefs;" Marcus, "Inference as Consciousness of Necessity," *Analytic Philosophy* (2020): 1-19, https://doi.org/10.1111/phib.12153. Kietzmann's particular strategy appeals to the robust view. Here I simply note that his strategy, according to which the regress is halted by conceiving of one's taking-attitudes as ontologically indistinct from their objects, might be made to work on a second-tier conception of taking-attitudes as well. I also believe—though I cannot argue it here—that Marcus's strategy does not suffer from the defects that Kietzmann points out for other strategies, and that Marcus's account of inference does not seem to depend on the robust view.

3.1 Inference (Good and Bad) and Association

Inference is but one way for agents to form a mental state. A non-inferential cause of many mental states is *association*. An agent might associate by being subject to some "bizarre psychology experiment" where she is conditioned to think that the sun will one day explode every time she thinks that Donald Trump is the President of the United States.[22] Or we might imagine a more ordinary case of a habitual depressive who, whenever she thinks about how much fun she is having, also thinks that there is so much suffering in the world.[23] Both cases involve mental states that are in *some* way sensitive to one another. But they are not sensitive in an inferential way. This is why it is not enough to say, as Hilary Kornblith does, that inference is simply a matter of "transitions involving the interaction among representational states on the basis of their content,"[24] for the interaction also needs to be of the right sort.

Might one argue that association and inference can be distinguished by pointing out that, in the latter case, one's mental states stand in epistemic support relations, whereas no such support obtains in the associative case? On the face of it, this proposal can even dispense with TC: it can account for the inference/association distinction in terms of the presence or absence of epistemic support relations, without countenancing any appreciation of these relations on the agent's part. One problem with this is that epistemic support relations can obtain between attitudes even when they are not occurrently involved in an inference.[25] Another problem is that that attitudes can be related inferentially despite the *absence* of genuine epistemic support relations between them. After all, as Boghossian says, this seems to be what happens with *bad* inferences.

In light of all this, Boghossian argues that:

> … something like a taking-based account seems not only natural, but forced: the depressive's thinking doesn't count as reasoning *not* because his first judgment doesn't *support* his second, but, rather, it would seem, because he doesn't *take* his first judgment to support his second. The first judgment simply *causes* the second one in him; he doesn't draw the second one because he takes it to be supported by

---

[22] Jake Quilty-Dunn and Eric Mandelbaum. "Inferential Transitions," *Australasian Journal of Philosophy* 96, 3 (2018): 532-547.

[23] Boghossian, "Inference, Agency, and Responsibility," 112.

[24] Hilary Kornblith, *On Reflection* (Oxford University Press, 2012), 55.

[25] E.g., one might associate thoughts in a way that matches the pattern of a modus ponens inference without actually being a modus ponens inference (Quilty-Dunn and Mandelbaum, "Inferential Transitions," 13).

the first.[26]

So, TC makes sense of the difference between associative and inferential mental state transitions as well as the difference between good and bad inferences. For, while in both good and bad inferences, the agent takes her premise(s) to support her conclusion, inference is only good when the epistemic support that she takes her premise(s) to provide is actual. Either way, the psychological relationship between the mental states at issue is not merely associative, because associations do not involve takings.

Supposing that Boghossian is right about TC's indispensability here,[27] what have we learned about the essential contents of taking-attitudes? I do not think that we have been led to the robust view. This is because proponents of the minimal view could argue that an agent who thinks that p, and that p supports q, infers q on the basis of appreciating (by her lights) an epistemic support relation between two world-directed propositions. In the good case, an agent takes p to support q and p does support q. In the bad case, she takes p to support q but p does not support q, say, because p is false or because there simply is no support relation between p and q even when both are true. In this way, it does not seem that she is required to think about her *beliefs* in p and q. Moreover, we can agree with Boghossian that she is not merely associating.

According to Nicholas Koziolek, however, this is too quick. He argues that we should prefer an account of the good/bad inference distinction that appeals to an agent's self-knowledge. To see this, consider first how he understands bad inferences: to infer badly is, on his view, to associate while taking it that you have actually formed a conclusion-belief on the basis of a premise-belief,[28] whereas good inferences involve no such mistakes. Koziolek grants that this is a paradoxical-sounding view, since it means describing bad inferences as, in fact, just perverse associations. Still, I will grant his claim that these deserve to be called inferences "if only by a sort of courtesy."[29]

On this new account, one's second-order perspective on one's attitudes is doing explanatory work, since appealing to this perspective allows us to countenance two mistakes involved in bad inferences: (1) a false belief that one's beliefs have inferentially caused one's conclusion in a non-deviant way (they haven't, since one's inference *does* involve deviant causation) and (2) a false belief that one has gained

---

[26] Boghossian, "Inference, Agency, and Responsibility," 112.

[27] Though see Siegel, "Inference Without Reckoning," for a non-TC account of the association/inference distinction.

[28] Koziolek, "Inferring as a Way of Knowing," 18.

[29] Koziolek, "Inferring as a Way of Knowing," 19.

knowledge via her inference (one hasn't, since the cause of one's conclusion-belief is not a rational cause). Crucially, while one surely *lacks* a degree of self-knowledge in making these mistakes, one still possesses some self-knowledge. One has self-knowledge insofar as one has knowledge *of the beliefs involved* in the defective inference, even though one lacks self-knowledge of the causal connections between these beliefs.[30]

This account allows us to make sense of bad inferences other than those where an agent takes *p* to support *q* even though, as it happens, *p* turns out to be false. This is especially interesting if Koziolek is right, as he argues, that inferences where *p* fails to support *q* merely because *p* turns out to be false are *not* actually bad. After all, though we can grant that such inferences fail to yield knowledge, it is still the case that, *were p to be true*, one's inference could yield knowledge. For this reason, Koziolek argues that such inferences are actually good because they are "potentially productive of knowledge."[31] So Koziolek's account helps us to distinguish other ways in which inferences can go awry. And since Koziolek's account appeals to the agent's self-knowledge, it may seem that we also need to accept the robust view.

Two concerns about Koziolek's account, for my purposes, are these. The first concerns a further feature of his account, which is that he avoids treating taking-attitudes as causal ingredients of inferences and argues instead that "takings are second-order beliefs about inferences…that are formed *in response to* the inferences themselves."[32] So the robust view of TC is not a consequence of his account, because his account actually rejects TC.

Of course, I have simply taken TC for granted in this paper in order to explore whether TC supports SAC. As a result, one may not be too worried about the above concern; after all, TC may turn out to be false. But even if TC is true, Koziolek's view may show that that proponents of the minimal view have an impoverished account of the good/bad inference distinction.

In response, I think that the minimal view can countenance at least two species of bad inference. First, bad inferences can involve agents who fail to accord *the right amount* of evidential weight to *p* in concluding *q*, even if they take *p* to support *q* to some degree and *p* really does support *q* to some degree.[33] In such a case, the agent's focus may strictly be on these world-directed evidential relations between propositions. Second, bad deductive inferences might have the following

---

[30] Koziolek, "Inferring as a Way of Knowing," 18.
[31] Koziolek, "Inferring as a Way of Knowing," 11.
[32] Koziolek, "Inferring as a Way of Knowing," 6.
[33] For a similar suggestion, not tethered to a second-tier conception of inference, see Siegel, "Inference Without Reckoning," 29.

structure: (1) an agent takes p to support q, (2) p does not in fact support q, and (3) the inference is *not* "potentially productive of knowledge." For example, someone who affirms the consequent infers badly, but this does not mean that her conclusion is deviantly caused by an association (as Koziolek would have it).[34] Rather, while she genuinely infers, she does so in accordance with a bad first-order inference schema, one that is not potentially productive of knowledge. In other words, affirming the consequent need not amount merely to associating p and q via, say, some sort of conditioning that produces a sort of mental "jogging"[35] that *looks like* affirming the consequent. A test for whether one merely associates in a way that looks like affirming the consequent, as opposed to actually affirming the consequent, is this: if an agent's mental activity satisfies the pattern of affirming the consequent for different values of p and q, this is evidence that she is not merely associating, since associations are content-based conditionings, whereas inference schemas generalize over different contents.

Perhaps only Koziolek's account can explain the bad inferences he describes. But whether being able to do this is something that *all* reasoners must be able to do simply in virtue of being reasoners is a different matter. The proponent of the minimal view can deny that second-tier reasoners can infer badly in Koziolek's sense while arguing that they can nevertheless draw bad inferences. Moreover, as we have seen, they can do so while accounting for the difference between association and inference. Two essential desiderata of an account of inference are hereby satisfied, but they appear to be satisfied along minimalist lines. The robust view has yet to follow.

3.2 Inference and Inferential Absurdity

Some philosophers have noted that inferences can figure into a version of Moore's Paradox. Thus, consider what Ulf Hlobil calls "Inferential Absurdity,"[36] or INFA:

> INFA. It is irrational, and transparently so from the agent's own perspective, to infer B from A1, ..., An and to believe also that these premises don't support B or to suspend judgment on whether they do.

Thus, if an agent infers q from p, but also believes that p does not support q (or is ambivalent about whether this is so), she is manifestly irrational. Hlobil points out that INFA has the air of a Moorean paradox. The reason, he says, is that inferences

---

[34] Contra Koziolek, "Inferring as a Way of Knowing," 19, fn. 29, who argues that there are no fallacious inferences without self-consciousness.

[35] Broome, *Rationality Through Reasoning*, 226.

[36] Hlobil, "Inferring by Attaching Force," 2.

are *acts*,[37] and acts are not content-bearing vehicles that can be in tension with *states* like beliefs (though of course inferences *operate on* content-bearing states like beliefs). So we face the question: "[h]ow can a doing that seems to have no content be in rational tension with a judgment or a belief?"[38]

Now, if TC is true, this absurdity is readily explained. For, as Christian Kietzmann puts it, "[i]f inference involves the thinker taking his premises to support his conclusion, this taking-attitude will clash with a belief or judgement that the premises do not support the conclusion."[39]

Once again, however, we can ask whether this explanation of INFA also motivates SAC and, hence, the robust view. On my view, the minimalist could accommodate INFA by explaining it in terms that do not require self-awareness, on the reasoner's part, of the tension between her inference and her belief that there is no epistemic support relation between the premise(s) and conclusion of her inference (or her agnosticism about this epistemic support relation). Perhaps this can be done by supposing that, while an agent draws an inference such as:

(**Minimal**): *p. p epistemically supports q*. Therefore, *q*.

While simultaneously believing either of:

(1) *p* may or may not support *q*, or;

(2) *p* does not support *q*

…then the contents of her mental states will contradict. But, *ex hypothesi*, (Minimal) and (1)-(2) are minimally contentful. They take propositional contents and relations between these as their objects, rather than taking mental states as their objects. The puzzle of how inferences (being acts, processes, or events) can stand in rational tension with beliefs (being states) is hereby explained.

What about the fact, as Hlobil sees it, that inferential absurdities are "transparently so from the agent's own perspective"? One way to go is to read this as a claim about a second-tier rather than a self-conscious perspective. But even if we are thinking about this perspective as a self-conscious one, the minimalist can also argue that it is unclear whether we must understand Hlobil's insistence on the transparency, to the subject herself, of inferential absurdities as an essential feature

---

[37] Though note that in an earlier paper he does not assume that inferential acts are intentional acts. See Ulf Hlobil, "Against Boghossian, Wright, and Broome on Inference," *Philosophical Studies* 167, 2 (2014): 419-429, specifically 421, fn. 1. Note that he does not assume that inferential acts are intentional acts. I take up the question of whether inferences are intentional acts in §3.3.

[38] Hlobil, "Against Boghossian, Wright, and Broome on Inference," 421.

