SURPRISING SUSPENSIONS: THE EPISTEMIC VALUE OF BEING IGNORANT

By

CHRISTOPHER WILLARD-KYLE

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfilment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Philosophy

Written under the direction of

Ernest Sosa

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

January 2021

ABSTRACT OF THE DISSERTATION

Surprising Suspensions: The Epistemic Value of Being Ignorant

by CHRISTOPHER WILLARD-KYLE

Dissertation Director:

Ernest Sosa

Knowledge is good, ignorance is bad. So it seems, anyway. But in this dissertation, I argue that some ignorance is epistemically valuable. Sometimes, we should suspend judgment even though by believing we would achieve knowledge. In this apology for ignorance (ignorance, that is, of a certain kind), I defend the following four theses:

1) Sometimes, we should continue inquiry in ignorance, even though we are in a position to know the answer, in order to achieve more than mere knowledge (e.g. understanding) while minimizing the effects of confirmation bias.

2) It's false that we should believe every proposition such that we are guaranteed to be right about it (and even such that we are guaranteed to know it) if we believe it.

3) Being in a position to know is the norm of assertion: importantly, this does not require belief or (thereby) knowledge, and so proper assertion can survive speaker-ignorance.

4) It can be permissible and conversationally useful to tell audiences things that it is logically impossible for them to come to know: Proper assertion can survive (necessary) audience-side ignorance.

Cumulatively, this project suggests that, properly understood, ignorance has an important role to play in the good epistemic life.

DEDICATION

For Danielle, whom I love

ACKNOWLEDGEMENTS

Single authorship is a lie. At best, it's a fiction. We cannot help but write in a way that is shaped by the voices around us.

This dissertation, too, has been a labor of many. Most immediately, I thank my committee—Jennifer Lackey, Matthew McGrath, Susanna Schellenberg, and Ernest Sosa. This dissertation has been markedly impacted by each of them. I wanted them to be on my committee not only because they would help me improve my philosophical writing (which, of course, they did) but because they are generous with their time, kind to their students, and exemplars within the profession. Each is the sort of philosopher I would like to be. I'm delighted that each agreed to be on my committee, and I count it as the luckiest thing that happened to me at Rutgers that the department hired Matt just in time to join it.

I'm especially grateful to Ernie, my principal advisor. I first read Ernie's work in a serious way in my final year of undergraduate study. His chapter in *Apt Belief and Reflective Knowledge* arguing against dream skepticism lurched me out of ambivalent defeatism about the prospects of warding off skepticism, and his two-volume set, in conjunction with a steady dose of Chisholm ingested that same semester, converted me to epistemology. My chapter "Against the Doctrine of Infallibility" engages with some of the very same material on dream skepticism of his that I first read at twenty-one. It has been a professional dream come true to work with Ernie, whose expansive curiosity and indefatigable graciousness have created conversations as enjoyable as they are thought-provoking.

I've learned from the many faculty, postdocs, and graduate students with whom I've overlapped at Rutgers and at neighboring universities, especially those who were a part of the various reading groups or seminars in epistemology. Those include (but are surely not limited to) D Black, Laura Callahan, Liz Camp, Sam Carter, Tez Clark, Chris Copan, Charles Côté-Bouchard, Morten Dahlback, Andy Egan, Megan Feeney, Will Fleisher, Danny Forman, Chris Frugé, Adam Gibbons, Thony Gillies, Michael Glanzberg, Verónica Gómez Sánchez, Chris Hauser, Caley Howland, Matt Jope, Tom Kelly, Savannah Kinkaid, Nico Kirk-Giannini, Zach Kofi, Ting-An Lin, Andrew Moon, Morgan Moyer, Dee Payton, John Phillips, Paul Pietrotski, Pamela Robinson, Ezra Rubenstein, Daniel Rubio, Philip Swenson, Jeff Tolly, Peter van Elswyk, Caroline von Klemperer, Isaac Wilhelm, and Elise Woodard.

I must make special mention of Dean Zimmerman, who first made me feel welcome at Rutgers and gave me a way to be a part of the community even before I had been officially hired as a graduate student. And also of Carolina Flores, who, in addition to being a good friend, read multiple drafts of every chapter in this dissertation, always with a constructive eye within the spirit of the project toward how the argument could be clearer, more compelling, more impactful. I consider her a peer committee member of this dissertation.

Before Rutgers, David Vander Laan, my first philosophy professor, gave me the gift of philosophical curiosity and the ambition to try to answer confounding questions. Brian Leftow advised my M.Phil. and modeled the humility to ask questions readily and the philosophical patience to work through an argument slowly. And before I knew what

philosophy was, my parents, Cindy and Doug, encouraged me to enjoy questions, both in the asking and in the answering.