[39] Kietzmann, "Inference and the Taking Condition," 295.

of the phenomenon. Hlobil seems to stipulate that it is. But if the minimalist's reply above is roughly correct, then inferential absurdities seem to be one thing and their self-conscious appreciation another. What proponents of the robust view need, then, is an argument for two claims: (1) that this appreciation is necessarily self-conscious (Hlobil does not argue for this) and (2) that this appreciation is necessary to being an agent-level reasoner. Unfortunately, (2) seems to implausibly assume that agent level reasoners are vulnerable to inferentially absurd states of mind unless they self-consciously ward them off. It is implausible because warding against irrationality is not always a matter of having to actively prevent one's self from adopting clashing attitudes, for part of what it is to be an agent is to be by and large rational.[40] But even if some such monitoring is required, why won't 'second-tier' monitoring do?

Perhaps it is worth closing this subsection by noting that other Moore-paradoxical phenomena might also be explained without appealing to an agent's self-knowledge, *even when* the phenomena necessarily invoke an agent's capacity to *self-attribute* mental states. Take, for example, the original Moore's Paradox, which concerns thoughts and utterances like "*p*, but I don't believe that *p*." Such thoughts or utterances seem highly irrational despite the fact that their conjuncts do not formally contradict. Now, as several philosophers have argued,[41] a defining featuring of ordinary, first-personal self-ascriptions of one's current mental states—what many call "avowals"[42]—is their *expressive* dimension. Thus, according to the expressivist view of avowals, I can express my first-order mental states by avowing them, whereas others can only *report* on them. What this mean is that my avowal, but not your ascription to me of the same state I avow, can directly manifest the avowed state. Expressivists have appealed to this function of avowals in order to account for our so-called first-person authority.[43] But they have also shown that this account of avowals lends itself to a "smooth account" of Moore's paradox.[44]

The expressivist account of Moore's paradox is, roughly, this. An avowal of "*p*, but I don't believe that *p*" *expresses* both a first-order belief that *p* and either (1)

---

[40] Cf. Donald Davidson, "Radical Interpretation," *Dialectica* 27 (1973): 314-28; McHugh and Way, "Against the Taking Condition," 322.

[41] See Rockney Jacobsen, "Wittgenstein on Self-Knowledge and Self-Expression," *The Philosophical Quarterly* 46, 182 (1996): 12-30; Kevin Falvey, "The Basis of First-Person Authority," *Philosophical Topics* 28, 2 (2000): 69-99; Dorit Bar-On, *Speaking My Mind: Expression and Self-Knowledge* (New York: Oxford University Press 2004).

[42] See Bar-On, *Speaking My Mind*.

[43] First-person authority is usually defined as a distinctive presumption of truth owed to or conferred on avowals. Alternatively, it is sometimes defined as the relative indubitability of by avowals.

[44] Jacobsen, "Wittgenstein on Self-Knowledge and Self-Expression," 28.

a first-order belief that not-*p*, or (2) a first-order agnostic attitude towards *p*. Because of this, the rational tension consists in a first-order "expressive conflict."[45] All the while, the utterance contains no contradiction at the level of its semantic meaning. Moore's paradox is dissolved when we recognize that expressive force and semantic meaning can come apart. But notice that the explanation does not invoke any claims about the speaker's self-knowledge. Rather, it invokes claims about a rational tension at the level of her first-order attitudes.

Crucially, these Moorean thoughts and utterances have an irreducibly *de se* conjunct. After all, if the target utterance or thought was merely "*p*, but not p" rather than "*p*, but I don't believe *p*", the irrationality manifest in such an utterance or thought would be visible at the semantic level, and would hardly constitute a puzzle of any sort. My point in raising the expressivist analysis of Moore's Paradox, then, is this. Because it is possible to give an account of it that does not depend on the subject's higher-order perspective on her attitudes (again, the tension exists as a first-order expressive conflict), despite the fact that the paradox cannot even be set up without granting that the subject self-ascribes her belief, this should make us doubly confident that INFA can be explained without appeals to self-knowledge. This is because there is at least an apparent possibility of understanding the semantic components of INFA (the inference and the incompatible belief) as second-tier rather than robustly contentful, unlike the original Moore's Paradox.

3.3 Inference and Practical (Self-)Knowledge in Action

Proponents of TC frequently argue that inference is *active*, and that its activeness explains why inference is attributable to the agent herself. The connection between taking and inferring can now be put as follows: "[a]ppreciating the support relation between premises and conclusion and drawing the conclusion on account of that appreciation seem to be things persons actively engage in. Inference will then count as something persons do because it involves the person-level activity of taking."[46]

Although I am confident that this thought motivates many philosophers to embrace TC, its implications for characterizing the essential contents of taking-attitudes are not straightforward. One might think that there are straightforward implications if one thinks that there is a constitutive connection between *mental action* and *practical knowledge* of what one is doing. One could take one's cue here from Elizabeth Anscombe, who famously argued that non-observational, non-testimonial, and non-inferential knowledge of what one is doing is constitutive of

---

[45] Bar-On, *Speaking My Mind*, 217-219.
[46] Kietzmann, "Inference and the Taking Condition," 295.

acting intentionally.[47] The idea now is that, if the activeness of inference is the activeness of intentional (mental) action, then TC could fall out as a consequence. For, if inference requires practical knowledge of what one is doing, one's taking-attitudes might be the very site of such knowledge: taking-attitudes might be the form of practical knowledge-in-inferring. Put differently: in inferring, one's practical knowledge of what one is doing takes the form of a taking-attitude to the effect that one's conclusion derives epistemic support from one's premise-beliefs.

In reply, one might question whether there really is a constitutive connection between intentional action and practical knowledge (however frequently these may come together as a matter of fact).[48] One might also wonder whether there can be another description under which one knows what one is doing, in inferring, that does not presuppose knowledge of one's mental states.[49]

I will not pursue these possibilities here. Instead, I will begin by raising the possibility that inference is not a species of action at all. Good evidence for this consists, as Kieran Setiya and Casey Doyle argue, in the grammar of inference-talk.[50] We do not say, for example, that we are in the middle of drawing an inference from p to q, even though we may be in the middle of considering whether p is evidence for q.[51] This suggests that the term 'inferring' and its cognates do not have the grammar of ordinary process verbs, in that they lack intelligible progressive aspects.[52]

But let us suppose, for the sake of argument, that inference may yet be a kind of action.[53] Another concern is that it does not seem to be a *voluntary* act whether one infers q from p, since one cannot simply *decide* to infer q from p; the alternative suggests an implausibly strong form of doxastic voluntarism that I will not bother to argue against here. But it is natural to think that intentional actions are standardly

---

[47] G. E. M. Anscombe, *Intention* (Oxford: Blackwell 1963).

[48] Juan Piñeros Glasscock, "Practical Knowledge and Luminosity," *Mind*, https://doi.org/10.1093/mind/fzz056

[49] Thanks to Dorit Bar-On for this suggestion.

[50] Setiya, "Epistemic Agency: Some Doubts;" Casey Doyle, *Four Essays on Self-Knowledge*, Dissertation: University of Pittsburgh. http://d-scholarship.pitt.edu/24970/

[51] Doyle, *Four Essays on Self-Knowledge*, 105. For similar observations, see Quilty-Dunn and Mandelbaum, "Inferential Transitions," 14.

[52] Similarly, believing also consists in taking a proposition to be true in some sense or other. But beliefs are states, not actions (Setiya, "Epistemic Agency: Some Doubts"; Nicholas Koziolek, "Belief as an Act of Reason," *Manuscrito* 41, 4 (2018): 287-318). Perhaps *judgements* are mental actions, but even these do not seem to be *intentional* ones.

[53] Perhaps Doyle's argument from the non-processual nature of inference to its non-actional nature is too quick if there can be instantaneous mental acts (this, it seems to me, is one way of reading Hlobil in "Inferring by Attaching Force").

voluntary actions, and so if inferences are not voluntary this may be another source of pressure against the claim that they are intentional.

It might be argued that inferences are intentional despite not being voluntary. On this argument, inferences involve *intentions-in* the act, analogously to how I can intentionally albeit reflexively (and so, in one sense, non-voluntarily) raise my arm as a basketball is being hurled at my head.[54] Such actions are not intentional because preceded by a decision, but because they have an appropriate means-end telos. However, it may seem that inferences are *never* such that one can simply decide to perform them, whereas even actions like raising my arm to block a basketball can *sometimes* be. Moreover, even when an action is reflexive and hence not voluntary, it is usually something that can be overridden: I can lower my arm quickly after it reflexively raises. But this is not possible with inference. With inference, one understands that there is an epistemic support relation between two propositions and draws a conclusion without choosing to do so or choosing to stop.[55] A plausible explanation is that inferences are not intentional actions.

It might be replied that inferences *can* in fact be voluntary, and so the argument that inferences are not intentional because not voluntary is a bad one. For example, David Hunter focuses on cases where an agent's evidence for or against *p* is strong enough to license an inference in either direction.[56] He concludes that it is up to you (i.e., is voluntary) whether you infer *p* or ~*p* from the evidence. But this at most shows that *some* inferences are voluntary. For, in cases where the evidence strongly favours only one conclusion *p*, Hunter agrees that one cannot voluntarily infer *p*. So, if voluntariness is a sign of intentional action, it will only be a sign that some corner cases of inference are intentional actions.

Alternatively, as Hunter argues, inference may involve self-knowledge simply in virtue of being voluntary, whether or not they are also intentional, since—following John Hyman[57]—he argues that voluntary actions depend on a lack of ignorance of what one is doing, even if they do not aim at desire-satisfaction in the way that intentional actions do. If correct, voluntary inferences will involve self-knowledge even if they are not intentional actions, but once again this will only be true for a small number of inferences, and only if there really are voluntary inferences. In any event, this will not follow from an obviously Anscombian thesis.

---

[54] The example comes from Parent, *Self-Reflection for the Opaque Mind: An Essay in Neo-Sellarsian Philosophy* (Routledge, 2017), 186

[55] Cf. Marcus, "Inference as Consciousness of Necessity."

[56] David Hunter, "Inference as a Mental Act," unpublished.

[57] John Hyman, *Action, Knowledge and Will* (Oxford University Press, 2015).

3.4 Inference and Cognitive Agency

Even if one agrees that inferences are not intentional actions, one might think that inferences are nevertheless active in *some* sense, and that TC can help us to understand this activeness. In other words, TC might still help us to understand how inference is a site of cognitive agency.

To see how we might proceed, note first that there exist fairly uncontroversial characterizations of the activeness of *mental states* that do not trade on the notion of an intentional action. Thus, consider Joseph Raz's claim that our "beliefs are a product and an aspect of our active nature because they are responsive to reasons."[58] The domain of active attitudes has contestable boundaries within and beyond the category of belief, but it is plausible that any attitudes that can figure into inferences are active.

Now, whatever philosophers typically mean when they talk about cognitive agency, Doyle doubts that they are merely parroting Raz's claim that many of our cognitive attitudes are active. This is because the existence of cognitive agency is typically treated as controversial,[59] whereas the responsiveness of (at least some of) our attitudes to reasons is not. A richer conception of cognitive agency will require, on Doyle's view, that we do better to explain what makes inference "*attributable* to the subject herself."[60] And to explain *this*, we must concede (so Doyle argues) that the agent's attitudes are not only rationally sensitive to other states that figure into her inferences, but that they are also "possessed in virtue of the agent's own assessment of their credentials."[61] With this point in mind, Doyle offers the following conception of inference: "I bring it about that I believe that *p* on the ground that *q* when I believe that *p* because I take it that this is what I should believe." In this way, "my sense of how things should be with my beliefs is explanatory of my believing as I do."[62] Notice that these taking-attitudes are *de se*: it is by taking it that *I ought to believe q* on the basis of *p* that I come to believe *q*.

One reply is simply that Doyle has not hereby ruled out a minimalist alternative to the picture he presents. Perhaps, then, the minimalist could say that

---

[58] Joseph Raz, "The Active and the Passive," *Aristotelian Society Supplementary Volume* 71, 1 (1997): 211-228. See also Scanlon's talk of "judgement-sensitive attitudes" in his *What We Owe to Each Other* (Harvard University Press, 1998).

[59] See, e.g., Setiya, "Epistemic Agency: Some Doubts" and Kornblith, "Epistemic Agency," in *Performance Epistemology: Foundations and Applications*, ed. Miguel Ángel Fernández (Oxford University Press, 2016), 167-182.

[60] Doyle, *Four Essays on Self-Knowledge*, 106.

[61] Doyle, *Four Essays on Self-Knowledge*, 106.

[62] Doyle, *Four Essays on Self-Knowledge*, 107.

inference involves cognitive agency to the extent that we are the ones who bring our mental states into cognitive contact with one another via second-tier taking attitudes in inferences. On this picture, my belief that q is responsive to my reason p because I take p to provide epistemic support for q, minimally understood.[63] If I am rational, we can imagine that my inference proceeds accordingly.

To better see why this minimalist alternative actually makes a good deal of sense, notice that it would be strange, if not outright disquieting, if a reasoner's awareness of p as providing rational support for q could never motivate an inference without self-knowledge. For, as David Owens rhetorically puts the point: "if you already have a non-reflective awareness of the reasons which ought to motivate you, how does the judgement that you ought to be moved by them help to ensure that you are so moved?"[64] An answer is that they do not. This means that beliefs like *q is what I should believe in light of p* look like "an idle wheel in our motivational economy."[65] If this is right, an agent's sense of how things should be with her beliefs is *not* explanatory of her believing as she does. Rather, her awareness *that p supports q* is explanatory. We can grant that the only reason why one would (rationally) take it that q is what one should believe on the basis of p is that one takes p to support q. And it is surely reasonable to believe that one ought to believe q on the basis of p if one takes p to support q. But then the question resurfaces: what additional motivational role, in coming to believe q, is played by this further (self-)judgement?[66] As I understand Owens' point, neither my sense of what I ought to believe nor the norms that constrain this sense must figure into the contents of my inference.

I am not denying that we sometimes place great stock in ensuring, from a self-conscious perspective, that we believe in accordance with our sense of how we ought to believe. I am only denying that such an aim is *constitutive* of agent-level inference, and that this aim must be represented in the contents of our taking-attitudes. Thus, Boghossian goes wrong in suggesting that the constitutive aim of inference is "figuring out what follows or is supported by other things one believes,"[67] where this is one's conscious aim *de dicto*.[68] An alternative aim of

---

[63] I am not arguing that mental states are *only* sensitive to each other in inferential episodes.

[64] Owens, *Rationality Without Freedom*, 18.

[65] Owens, *Rationality Without Freedom*, 18.

[66] Owens is originally responding to Burge's account of the importance of "critical reasoning" (see §2).

[67] Boghossian, "What is Inference?," 5.

[68] For a similar view see Nishi Shah and David Velleman, "Doxastic Deliberation," *Philosophical Review* 114 (2005): 497-534.

inference might just be to find out what is true, rather than what follows from what one believes.[69]

3.5. Inference and Rational Responsibility

Philosophers sometimes claim that an agent's *responsibility* for her inferences is a key motivator for TC. As Boghossian argues, while inferences are epistemically evaluable; *we* are also responsible for our inferences.[70] In Markos Valaris' words: "if you make a bad inference, we can legitimately criticize *you* as having been hasty, irresponsible, biased, and so on."[71] Put differently, we are *rationally responsible* for our inferences, such that we can be criticized for violating norms of inference. TC can purportedly explain this, for it enables us to say that we are rationally responsible for our inferences, and can be criticized accordingly, because *our* takings produce our inferences.

Crucially, if we can be criticized for inferring badly, it follows that we must be able to think about *ourselves* and, accordingly, our inferences *themselves*. So if Boghossian is right that it is constitutive of being a reasoner that we can be criticized for our inferences, then perhaps we really must always understand our inferences robustly: we must understand them as *being drawn by us*, which means understanding them as taking place in a mind. Alternatively, it may be that although the very process of inference can be second-tier, the second-tier reasoner must always be *in a position* to robustly conceptualise her inferences after the fact, insofar as she is already, at the moment of second-tier inference, a possible target for epistemic criticism down the line. In sum, we are rationally responsible for our inferences, but we could not be so if we were not in a position to know our minds. So, as reasoners, we are in a position to know our own minds.[72]

Besides the concern that this may weaken the agentialist's conclusion beyond any point of serious interest, since it never actually requires the agent to *have* self-knowledge, the main problem with this argument is that it doesn't tell us whether all reasoners really are rationally responsible for their inferences. Thus, consider Anna-Sara Malmgren's view:

> Small children can engage in effectual deliberation, at least of the primitive sort…

---

[69] This comes from McHugh and Way, "Against the Taking Condition," 325, despite their skepticism that this is the best way to characterize the aim of inference. Consider also Anna-Sara Malmgren: "the inferring agent needn't think of herself as settling, or trying to settle, what to believe…She just has to try and do it." ("On Fundamental Responsibility," 206).

[70] Boghossian, "Inference, Agency and Responsibility," 113.

[71] Valaris, "What Reasoning Might Be," 2010.

[72] This brings to light my qualification about the robust view in §1.

> But small children and most animals lack the kind of responsibility that rounds rational evaluability…
>
> …One might suggest that what makes the difference is the capacity for second-tier thought. But it's not entirely clear why, by itself, it would. The ability to (say) make judgments to the effect that such-and-such is a reason to ϕ, that this evidence outweighs that, or that a given claim implies another, might make one *better* at deliberation—in that it makes one better at conforming to the governing norms. But how is that, in turn, supposed to explain fundamental responsibility?[73]

Malmgren speaks of "effectual deliberation" of a "primitive sort" as a kind of inference that is not yet second-tier. If such reasoning is understood as agent-level but not taking-mediated, then proponents of TC must deny its possibility. No matter how they manage this, however, the more important bit in our dialectical context is Malmgren's insistence that even having a capacity for second-tier inference does not entail that one has a "fundamental responsibility" for one's inferences or the mental states that figure into them. Malmgren's use of the term 'fundamental responsibility' essentially refers to the appropriateness of being subject to various deontic norms in the formation and maintaining of one's attitudes, such that one is criticisable when one violates them. So Malmgren's view seems to be a relevant alternative to Boghossian's. They adopt a seemingly similar conception of rational responsibility, but Malmgren denies that such responsibility follows directly from our capacity to infer.

According to Malmgren, "[w]hat makes the difference" to being rationally responsible for our inferences or their products "are our *introspective* and *self-reflective* capacities."[74] This may be so, but it can be no help for proponents of the robust view. The reason is straightforward: if rational responsibility requires (a capacity for) self-knowledge, and (a capacity for) self-knowledge is something that must be added to the cognitive repertoire of individuals who are antecedently capable of drawing inferences, then the robust view is false.

Perhaps Boghossian could reply by simply *defining* agent-level inferences as those for which we are rationally responsible.[75] But why should we accept this definition? After all, if Malmgren is right, then it isn't a feature of *inference itself* that explains our rational responsibility for our inferences. Rather, it is an additional set of self-reflective capacities that explains this. In that case, Boghossian is not obviously capturing a desideratum of an account of agent-level inference; rather, he is capturing a desideratum of an account of self-reflective agent-level inference.

---

[73] Malmgren, "On Fundamental Responsibility," 207.
[74] Malmgren, "On Fundamental Responsibility," 207.
[75] Thanks to Ryo Tanaka for pressing this point.

What Boghossian needs to argue is that the TC should lead us to think that inference, at the agent level, already involves these self-reflective capacities. If he cannot do this, opponents of the robust view are free to deny this definitional maneuver.

4. Conclusion

In this paper I have evaluated six arguments for the robust view, which is the view that agent-level inference requires self-knowledge (or being in a position to have self-knowledge) of one's inferentially embedded mental states. These arguments are implausible. So far, then, agentialists who hope to vindicate deep connections between our agency and our self-knowledge have failed to do so by appealing to TC, even if TC is true.

Notably, this conclusion is not entirely negative, for the arguments of this paper simultaneously confer greater plausibility on Peacocke's suggestion, originally made in response to Burge, that our capacity for higher forms of self-aware inference is scaffolded onto a more "primitive" capacity for second-tier inference.[76] So long as this more primitive capacity captures core desiderata for our account of (agent-level) inference, inference and self-knowledge are not bound together by necessity. No doubt, we are often self-conscious when we infer, or can easily become such, but these do not appear to be necessary conditions for the basic process called inference that so many philosophers have sought to illuminate, this being inference that: is attributable to the agent herself, is distinct from association, has good and bad instances, has intelligible Moore-paradoxical instances, and is active.

---

[76] Peacocke, "Our Entitlement to Self-Knowledge," 129.

DISCUSSION NOTES/ DEBATE

WRIGHT ON MCKINSEY ONE MORE TIME

Simon DIERIG

ABSTRACT: In this essay, Crispin Wright's various attempts at solving the so-called McKinsey paradox are reconstructed and criticized. In the first section, I argue against Anthony Brueckner that Wright's solution does require that there is a failure of warrant transmission in McKinsey's argument. To this end, a variant of the McKinsey paradox for earned a priori warrant is reconstructed, and it is claimed that Wright's putative solution of this paradox is best understood as drawing on the contention that there is a transmission failure in the argument in question. In section II, I focus on Wright's views in the second part of his pivotal article on the McKinsey paradox (published in 2003). It is argued that the solution to the paradox proposed there by Wright is convincing *if* his theory of entitlements is accepted. In the third section, however, I raise an objection against Wright's account of entitlements. Finally, in section IV, Wright's views in his most recent essay on the McKinsey paradox are examined. It is shown that his new solution to this problem does not work any better than his earlier attempts at solving it.

KEYWORDS: McKinsey paradox, Wright's solution, externalism, self-knowledge, epistemology

I.

In an article on Crispin Wright's solution to the McKinsey problem, Anthony Brueckner has claimed that Wright's solution to this problem does not rely on the contention that there is a failure of warrant transmission in the McKinsey argument.[1] This claim is surprising given that Wright makes it quite clear that his

---

[1] See Anthony Brueckner, "Wright on the McKinsey Problem," *Philosophy and Phenomenological Research* 76 (2008): 389–390. For the McKinsey problem, see Michael McKinsey, "Anti-individualism and Privileged Access," *Analysis* 51 (1991): 9–16; Jessica Brown, "The Incompatibility of Anti-Individualism and Privileged Access," *Analysis* 55 (1995): 149–156; and Paul A. Boghossian, "What the Externalist Can Know A Priori," *Proceedings of the Aristotelian Society* 97 (1997): 161–175. For Wright's solution to this problem, see Crispin Wright, "Cogency and Question-Begging: Some Reflections on McKinsey's Paradox and Putnam's Proof," in *Skepticism*, eds. Ernest Sosa and Enrique Villanueva (Boston and Oxford: Blackwell Publishers, 2000), 140–163; Crispin Wright, "Replies," in *Skepticism*, eds. Ernest Sosa and Enrique Villanueva (Boston and Oxford: Blackwell Publishers, 2000), 201–219; Crispin Wright, "Some Reflections on the Acquisition of Warrant by Inference," in *New Essays on Semantic Externalism and Self-*

dissolution of McKinsey's paradox is, at least in part, based on the thesis that there is a transmission failure in the pertinent argument.[2] In this section, it will be shown that Brueckner's interpretation of Wright's solution is misguided because he has too narrow an understanding of what the McKinsey paradox amounts to.

Let me begin by explaining the distinction between warrant transmission and warrant closure as Wright has drawn it in a series of essays published at the beginning of this century.[3] The principle of *closure* of (a priori) warrant says that someone who has an (a priori) warrant for the premises of an argument and knows that these premises entail its conclusion has an (a priori) warrant for its conclusion. According to Wright, this principle must be distinguished from the principle of warrant *transmission*, which says that someone who acquires a warrant for the premises of an argument and, recognizing its validity, infers its conclusion from the premises, thereby acquires a warrant for the argument's conclusion.

In Wright's view, warrant transmission, but not warrant closure, fails in circular arguments as well as in Fred Dretske's well-known zebra argument.[4] The philosophical significance of Wright's distinction becomes clearer if one takes into account that the principle of warrant transmission also fails, in his opinion, in the following argument, which may be called the "McKinsey argument:"[5] I think that water is wet. If I think that water is wet, I have encountered water. Therefore, I have encountered water. (Wright holds that there is a failure of warrant transmission in

---

Knowledge, ed. Susana Nuccetelli (Cambridge, Mass., and London: The MIT Press, 2003), 57–77; Crispin Wright, "McKinsey One More Time," in *Self-Knowledge*, ed. Anthony Hatzimoysis (Oxford and New York: Oxford University Press, 2011), 80–104. For a related, though ultimately different, dissolution of McKinsey's paradox, see Martin Davies, "Externalism, Architecturalism, and Epistemic Warrant," in *Knowing Our Own Minds*, eds. Crispin Wright, Barry C. Smith and Cynthia Macdonald (Oxford and New York: Oxford University Press, 1998), 321–361; Martin Davies, "Externalism and Armchair Knowledge," in *New Essays on the A Priori*, eds. Paul A. Boghossian and Christopher Peacocke (Oxford and New York: Oxford University Press, 2000), 384–414; Martin Davies, "The Problem of Armchair Knowledge," in *New Essays on Semantic Externalism and Self-Knowledge*, ed. Susana Nuccetelli (Cambridge, Mass., and London: The MIT Press, 2003), 23–55.

[2] See Wright, "Cogency," 153–157; and Wright, "Some Reflections," 63–64.

[3] See Crispin Wright, "(Anti-)Sceptics Simple and Subtle: G. E. Moore and John McDowell," *Philosophy and Phenomenological Research* 65 (2002): esp. 331–332, and Wright's papers mentioned in fn. 1.

[4] See Fred Dretske, "Epistemic Operators," *The Journal of Philosophy* 67 (1970): 1015–1016; and Wright, "(Anti-)Sceptics Simple and Subtle," 332 and 342.

[5] See Wright, "Cogency," 153–157.

Moore's famous proof of an external world, too.[6] But in this essay I will not be concerned with this thesis.)

The McKinsey paradox can now be formulated as follows: I have an a priori warrant for the claim that I believe that water is wet. I also have an a priori warrant for the externalist contention that, if I believe that water is wet, I have encountered water.[7] Finally, closure holds for a priori warrant. Therefore, I have an a priori warrant for the claim that I have encountered water. But I cannot have an a priori warrant for such a specific claim about the external world.

Brueckner contends that Wright tries to dissolve this paradox by alleging that its last premise is wrong: I have an a priori "entitlement" to the claim that I have encountered water.[8] The term "entitlement" is Wright's technical term for a non-evidential and unearned a priori warrant.[9] According to Wright, an entitlement to a certain proposition "is conferred not by positive evidence for the proposition in question but by the operational necessity, so to speak, of proceeding on the basis of such an untested assumption if one is to proceed at all."[10]

But why does Wright think that I enjoy an entitlement to the claim that I have interacted with water? For Wright, I have an a priori entitlement to the "integrity of the concepts in terms of which I essay to formulate items of my self-knowledge (…)."[11] If externalism is true, a necessary condition for this integrity is the satisfaction of certain external conditions, in the present example: my having causally interacted with water. Thus, it can be inferred, according to Wright, that I have an a priori presumption or entitlement to the claim that I have encountered water.

---

[6] See Wright, "(Anti-)Sceptics Simple and Subtle," 330–348.

[7] The classical sources for externalism regarding the mental are Hilary Putnam, "The Meaning of 'Meaning'," in *Mind, Language and Reality. Philosophical Papers, Volume 2* (Cambridge, New York and Melbourne: Cambridge University Press, 1975), 215–271; Tyler Burge, "Individualism and the Mental," *Midwest Studies in Philosophy* 4 (1979): 73–121; Tyler Burge, "Other Bodies," in *Thought and Object. Essays on Intentionality,* ed. Andrew Woodfield (Oxford and New York: Clarendon Press, 1982), 97–120; and Tyler Burge, "Intellectual Norms and Foundations of Mind," *The Journal of Philosophy* 83 (1986): 697–720.

[8] See Brueckner, "Wright on the McKinsey Problem," 389–390.

[9] See Wright, "Some Reflections," 68–69; and Crispin Wright, "Warrant for Nothing (and Foundations for Free)?" *The Aristotelian Society. Supplementary Volume* 78 (2004): 174–175. For more on entitlements, see also Crispin Wright, "On Epistemic Entitlement (II). Welfare State Epistemology," in *Scepticism and Perceptual Justification*, eds. Dylan Dodd and Elia Zardini (Oxford and New York: Oxford University Press, 2014), 213–247.

[10] Wright, "Some Reflections," 68.

[11] Wright, "Some Reflections," 68.

Simon Dierig

One may object that I can only have empirical evidence for such a specific claim about the external world and that I cannot therefore have an a priori entitlement to it. For Wright, this argument is a *non sequitur*. He holds that "an a priori entitlement to a belief is quite consistent with the only envisageable kind of positive evidence for it being empirical. I can be a priori entitled to suppose that my senses are functioning adequately right now, though a check would need to be empirical."[12] Once it is realized that having an a priori entitlement to a claim about the external world does not require having a priori evidence for it, the thesis that I have such an entitlement does not seem absurd any longer.[13]

By employing the notion of an earned (and evidential) a priori warrant, Wright constructs the following variant of the McKinsey paradox:[14] I have an earned a priori warrant for the claim that I believe that water is wet. I also have an earned a priori warrant for the externalist contention that, if I believe that water is wet, I have encountered water. Finally, closure holds for earned a priori warrant. Therefore, I have an earned a priori warrant for the claim that I have encountered water. But I cannot have an earned a priori warrant for such a specific claim about the external world.

McKinsey's original paradox can be dissolved by maintaining that I enjoy an a priori *entitlement* to the claim that I have interacted with water. As already mentioned, this way of solving the original problem no longer seems absurd once it is realized that having an a priori entitlement to a proposition does not require having a priori evidence for it. What remains absurd, according to Wright, is, however, the idea that I could have an *earned* a priori warrant for a claim about the external world.[15] The variant of McKinsey's paradox just reconstructed cannot therefore be dissolved in the same way as the original paradox, viz, by claiming that its last premise is wrong.

Wright does not explain in detail how, in his opinion, the above variant of McKinsey's problem is to be solved. All he says is that its solution must have something to do with the fact that closure does not hold for earned a priori warrant.[16]

---

[12] Wright, "Some Reflections," 68–69.

[13] Jesper Kallestrup disagrees (see Jesper Kallestrup, "Recent Work on McKinsey's Paradox," *Analysis* 71 (2011): 168). In his view, the contention that one can possess an a priori warrant for a claim about the external world is absurd, no matter whether this warrant is evidential or non-evidential.

[14] "(…) but McKinsey is quite right that closure for reflectively earned armchair warrant would suffice to set the problem up." (Wright, "Some Reflections," 68)

[15] See Wright, "Some Reflections," 68.

[16] "(…) closure for reflectively earned armchair warrant (…). However, we don't have *that* principle – quite." (Wright, "Some Reflections," 68)

Wright endorses the closure principle both for warrant and for a priori warrant.[17] Indeed, he argues explicitly that closure for a priori warrant is true no matter whether transmission holds or fails.[18] So why not submit that closure holds for earned a priori warrant too?

The answer seems to be that, according to Wright, the chief argument for closure of earned a priori warrant is untenable because warrant transmission fails in certain cases. The argument in question proceeds as follows: First, if I have an earned a priori warrant for the premise "p" and also for the premise "If p, then q," and if warrant is transmitted from these premises to the conclusion "q," then I have an earned a priori warrant for this conclusion. Second, warrant is transmitted from the two premises to the conclusion. Therefore, closure is true for earned a priori warrant.

Provided the presupposed transmission principle founders, the main argument for closure of earned a priori warrant cannot be sustained. The above variant of McKinsey's paradox can therefore be dissolved by claiming that this closure principle is mistaken. Brueckner thus errs when he alleges that the issue of warrant transmission is irrelevant for Wright's solution to the McKinsey problem. To be sure, Wright's solution to McKinsey's original problem is independent of the issue of warrant transmission. But Wright considers a variant of the McKinsey problem for earned a priori warrant, and his solution to this variant is best understood as relying on the claim that there is a failure of warrant transmission in the McKinsey argument.

Before I proceed to the next section, let me sum up the textual evidence for my interpretation of Wright's position. First, he claims that his solution of McKinsey's problem draws partly on the thesis that there is a transmission failure in the McKinsey argument.[19] Second, since he endorses closure of a priori warrant,[20] the contention that there is a transmission failure plays in his opinion no role for solving McKinsey's original paradox. Third, he holds that the variant of McKinsey's paradox for earned a priori warrant must be dissolved by claiming that closure for earned a priori warrant is wrong.[21] From these three hypotheses it can be inferred that, according to Wright, the assumption of a transmission failure in the McKinsey argument must somehow lead to the rejection of closure for earned a priori warrant and thus to the solution of the variant of the McKinsey paradox for this kind of warrant. My proposal is that Wright thinks that, once it is realized that there is a

---

[17] See Wright, "Some Reflections," 67–69.
[18] See Wright, "Some Reflections," 68, first paragraph.
[19] See Wright, "Some Reflections," 63–64.
[20] See Wright, "Some Reflections," 67–69.
[21] See Wright, "Some Reflections," 68.

transmission failure in the McKinsey argument, the chief argument for closure of earned a priori warrant is inconclusive and this principle can therefore be repudiated. In brief, the issue of warrant transmission is, *contra* Brueckner, not irrelevant to Wright's solution of the McKinsey paradox.

II.

In the preceding section, I have focused on Wright's position in the first part of his essay "Some Reflections on the Acquisition of Warrant by Inference." In this and the next section, I shall concentrate on Wright's differing views in the second part of that article, more precisely, in section V. For the sake of convenience, I will henceforth talk of "Wright I" and "Wright II," respectively.

According to what has been called the "illusion version of anti-individualism,"[22] I suffer an illusion of content if I am in a no-reference situation such as Paul Boghossian's "Dry Earth."[23] On this planet, "all apparent interaction with watery substance is multisensory communal hallucination (…)."[24] The term "water" therefore has no reference there. Proponents of the illusion version of anti-individualism claim that an inhabitant of Dry Earth who says to herself "Water is wet" does not thereby express a thought even though her experience is subjectively indistinguishable from the experience of her actual-world counterpart.[25] In other words, our protagonist suffers an illusion of content.

Armed with these terminological explanations, Wright's account of entitlements in the second part of his article can be reconstructed as follows:[26] The proposition that all is in order with my concepts implies that I do not suffer an illusion of content. Provided the illusion version of externalism is true, this in turn implies that I am not an inhabitant of Dry Earth. My a priori entitlement to the claim that all is in order with my concepts therefore becomes an a priori entitlement to the contention that I am not on Dry Earth, i.e., that I have encountered either water or some other watery substance (such as, for example, twater).

This line of argument differs from Wright's argument in the first part of his essay in one important respect. Wright I claims that, given externalism, freedom from content illusion implies that I have interacted with water. He concludes that I have an a priori entitlement to the claim that I have had water encounters in the

---

[22] See Jessica Brown, *Anti-Individualism and Knowledge* (Cambridge, Mass., and London: The MIT Press, 2004), chap. 4, esp. 113–114.

[23] See Boghossian, "What the Externalist," 170.

[24] Wright, "Some Reflections," 64.

[25] See Wright, "Some Reflections," 64–65.

[26] See Wright, "Some Reflections," 68 and 71–72.

past. In contrast, Wright II holds that lack of water encounters does not imply that I suffer content illusion.[27] For I might be an inhabitant of Twin Earth, who has encountered twater, not water, and who does not suffer an illusion of content (because he has twater instead of water thoughts). But if freedom from content illusion does not imply that I have encountered water, one has no reason to think that one enjoys an a priori entitlement to the thesis that one has interacted with water. What one is a priori entitled to believe, according to Wright II, is rather that one does not inhabit Dry Earth. For freedom from content illusion implies – if the illusion version of anti-individualism is true – that one is not an inhabitant of that planet.

Wright's just-described innovation in the second part of his article has repercussions on the dissolution of McKinsey's paradox proposed by Wright in the first part of his essay. McKinsey's original paradox cannot be solved in the way suggested by Wright I, that is, by alleging that its last premise is wrong because I have an a priori entitlement to the proposition that I have encountered water. For, according to Wright II, I do not have such an entitlement. Moreover, the variant of McKinsey's problem for earned a priori warrant cannot by solved in the way proposed by Wright I either. The crux of his solution to the variant of McKinsey's paradox was that there is a transmission failure in the McKinsey argument. But, for Wright II, one has no reason to believe that there is such a failure in this argument.

Wright advances two reasons for the claim that warrant is not transmitted from an argument's premises to its conclusion. The first reason is that all four conditions of his "disjunctive template" are fulfilled.[28] The second reason can be formulated as follows:[29] I have an entitlement to, that is, an unearned a priori warrant for, an argument's conclusion. An unearned warrant cannot be transmitted from an argument's premises to its conclusion. Therefore, transmission fails in the argument in question.

According to Wright II, neither the first nor the second of these reasons can be used to establish that there is a transmission failure in the McKinsey argument. As to the first reason, the fourth condition of the disjunctive template is not fulfilled with regard to the McKinsey argument.[30] For the counterfactual "If I lacked water encounters, I would suffer an illusion of content" is not true: In some of the closest

---

[27] See Wright, "Some Reflections," 68 and 71–72.

[28] See Wright, "Some Reflections," 62–63 and 65–66.

[29] Here I follow Brueckner's interpretation of Wright's position (see Brueckner, "Wright on the McKinsey Problem," 388).

[30] See Wright, "Some Reflections," 69–70.

counterfactual situations in which I lack water encounters I am on Twin Earth, rather than on Dry Earth, and therefore do not suffer content illusion.

As to the second reason, in Wright II's view I do not have an a priori entitlement to the conclusion of the McKinsey argument, that is, to the proposition that I have encountered water. The first premise of the second reason is therefore wrong. In brief, neither the first nor the second reason shows that warrant is not transmitted from the premises of the McKinsey argument to its conclusion.[31] Wright I's strategy of dissolving the variant of McKinsey's paradox for earned a priori warrant is thus unsuccessful.

So much for the destructive consequences of Wright's philosophical innovations in the second vis-à-vis the first part of his essay. What about the constructive consequences? Does Wright II have the resources for a solution to the McKinsey problem which is more convincing than the solution proposed by Wright I? To answer this question, I will now try to reconstruct a dissolution of McKinsey's paradox which is suggested by some of Wright II's remarks and which seems persuasive, at least if one accepts Wright's general epistemological framework, that is, his theory of entitlements.[32]

According to Wright II, I have an a priori entitlement to the claim that I am not an inhabitant of Dry Earth, i.e., to the proposition that it is not the case that I have never encountered water or any other watery substance. It follows that I have an a priori entitlement to the claim that I have interacted with water or some other watery substance. Wright II also contends that I have an a priori warrant for alleging that the watery stuff of our actual acquaintance is water. Finally, closure holds for a priori warrant. From these claims Wright II infers that I have an a priori warrant for believing that I have encountered water. Hence, McKinsey's original paradox can be dissolved by maintaining that its last premise is wrong.

This solution to the McKinsey problem differs from Wright I's solution in that he argues that I enjoy an *a priori entitlement* to the contention that I have encountered water, whereas Wright II's just-reconstructed argument only shows that I have an *a priori warrant* for this contention. The question whether this a priori warrant is earned or unearned, and therefore an entitlement, can be settled if one takes into account that this warrant is transmitted from the premises to the conclusion of the following argument:

(1) I have encountered water or some other watery substance.

(2) The watery stuff of our actual acquaintance is water.

---

[31] See Wright, "Some Reflections," 72.

[32] See Wright, "Some Reflections," 71–72.

(3) Therefore, I have encountered water.

If we have an a priori warrant for the conclusion (3) (as has been shown above) and if this warrant is transmitted from the premises to the conclusion of the argument just put forward, then I have a transmitted – and therefore earned – a priori warrant for the conclusion (3), that is, the proposition that I have encountered water. But this means that the variant of McKinsey's paradox for earned a priori warrant can be dissolved by claiming that its last premise is mistaken.

There is, however, a rather obvious objection to this solution of the variant of the McKinsey problem (one I have already mentioned in the first section). The thesis that one has an a priori *entitlement* to a claim about the external world no longer seems absurd once it is realized that having an a priori entitlement to such a claim does not presuppose possessing a priori evidence for it. But this strategy of dispelling the impression of absurdity does not work with regard to the contention that one has an *earned* a priori warrant for a claim about the external world. For having an earned a priori warrant for a proposition presupposes having a priori evidence for it. It seems, then, that Wright I is right when he holds that having an earned a priori warrant for a claim about the external world is "paradoxical."[33]

Wright II might reply – and perhaps actually replies – that the thesis that one has an earned a priori warrant for the conclusion (3) only seems paradoxical if this conclusion is "taken in isolation."[34] If one bears in mind that I can deduce the conclusion (3) from premises about the external world to which I have an a priori entitlement – that is true for the premise (1) – or an a priori warrant – that is true for the premise (2) – it is no longer mysterious how I can have a priori evidence and therefore an earned a priori warrant for a claim about the external world such as the conclusion (3).

III.

It has been argued that Wright II has the resources for a convincing solution to the McKinsey problem if one is prepared to grant him his general epistemological framework. In this section, I will, however, level an objection against this framework, more precisely, against his theory of entitlements. I will not call into question Wright's claim that we are a priori entitled to assume that our perceptual system is reliable and that all is in order with our concepts. Rather, I want to raise a more specific doubt about the way he derives entitlements concerning the external

---

[33] Wright, "Some Reflections," 68.
[34] Wright, "Some Reflections," 72.

world from relatively unproblematic entitlements pertaining to the integrity of our conceptual system.

As already noted, Wright II argues that my a priori presumption or entitlement to the claim that all is in order with my concepts, together with the "known conceptual necessity" of the illusion version of externalism, "becomes an a priori presumption"[35] or entitlement to the claim that I do not inhabit Dry Earth and therefore have encountered either water or some other watery substance.[36]

This argument sketch may be reconstructed as follows: I have an a priori entitlement to the claim that all is in order with my concepts and that I do not suffer content illusion. I am also a priori entitled to contend that, if I do not suffer content illusion, I am not on Dry Earth. Finally, a priori entitlement is closed under logical implication. Therefore, I have an a priori entitlement to the claim that I am not an inhabitant of Dry Earth.

If formulated in this way, Wright's argument is questionable for at least two reasons. First, Wright has not given us any grounds for believing that a priori entitlement is closed under logical entailment. Second, the illusion version of externalism may be known a priori, but it is certainly not a claim to which one has an a priori entitlement.

Wright may respond that the claim that one has an a priori entitlement to the illusion version of anti-individualism need not be among the premises of his argument, and that it must rather be understood along the following lines: I enjoy an a priori entitlement to the claim that all is in order with my concepts and that I do not suffer content illusion. Given the conceptual necessity of the illusion version of anti-individualism, this claim conceptually entails that I am not on Dry Earth. Finally, a priori entitlement is closed under conceptual implication. Thus, I am a priori entitled to contend that I am not an inhabitant of Dry Earth.

The weak point in Wright's argument, if reconstructed in the way proposed, is its second premise. For the claim that I am an inhabitant of Dry Earth does not seem to imply conceptually that I suffer an illusion of content. If one is an inhabitant of Dry Earth, one might still think with the notion *water*, viz, in case one possesses theoretical knowledge of the constitution of water. In other words, it is possible that one thinks with the water concept even though one is an inhabitant of Dry Earth and has never encountered water or any other watery substance, namely if one knows, in this possible world, that water consists of hydrogen and oxygen.[37]

---

[35] Wright, "Some Reflections," 68.

[36] See Wright, "Some Reflections," 68 and 71–72.

[37] See Burge, "Other Bodies," 116.

The second premise of Wright's argument must therefore be revised so as to assert that the claim that I inhabit Dry Earth and have no knowledge of the chemical composition of water conceptually implies that I suffer content illusion with regard to the notion *water*. By contraposition, it follows that freedom from content illusion conceptually entails that it is not the case *that I am an inhabitant of Dry Earth and have no knowledge of the composition of water*. If one substitutes this contention for the original second premise, the conclusion of Wright's argument has to be modified as well, viz, like this: I am a priori entitled to believe that it is not the case *that I am on Dry Earth and have no knowledge of the composition of water*.

If McKinsey's protagonist had an a priori entitlement to the claim that he has no knowledge of the composition of water, he could infer from the conclusion of Wright's argument that he is a priori entitled to believe that he is not an inhabitant of Dry Earth. But the protagonist has at best an a priori warrant for, not an entitlement to, the claim just mentioned. Thus, Wright II is wrong when he alleges that one has an a priori entitlement to the contention that one does not inhabit Dry Earth.

At this juncture, one may wonder whether it is not sufficient for the success of Wright II's attempt to solve the McKinsey problem that one has an earned a priori warrant for, not an a priori entitlement to, the claim that one is not on Dry Earth and has encountered water or some other watery substance. For if one has an a priori warrant for, though not necessarily an entitlement to, the premises (1) and (2), one also has an a priori warrant for the conclusion (3), and this suffices for solving the McKinsey problem.

This way of advancing Wright II's dissolution of the McKinsey paradox is questionable on the ground that the strategy of dispelling an impression of absurdity used above to defend Wright II's original solution to the McKinsey problem cannot be employed to dispel the impression that it is absurd to suppose that one has an earned a priori warrant for the claim that one is not an inhabitant of Dry Earth. Recall what this strategy was. According to Wright II, I have an earned a priori warrant, and therefore a priori evidence, for the claim (3), that is, the proposition that I have had contact with water. The impression that it is absurd to suppose that one has a priori evidence for the claim (3) can be dispelled if one takes into account that one can deduce it from claims about the external world, viz, the propositions (1) and (2), for which I have an a priori warrant.

But this strategy of dispelling an impression of absurdity cannot be applied to the claim that I am not on Dry Earth and have had contact with water or some other watery substance. The impression that it is absurd that one has an earned a priori warrant and therefore a priori evidence for this claim cannot be dispelled by holding

that it can be deduced from premises about the external world for which one has an a priori warrant. For it seems that the only premises for which I have an a priori warrant and from which the claim in question can be inferred are "It is not the case *that I am on Dry Earth and have no knowledge of the composition of water*" and "I have no knowledge of the composition of water," and these premises do not exclusively concern the external world.

One may object that the requirement that the premises in question concern the external world is unnecessary. Why not dispel the impression that it is absurd that one has a priori evidence for the claim that I am not on Dry Earth by calling attention to the fact that this claim can be deduced from a priori warranted premises? This proposal will not do either. For someone who knows the variant of the McKinsey paradox for earned a priori warrant also knows that the claim that he has encountered water can be deduced from a priori warranted premises. Still, he is puzzled by the variant of the McKinsey paradox, more precisely, by the alleged fact that he has an evidential (earned) a priori warrant for his having encountered water. Therefore, calling attention to the fact that the claim in question can be deduced from a priori warranted premises alone does not dispel the impression that it is absurd that one has a priori evidence for this claim.

One might suggest, finally, that the impression that it is absurd that one has a priori evidence for the claim that one is not on Dry Earth can be dispelled by drawing attention to the fact that this claim can be deduced from a priori warranted premises *different from the premises of the McKinsey argument.* Suppose someone doubts that one can have a priori evidence for the claim in question. Suppose further that she still doubts this even after having been told that the claim in question can be deduced from the a priori warranted premises of the McKinsey argument. The doubt of this person might be dispelled by calling attention to the fact that the claim in question can *also* be deduced from a priori warranted premises which are *different* from the premises of the McKinsey argument, viz, the premises "It is not the case *that I am on Dry Earth and have no knowledge of the composition of water"* and "I have no knowledge of the composition of water."

This strategy of dispelling the impression of absurdity associated with the thesis that there is a priori evidence for the claim that I am not on Dry Earth is not conclusive. Strictly speaking, the second premise of the McKinsey argument runs as follows: If I believe that water is wet, it is not the case *that I have no knowledge of the composition of water and have never encountered water.*[38] The conjunction of this premise and the implicit premise that I have no knowledge of the composition of water implies that, if I believe that water is wet, I have at some time encountered

---

[38] See Burge, "Other Bodies," 116.

water. In brief, one of the premises of the McKinsey argument, correctly understood, is the contention that I have no knowledge of the composition of water. But this contention is also among the premises of the aforementioned alternative argument for the claim that I do not inhabit Dry Earth. There is thus no argument for this claim whose premises are a priori warranted and different from the premises of the (correctly understood) McKinsey argument. It seems, then, that the impression of absurdity associated with the idea that I have an earned a priori warrant for the claim that I am not on Dry Earth cannot be dispelled. Wright II's solution to the McKinsey problem appears to fail.

IV.

In his essay "McKinsey One More Time," Wright qualifies his earlier diagnosis of the McKinsey paradox in a number of ways. The most important of them is that Wright no longer contends that I am a priori entitled to claim that I am not an inhabitant of Dry Earth. Instead, he relies on the following case distinction:[39] Either I am entitled to the claim that I am not on Dry Earth, then Wright II's solution to the McKinsey problem is correct; or I am *not* entitled to that claim, then the McKinsey paradox can be dissolved by alleging that its first premise is wrong, i.e., by maintaining that I do not have an a priori warrant for the proposition that I believe that water is wet.

According to Wright, the second horn of this dilemma can be substantiated as follows:[40] If I am warranted to contend that I believe that water is wet, I have a warrant for the claim that I am not an inhabitant of Dry Earth. But my warrant for this claim is only empirical for I lack an a priori entitlement to it. Therefore, my warrant for the contention that I believe that water is wet is only empirical too. In other words, the traditional view that I have a priori knowledge of my own mental states is wrong.

One may rejoin that the proposition that I lack an a priori entitlement to the claim that I am not on Dry Earth does not imply that my warrant for this claim is at most empirical. For I might have an *earned* a priori warrant for the claim in question. This rejoinder is, however, not available to me because it has been argued in section III that the impression of absurdity associated with the idea that one has a priori evidence and therefore an earned a priori warrant for the claim in question cannot be dispelled.

---

[39] See Wright, "McKinsey One More Time," 101–102. (What follows in this and the subsequent paragraphs is a rational reconstruction of Wright's position.)
[40] See Wright, "McKinsey One More Time," 97; see also Wright, "Some Reflections," 75.

Simon Dierig

What about the first premise of Wright's defense of the second horn? Do I have a warrant for the claim that I am not on Dry Earth if I have a warrant for the contention that I believe that water is wet? The following three-step argument may have convinced Wright that the answer to this question is in the affirmative. First step: "In taking it that I have recognized (…) that I have a certain specific belief, I take it for granted that the apparent concepts configured in that belief are in good standing (…)."[41] From this Wright concludes that, if I have a warrant for the contention that I believe that water is wet, I have a warrant for the claim that all is in order with the concepts in question and that I do not suffer content illusion regarding them.[42] Second step: If (i) I have a warrant for the claim that I do not suffer content illusion, (ii) this claim conceptually entails that I am not on Dry Earth and (iii) warrant is closed under conceptual entailment, then I have a warrant for believing that I am not on Dry Earth.[43] Moreover, the claims (ii) and (iii) are true. Therefore, if I have warrant for the claim that I do not suffer content illusion, I have a warrant for believing that I am not an inhabitant of Dry Earth. Third step: From the conditionals argued for in step 1 and 2 it follows that, if I have a warrant for the contention that I believe that water is wet, I have a warrant for the claim that I am not on Dry Earth.

A refined version of this argument does indeed show that the first premise of Wright's defense of the second horn is true.[44] Moreover, it follows from the considerations put forward in section III that the second premise is true as well. Should one, then, endorse Wright III's diagnosis of the McKinsey paradox? In the following, I will advance two lines of argument which suggest a negative answer to this question.

First, the McKinsey paradox cannot be dissolved simply by denying that our knowledge of our own propositional attitudes is a priori. What is needed in addition is a strategy for dispelling the impression that it is absurd to claim that one must investigate the external world in order to know what one believes, thinks etc. But Wright has not provided such a strategy. His proposal for a solution to the McKinsey problem is therefore at best incomplete.

Second, the premises of Wright's argument in defense of the second horn are admittedly true. Still, his argument is not sound since its conclusion does not follow from its premises. One might think that Wright's argument *is* valid because one

---

[41] Wright, "McKinsey One More Time," 96.
[42] "The claim to warrant for the premises rests on the reasonableness of these presuppositions." (Wright, "McKinsey One More Time," 96)
[43] See Wright, "Some Reflections," 68, second paragraph.
[44] For the needed refinement, see section III.

confuses it with the following closely related, and valid, reasoning: To acquire a warrant for the contention that I believe that water is wet, I have to acquire a warrant for the claim that I am not an inhabitant of Dry Earth. But to acquire a warrant for this claim, I have to explore the external world. Therefore, I must examine the external world in order to acquire a warrant for the contention that I believe that water is wet. In other words, my warrant for this contention is not a priori but merely empirical.

This reasoning is valid, but it is not equivalent to Wright's original argument. For the first premise of that argument does not imply the first premise of the above reasoning. In general, a proposition of the type "If I have a warrant for A, I have a warrant for B" does not entail a proposition of the type "To acquire a warrant for A, I have to acquire a warrant for B." But if the reasoning outlined above is valid, whereas Wright's original argument is not, why not abandon his argument in favour of the above reasoning?

The answer is that Wright's defense of the first premise of his argument cannot be modified so as to support the first premise of the reasoning in question. In particular, the second step of his defense, with its reliance on warrant closure, cannot be amended in this vein. The only way to circumvent this difficulty seems to consist in dropping the second step altogether and in instead "extending" the first step. The modified version of Wright's defense would then look as simple as this: In taking it that I have a warrant for the contention that I believe that water is wet, I take it for granted that all is in order with my concepts and that I am *therefore* not an inhabitant of Dry Earth. Thus, to acquire a warrant for the contention that I believe that water is wet, I have to acquire a warrant for the claim that I am not on Dry Earth.

The premise of this defense is of course only true if the protagonist is an externalist. But this does not pose a problem for Wright since the McKinsey paradox can only be generated for proponents of externalism anyway. There is, however, a second, related, worry. According to Tyler Burge, the proposition that one is on Dry Earth does *not* imply that one's concepts are not in order and one suffers content illusion.[45] A Burgean externalist holds that an inhabitant of Dry Earth might think with the concept *water* provided he knows about the chemical composition of water.

If one claims that causal contact with water is necessary for the possession of the concept *water*, and that causal contact with twater is necessary for the possession of the concept *twater* (and so on for all concepts of watery substances), then one will quite likely also contend that being on Dry Earth implies that one suffers content illusion. But a Burgean externalist does *not* claim that causal contact with water is

---

45 See Burge, "Other Bodies," esp. 116.

necessary for the possession of the concept *water*. He only holds that what is necessary for A's possession of the concept *water* is the truth of the proposition "A has encountered water *or has knowledge of its chemical composition*." Therefore, a Burgean externalist need not contend that being on Dry Earth implies that one suffers content illusion. But if there are (Burgean) externalists who do not submit that the proposition "All is in order with my concepts" implies the proposition "I am not an inhabitant of Dry Earth," then the premise of Wright's defense is not true for everyone it is supposed to be true for, and his defense founders.

In sum, the second horn of Wright's dilemma can be substantiated neither with recourse to his original argument – because it is not valid – nor with recourse to the modified version of that argument – because his defense of its first premise is unconvincing. It seems, then, that Wright's diagnosis of the McKinsey paradox in his most recent article on this issue is not doing any better than his earlier attempts at dissolving the paradox.[46]

---

[46] Two final remarks. It has been argued in the first section of this essay that Brueckner is wrong in holding that Wright's solution to the McKinsey problem is independent of the issue of warrant transmission. This remains true with regard to Wright I's solution to the McKinsey problem. But if we consider Wright's later views, my earlier diagnosis must be qualified. For neither Wright II's nor Wright III's dissolution of McKinsey's paradox presupposes that there is a transmission failure in the McKinsey argument.

Wright III's solution to the McKinsey problem is also independent from how one answers the question whether closure holds for a priori warrant. This is obscured by the fact that Wright rejects this principle in the last section of his 2011 article (see Wright, "McKinsey One More Time," 102–103). Note, by the way, that this indicates a change of mind on Wright's part. For in his essay "Some Reflections" he agreed with McKinsey that closure of a priori warrant is true (see Wright, "Some Reflections," 68–69; and Michael McKinsey, "Transmission of Warrant and Closure of Apriority," in *New Essays on Semantic Externalism and Self-Knowledge,* ed. Susana Nuccetelli (Cambridge, Mass., and London: The MIT Press, 2003), 106–107).

THE POSS-ABILITY PRINCIPLE, G-CASES, AND FITCH PROPOSITIONS

Noah GORDON

ABSTRACT: There is a very plausible principle linking abilities and possibilities: If S is able to Φ, then it is metaphysically possible that S Φ's. Jack Spencer recently proposed a class of counterexamples to this principle involving the ability to know certain propositions. I renew an argument against these counterexamples based on the unknowability of Fitch propositions. In doing so, I provide a new argument for the unknowability of Fitch propositions and show that Spencer's counterexamples are in tension with a principle weaker than the one linking abilities and possibilities.

KEYWORDS: Poss-Ability, Fitch propositions, G-cases, abilities, knowability

Here's a highly intuitive principle linking abilities and possibilities:

> **Poss-Ability**: (For all S and all Φ) If S is able to Φ, then it is metaphysically possible that S Φ's.

In addition to being intuitive, Poss-Ability underwrites much contemporary theorizing about abilities. For example, it is a consequence of both of the leading metaphysical analyses of abilities,[1] and it mediates the connection between 'ought' and the 'can' of metaphysical possibility.[2]

Much hangs, then, on the truth of Poss-Ability. In this note, I argue against Jack Spencer's class of counterexamples to Poss-Ability by drawing on considerations from epistemic logic to renew an argument against these counterexamples based on the unknowability of Fitch propositions.[3]

Spencer's class of counterexamples centers around the following case:[4]

> **Simple G**: Consider a fully deterministic world w.[5] Let 'H' name the complete

---

[1] See 481-483 within Jack Spencer, "Able to Do the Impossible," *Mind* 126, 502 (2017):466-497.

[2] Spencer, "Able to Do the Impossible," 486-488.

[3] For other arguments against the G-cases, see Anthony Nguyen, "Unable to Do the Impossible," *Mind* 129, 514 (2020): 585-602.

[4] Spencer, "Able to Do the Impossible," 468.

[5] A world w is deterministic just in case a complete specification of the laws of w and the state of w at any time entails a complete specification of the state of w at any other time. See Carl Hoefer "Causal Determinism," *The Stanford Encyclopedia of Philosophy* (2016) for a similar definition of 'determinism.'

specification of the initial conditions of w and 'L' name the complete specification of the laws of nature of w. Suppose that H ∧ L is common knowledge in G's community; it is taught in high school physics classes. G is one the top students in her class and is excellent in physics as well. However, for one accidental reason or another G misses the class in which her teacher covers the laws and initial conditions and so never comes to believe H ∧ L or know H ∧ L, though many of her classmates do.

Spencer judges that in Simple G, G is able to know H ∧ L. This quickly leads to the falsity of Poss-Ability, since it is metaphysically impossible that G knows H ∧ L.[6] Simple G serves as something of a template for generating counterexamples to Poss-Ability. For example, in Gettier G, G truly believes H ∧ L but fails to know H ∧ L for Gettier-like reasons.[7] Following Spencer, I call cases like these G-cases.[8]

There is an argument from the unknowability of Fitch propositions against the G-cases.[9] Letting 'K_GΦ' regiment the claim 'G knows that Φ,' Fitch propositions are of the form 'Φ ∧ ¬K_GΦ.' The unknowability of Fitch propositions is the thesis that no one is able to know a Fitch proposition. In other words, it is the claim that:

Unknowability of Fitch propositions: ∀S∀Φ(S is not able to know that Φ ∧ ¬K_SΦ).

Spencer considers the following argument from the unknowability of Fitch propositions to the claim that G is not able to know H ∧ L: Suppose that G is able to know H ∧ L. H ∧ L entails ¬K_G(H ∧ L). Now if knowability for G is closed under entailment, then G is able to know that ¬K_G(H ∧ L). If knowability for G is further closed under conjunction, it follows that G is able to know (H ∧ L) ∧ (¬K_G(H ∧ L). But this is a Fitch proposition. It follows from the unknowability of Fitch propositions that our assumption is false, so G is not able to know H ∧ L.

I agree with Spencer that this argument is no good. The problem is that there is no reason to think that knowability for G is closed under entailment or conjunction. There are many ordinary agents for whom it is not plausible that they are able to know every consequence of anything they know. Moreover Spencer

---

[6] If H ∧ L is false, then G does not know H ∧ L, since knowledge requires truth. But if H ∧ L is true, then G does not know H ∧ L, since H ∧ L entails that G does not come to know H ∧ L. C.f. Spencer, "Able to Do the Impossible." 469 and 490.

[7] Spencer, "Able to Do the Impossible," 477-478.

[8] In this note, I restrict my attention to G-cases involving the ability to know. Spencer does discuss some counterexamples to Poss-Ability that do not involve the ability to know. For discussion of these, see Nguyen, "Unable to do the Impossible," section 4.

[9] Spencer, "Able to Do the Impossible," 489-491.

argues that if Fitch propositions are unknowable, then knowability is not closed under either conjunction or known entailment.[10]

However, there is an improved version of the objection which circumvents these problems. The crucial point is that the properties of knowability may differ when the abilities of different agents are in question. While knowability *in general* is not closed under entailment or conjunction, there is still a limited class of agents for whom their ability to know is closed under entailment and conjunction. In particular, consider the class of idealized agents who are both metaphysically omniscient and have perfect introspective access to their own epistemic states. A metaphysically omniscient agent C is such that:

> Whenever C knows all of the formulas in a set Γ (including the empty set) and A follows logically/metaphysically from Γ, then C also knows A.[11]

Metaphysical omniscience guarantees that knowability is closed under logical and metaphysical entailment. Perfect introspection (for an agent S) is characterized by at least the following principles:

KK Principle: $K_S\Phi \rightarrow K_S K_S\Phi$

Consistent Introspection: $B_S\neg B_S\Phi \rightarrow \neg B_S\Phi$.[12]

Given metaphysical omniscience, whenever S knows Φ and Φ entails Ψ, S knows Ψ. Given the KK principle, if S knows Φ and S knows Ψ, then S knows that S knows both of these. Thus if S is able to know Φ and Φ entails Ψ, S is in a position to know $\Phi \wedge \Psi$. So the relevant kind of closure invoked by the argument against G-cases does hold for this class of idealized agents.

The new and improved version of the argument then says that if an idealized version of G, namely G+, for whom the above conditions held were able to know H ∧ L, then they would be able to know a Fitch proposition, namely (H ∧ L) ∧ ($\neg$ K_{G+}(H ∧ L)). But since no one is able to know a Fitch proposition, G+ is not able to know H ∧ L. The final step of the argument is simply the claim that if a highly

---

[10] Spencer, "Able to Do the Impossible," 490-491.

[11] This definition appears in Rendsvig and Symons, "Epistemic Logic," *The Stanford Encyclopedia of Philosophy* (2019).

[12] Consistent Introspection follows from the belief analogue of the KK principle plus a weak principle on rational belief:

(1) $B_G\Phi \rightarrow B_G B_G\Phi$	(BB Principle)
(2) $B_G B_G\Phi \rightarrow \sim B_G \sim B_G\Phi$	(Weak Rational Belief)
(3) $B_G\Phi \rightarrow \sim B_G \sim B_G\Phi$	(1, 2, PL)
(4) $B_G \sim B_G\Phi \rightarrow \sim B_G\Phi$	(3, PL)

idealized version of G would be unable to know H ∧ L, then plain old G is not able to know H ∧ L simply for failure to satisfy the relevant closure conditions.

Against this improved version of the Fitch objection, several of Spencer's responses simply fall flat. First, I already mentioned that Spencer argues that if Fitch propositions are unknowable, then knowability is not closed under either conjunction or known entailment. But the relevant question is, knowability for whom? Both of Spencer's counterexamples here are for agents who lack knowledge of certain mathematical truths.[13] Assuming that mathematical truths are metaphysically necessary, these agents fail to be metaphysically omniscient. Therefore, these counterexamples cannot show that knowability is not closed under conjunction or entailment for the relevant idealized kind of agent.

Second, Spencer objects that:

> Although ¬K(H ∧ L) may be necessitated by H ∧ L, it need not be scrutable on the basis of H ∧ L, so there is no reason to think that G (or G+) ought to be able to know that ☐(H ∧ L →¬K(H ∧ L)). Thus, even if the threatening argument is sound – which I doubt very much – there would still be some G-cases that did not lead to the knowability of Fitch propositions, and so the main claims of this paper would stand.[14]

But I respond that if G+ is metaphysically omniscient, then there is eminently good reason to think that G+ ought to be able to know that ☐(H ∧ L →¬K_{G+}(H ∧ L)), since metaphysically omniscient agents know every metaphysically necessary truth. The second part of the quoted passage seems to suggest that even if the fact that H ∧ L entails ¬K_G(H ∧ L) is knowable for some G in some G case, say for metaphysically omniscient G+, we still might maintain that there are some G-cases which do not lead to knowability of Fitch propositions, particularly where G is not able to know ☐(H ∧ L →¬K_G(H ∧ L)). But, as I have suggested, this sort of reasoning is perverse. If a metaphysically omniscient version of G is not able to know H ∧ L, surely G is not able to know H ∧ L merely out of ignorance of the metaphysical truth that H ∧ L entails ¬K_G(H ∧ L).

Spencer's third response to objections based on the unknowability of Fitch propositions is the most interesting. He proposes to bite the bullet on the knowability of Fitch propositions, since the main argument for their unknowability depends on Poss-Ability. There is a simple proof that it is not metaphysically possible to know a Fitch proposition, first reported by Frederich Fitch.[15] The proof

---

[13] Spencer, "Able to Do the Impossible," fn 43.

[14] Spencer, "Able to Do the Impossible," 490.

[15] Frederich Fitch, "A Logical Analysis of Some Value Concepts," *The Journal of Symbolic Logic* 28, 2 (1963).

demonstrates that knowledge of a Fitch proposition would require both knowing and not knowing some proposition, which is of course impossible:

(1) $\Box(K(\Phi \wedge \neg K\Phi) \rightarrow (K\Phi \wedge K\neg K\Phi))$ (Knowledge distributes over conjunction)

(2) $\Box((K\Phi \wedge K\neg K\Phi) \rightarrow (K\Phi \wedge \neg K\Phi))$ (Factivity of knowledge)

(3) $\Box(K(\Phi \wedge \neg K\Phi) \rightarrow (K\Phi \wedge \neg K\Phi))$ (1, 2, PL, K axiom)[16]

(4) $\Box\neg(K\Phi \wedge \neg K\Phi)$ (PL)

(5) $\Box\neg(K(\Phi \wedge \neg K\Phi))$ (3, 4, PL, K axiom)

Interpreting '$\Box$' as a metaphysical necessity operator, the above result shows that it is metaphysically impossible to know a Fitch proposition, provided agents satisfy the plausible distribution over conjunction condition. What Spencer points out is that to conclude that no one is able to know a Fitch proposition on the basis of this argument depends on an application of Poss-Ability. Spencer could allow that it is metaphysically impossible to know a Fitch proposition, yet resist the claim that they are unknowable in the sense of ability.

Spencer calls for a new argument to be given for the unknowability of Fitch propositions. This seems something of a double standard. No argument is given for the claim that G is able to know H ∧ L. If it came down to a clash of intuitions, I would favor the unknowability of Fitch propositions over Spencer's judgment in G-cases. Nonetheless, there is at least one independent argument for the unknowability of Fitch propositions. It relies on a principle weaker than Poss-Ability, namely:

Believe-Ability: If S is able to know Φ, then it is metaphysically possible for S to believe Φ.

This principle is highly plausible and nothing in the G-cases *prima facie* rules it out, since there may be many worlds where G falsely believes H ∧ L, and Gettier-G explicitly relies on the construction that G believes H ∧ L.

---

[16] 'PL' stands for 'propositional logic' and the K axiom of normal modal logics is as follows:

K Axiom: $\Box(\Phi \rightarrow \Psi) \rightarrow (\Box\Phi \rightarrow \Box\Psi)$

Many proofs in this paper appeal to the combination of PL and the K axiom to make inferences valid in propositional logic within the scope of the metaphysical necessity operator. To see that this is valid, suppose that $\Phi \rightarrow \Psi$ is provable in PL. Since anything provable in PL is metaphysically necessary, $\Box(\Phi \rightarrow \Psi)$ follows. Then by the K axiom, $\Box\Phi \rightarrow \Box\Psi$ is true. So if we begin with some formula $\Box\Phi$ and $\Phi \rightarrow \Psi$ is provable in PL, we may conclude $\Box\Psi$ by applications of PL and the K axiom. Alternatively, to see that this is valid one may simply reflect on the fact that the logical consequences of any metaphysically necessary truth are also metaphysically necessary.

Noah Gordon

We can show if Believe-Ability is true, then any agent S who satisfies a kind of epistemic anti-akrasia constraint is not able to know a Fitch proposition:

Anti-Akrasia: $B_S \neg K_S \Phi \rightarrow \neg B_S \Phi$.

Anti-Akrasia says that if S introspects and comes to believe that they do not know Φ, they will not believe Φ. Just as doing what's best is the aim of action, knowledge is the aim of belief. So as an akratic person Ψ's whilst believing that Ψ-ing is not the best option overall, an epistemically akratic person believes Φ whilst believing that they do not know Φ. I take Anti-Akrasia to express a plausible principle of perfect rationality.

Anti-Akrasia and Believe-Ability together imply the unknowability of Fitch propositions, since believing a Fitch proposition involves believing both Φ and that you don't know Φ. Letting '$A_G \Phi$' regiment the claim that G is able to bring it about that Φ:

(1) $\square(B_G(\Phi \wedge \neg K_G \Phi) \rightarrow (B_G \Phi \wedge B_G \neg K_G \Phi))$	(B_G distribution over conjunction)
(2) $\square \neg(B_G \Phi \wedge B_G \neg K_G \Phi)$	(Anti-Akrasia, PL, K axiom)
(3) $\square \neg(B_G(\Phi \wedge \neg K_G \Phi))$	(1, 2, PL, K axiom)
(4) $\neg \Diamond \neg \neg(B_G(\Phi \wedge \neg K_G \Phi))$	(3, Duality of $\square$)
(5) $\neg \Diamond(B_G(\Phi \wedge \neg K_G \Phi))$	(4, PL)
(6) $A_G K_G(\Phi \wedge \neg K_G \Phi) \rightarrow \Diamond B_G(\Phi \wedge \neg K_G \Phi)$	(Believe-Ability)
(7) $\neg A_G K_G(\Phi \wedge \neg K_G \Phi)$	(5, 6, PL)

The above proof shows that given Believe-Ability, no agent satisfying distribution of belief over conjunction and Anti-Akrasia is able to know a Fitch proposition. A metaphysically omniscient agent satisfies the distribution constraint, and a perfectly rational agent satisfies Anti-Akrasia. Therefore, no metaphysically omniscient and perfectly rational agent is able to know a Fitch proposition.[17] And here I extend this argument in a familiar way – if such idealized agents cannot know Fitch propositions, then we mere mortals cannot either.

These results can be extended in several dimensions. First, I note that in non-Gettier G-cases where G does not believe H $\wedge$ L, that conjunction will entail that G

---

[17] Since we rely on necessitated versions of principles such as Anti-Akrasia in these proofs, what really follows is that no necessarily metaphysically omniscient and necessarily perfectly rational agent is able to know a Fitch proposition. This does not seem to me to diminish at all the strength of the inference to our inability to know such propositions, since all we have done is idealized these agents to be even more godlike. In what follows I leave implicit the clarification made here, but it applies throughout.

does not believe H ∧ L. Therefore, by another argument relying on closure over entailment and conjunction, idealized G+ is able to know H ∧ L in non-Gettier G-cases only if G+ is able to know $(H \wedge L) \wedge \neg B_{G+}(H \wedge L)$. By a structurally identical proof as above, we can show that this violates Believe-Ability whilst relying on weaker assumptions. We trade out Anti-Akrasia for Consistent Introspection:

(1) $\boxempty(B_G(\Phi \wedge \neg B_G\Phi) \rightarrow (B_G\Phi \wedge B_G\neg B_G\Phi))$ (B_G distribution over conjunction)

(2) $\boxempty\neg(B_G\Phi \wedge B_G\neg B_G\Phi)$ (CI, PL, K axiom)

(3) $\boxempty\neg(B_G(\Phi \wedge \neg B_G\Phi))$ (1, 2, PL, K axiom)

(4) $\neg\Diamond\neg\neg(B_G(\Phi \wedge \neg B_G\Phi))$ (3, Duality of $\square$)

(5) $\neg\Diamond(B_G(\Phi \wedge \neg B_G\Phi))$ (4, PL)

(6) $A_G K_G(\Phi \wedge \neg B_G\Phi) \rightarrow \Diamond B_G(\Phi \wedge \neg B_G\Phi)$ (Believe-Ability)

(7) $\neg A_G K_G(\Phi \wedge \neg B_G\Phi)$ (5, 6, PL)

Since Consistent Introspection is guaranteed simply from having perfect access to one's epistemic states, the above shows no metaphysically omniscient agent G with perfect introspection can know a proposition of the form $\Phi \wedge \neg B_G\Phi$. But non-Gettier G-cases require this ability.

     Second, it can actually be shown that G-cases are more generally in tension with Believe-Ability. The following principle follows by metaphysical omniscience:

Belief Follows Entailment: $\boxempty(\Phi \rightarrow \Psi) \rightarrow \boxempty(B_G\Phi \rightarrow B_G\Psi)$.

This fact may be exploited to show that the conjunction of Believe-Ability and Anti-Akrasia rules out all G-cases, and the conjunction of Believe-Ability and Consistent Introspection rules out all non-Gettier G-cases.

     As we have already noted, in all G-cases, H ∧ L entails that $\neg K_G(H \wedge L)$. Given this, we show that no agent G satisfying Belief Follows Entailment and Anti-Akrasia can know H ∧ L, since G believing H ∧ L would also require G believing that G does not know H ∧ L:

(1) $\boxempty(H \wedge L \rightarrow \neg K_G(H \wedge L))$ (H ∧ L entailment)

(2) $\boxempty(B_G(H \wedge L) \rightarrow B_G\neg K_G(H \wedge L))$ (1, Belief Follows Entailment)

(3) $\boxempty(B_G\neg K_G(H \wedge L) \rightarrow \neg B_G(H \wedge L))$ (Anti-Akrasia)

(4) $\boxempty(B_G(H \wedge L) \rightarrow \neg B_G(H \wedge L))$ (2, 3, PL, K axiom)

(5) $\boxempty(\neg(B_G(H \wedge L) \wedge \neg\neg B_G(H \wedge L)))$ (4, PL, K axiom)

(6) $\boxempty(\neg B_G(H \wedge L))$ (5, PL, K axiom)

(7) $\neg\Diamond\neg\neg B_G(H \wedge L)$ (6, Duality of $\boxempty$)

(8) $\neg\Diamond B_G(H \wedge L)$ (7, PL)

(9) $A_G K_G(H \wedge L) \rightarrow \Diamond B_G(H \wedge L)$ (Believe-Ability)

(10) $\neg A_G K_G(H \wedge L)$ (8, 9, PL)

Plausibly, a metaphysically omniscient, perfectly rational agent would satisfy the principles relied on above. Therefore no such agent is able to know $H \wedge L$. Again, I conclude from this that no lesser agent can either.

Finally, in any non-Gettier G-case $H \wedge L$ entails that $\neg B_G(H \wedge L)$. This means that the above argument can be run against any non-Gettier G-case again trading Anti-Akrasia for Consistent Introspection, since G believing $H \wedge L$ would then require G believing that G does not believe $H \wedge L$ (the proofs are identical in steps 4-10):

(1) $\square(H \wedge L \rightarrow \neg B_G(H \wedge L))$ (H ∧ L entailment)

(2) $\square(B_G(H \wedge L) \rightarrow B_G\neg B_G(H \wedge L))$ (1, Belief Follows Entailment)

(3) $\square(B_G\neg B_G(H \wedge L) \rightarrow \neg B_G(H \wedge L))$ (Consistent Introspection)

(4) $\square(B_G(H \wedge L) \rightarrow \neg B_G(H \wedge L))$ (2, 3, PL, K axiom)

(5) $\square(\neg(B_G(H \wedge L) \wedge \neg\neg B_G(H \wedge L)))$ (4, PL, K axiom)

(6) $\square(\neg B_G(H \wedge L))$ (5, PL, K axiom)

(7) $\neg\Diamond\neg\neg B_G(H \wedge L)$ (6, Duality of $\square$)

(8) $\neg\Diamond B_G(H \wedge L)$ (7, PL)

(9) $A_G K_G(H \wedge L) \rightarrow \Diamond B_G(H \wedge L)$ (Believe-Ability)

(10) $\neg A_G K_G(H \wedge L)$ (8, 9, PL)

I reiterate that an agent with total introspective access satisfies Consistent Introspection. So no metaphysically omniscient, perfectly introspective agent can know $H \wedge L$ in any non-Gettier G-case.

It is worth noting that even if you disagree with some of my assessments about what principles certain idealized agents satisfy, the broader point is that these results show that if Believe-Ability holds, then any agent who did satisfy these constraints could not know a Fitch proposition or $H \wedge L$. And it seems dubious to hold that someone is able to know $H \wedge L$ only by failing to satisfy some of these constraints.

To summarize, I argued that agents satisfying certain closure principles on knowability are able to know $H \wedge L$ only if they are able to know a Fitch proposition. Contra Spencer, if one thinks, as seems intuitive, that Fitch propositions are unknowable, one should conclude not only that such agents are unable to know $H \wedge L$, but that even lesser agents who do not satisfy these conditions are unable to

know H ∧ L, since one surely cannot know a proposition simply because of ignorance of its consequences. I then gave an independent argument for the unknowability of Fitch propositions based on Believe-Ability, a principle even weaker than Poss-Ability. Finally, I pointed out that Believe-Ability rules out the G-cases even independently of the argument about Fitch propositions. Therefore, both the unknowability of Fitch propositions and Believe-Ability are good reasons to reject Spencer's counterexamples to Poss-Ability.[18]

NOTES ON THE CONTRIBUTORS

Simon Dierig is a Lecturer of Philosophy and a "Privatdozent" at Heidelberg University. He received his PhD from the LMU Munich and his "Habilitation" from the University of Heidelberg. He works mainly in epistemology, the philosophy of mind, action theory and the history of philosophy. He has published on Descartes' philosophy of mind, Berkeley's metaphysics, Moore's paradox and self-knowledge, the compatibility of semantic externalism and a priori self-knowledge, the individuation of actions and the role of the principle of charity in interpretation. His recent papers include "Moores Beweis, Wahrnehmung und Skeptizismus" (*Grazer Philosophische Studien* 94, 2017), "Descartes' Sum-Res-Cogitans-Argument in der Zweiten Meditation" (forthcoming in *Archiv für Geschichte der Philosophie*) and "Falsifikationismus ohne Festsetzungen" (forthcoming in *Zeitschrift für philosophische Forschung*). Contact: simon_dierig@yahoo.de.

Claire Field is a Postdoctoral Research Fellow at the University of Stirling on the AHRC-funded project Varieties of Risk, and an Associate Research Fellow on the European Research Council-funded project Roots of Responsibility. Her research interests concern irrationality, blame, and recklessness in ethics and epistemology. She recently completed her PhD at the University of St Andrews. Website: claire-field.com. Contact: claire.a.field@gmail.com.

Eric Gilbertson is a Senior Lecturer in Philosophy at Texas State University. His current research interests include bioethics, philosophy of sport, and the epistemology of theistic belief. Contact: eng18@txstate.edu.

Noah Gordon is currently studying for the PhD at the University of Southern California. His main interests lie in metaphysics, philosophy of action, and philosophy of religion. He also has published other recent work in epistemology: "What Does It Take to Know that You Know?" (*Acta Analytica*, forthcoming). Contact: noahgord@usc.edu.

Ron Wilburn received his B.A. in philosophy from the University of California, Los Angeles, and his Ph.D. in philosophy from the University of Pennsylvania. Now retired from the University of Nevada, Las Vegas, he has published numerous papers, mainly concerning epistemological and realist issues, with an emphasis on skepticism. Contact: wilburn.ron@gmail.com.

Benjamin Winokur received his PhD in philosophy from York University (Toronto) in 2021. His research is primarily focused on questions relating to self knowledge: whether, how, and how often we can know our own current mental states. He is also interested in what roles self-knowledge plays in our psychological economies, and in what entitles other people to trust what we say about our current states of mind. Finally, he has wider interests in social epistemology, digital epistemology, philosophy of mind, metaethics, and philosophy of language. Contact: ben.i.winokur@gmail.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 knowledge in 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 Cercetări Economice şi Sociale „Gh. Zane" Academia Română, Filiala Iaşi, 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. Copyright & Publishing Rights

The journal holds copyright and publishing rights under the terms listed by the CC BY-NC License ( ). Authors have the right to use, reuse and build upon their papers 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.

No manuscripts will be returned to their authors. The journal does not pay royalties.

14. 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.