The chapter "Being in a Position to Know is the Norm of Assertion" is used with permission from *Pacific Philosophical Quarterly*.[1] Penstock Coffee Roasters and Hidden Grounds jointly provided the caffeine required for sustained writing.

To my dog, Pemberley, I'm grateful for acceptance without revisions. To Rowan, for joy. To Danielle, for everything.

Finally, after having puzzled over them from afar, I am delighted to have a preface paradox of my very own: Over the course of the following dissertation, I have undoubtedly said something false. Those are exactly the parts of the dissertation that are truly single-authored.

---

[1] Willard-Kyle, C. (2020). "Being in a Position to Know is the Norm of Assertion," *Pacific Philosophical Quarterly* 101 (2):328–352.

TABLE OF CONTENTS

INTRODUCTION

1. Theses

Sometimes you're better off not knowing—even from a strictly epistemic point of view. This project articulates a vision of the good (epistemic) life in which agents can rightly choose ignorance over knowledge on purely epistemic grounds. This is not to say that ignorance is intrinsically valuable or knowledge worthless—that would be far too strong. But it suggests a picture wherein agents do not merely aim to know but to know *the right amount*. The honest know when to lie. The brave know when to run away. The epistemically virtuous know when to remain ignorant.

Ignorance can be valuable, I contend, for two main reasons. First, proper ignorance can create space for inquiry. We inquire in order to figure things out. So, it's tempting to think that once we are capable of having knowledge, we should always end inquiry there. But I argue that in epistemology, as in other domains, delayed gratification is a powerful strategy: it's often better to continue inquiry rather than to settle for immediate knowledge. For instance, a detective might reasonably wait to form a judgment about the innocence of a suspect until all the evidence has been considered, even if a subset of the evidence is sufficient to form a knowledgeable answer now. By continuing inquiry, one can often come to know with extra justification or to understand or to be rationally certain in the future, whereas closing inquiry can, by activating certain kinds of confirmation bias, make these better epistemic outcomes less likely.

Second, ignorance can be valuable because knowing certain propositions is epistemically harmful. Believing propositions like, "I have irrational beliefs about *p*," or

"My beliefs about *p* are problematic" can function as self-fulfilling prophecies. In fact, I argue that in certain specified contexts, forming these kinds of beliefs can guarantee that one will be right about them but at the cost of downgrading one's overall epistemic position. It's surprising that agents should thus forego guaranteed truths (and guaranteed knowledge), but ignorance in these cases can keep us safe from epistemic self-harm.

The project concludes by examining some of the social aspects of ignorance, arguing that assertion can be productive even when either the speaker or the audience is stuck in ignorance. I defend a knowledge-friendly norm (the position-to-know norm) that, nevertheless, captures the audience-serving insights manifest in Lackey's (2007) selfless assertions, in which the speaker does not believe and so does not know the asserted content. I then explore second-person Moorean assertions (e.g. "*p*, but you don't believe that *p*"), which can be properly asserted despite necessarily failing to give addressees knowledge of their content. I use second-person variants of Moorean paradoxes as a tool to explore the reasons we might have for telling audiences things they can't come to know. By doing so, we learn not to confuse the *aims* of assertion with its *norms*.

2. Assumptions

Those are the theses that I will defend in the following chapters. But before diving into the arguments, I want to lay out some of the guiding assumptions that recur throughout the project. These aren't intended to be anything like *axioms* or *self-evident principles* in the grand old style of the rationalists: they are neither obvious to all nor fully precise. Nor do they entail the theses. But these assumptions do express some of the core, general commitments upon which this project is built. I won't fully defend them—they are *assumptions*, after all!—but I will briefly characterize them.

A. *Knowledge is epistemically special*. "[I]t makes me wonder, Socrates, ...why knowledge is prized far more highly than right opinion?" (Plato 2002, 89). Meno asks a difficult question but a sensible one: the presupposition—that knowledge *is* more valuable than right opinion—is importantly right. And it isn't just that knowledge is more valuable than true belief. Knowledge is a special kind of epistemic success that parallels the kinds of success we care about in other domains (cf. Sosa 2011). And it plays a central normative role in epistemology. If you know that you can't be in a position to *know* something, that's a good reason not to *believe* it.

Plausibly, the post-Gettier landscape of epistemology has been *overly* focused on knowledge. But to the extent that this is so, it is because the wrong things have been ignored, not because post-Gettier epistemology has failed to latch onto something central to the field.

B. *Epistemic pluralism about value.* Although knowledge is special, it's not the only important thing epistemically. True belief matters. Rationality matters. Epistemic justice matters. Understanding matters. Any compelling theory of epistemology must balance the idea that knowledge is epistemically exceptional with the idea that knowledge is not unique in being epistemically valuable.

C. *Knowledge entails belief.* You can't know what you don't even believe. Note that it's natural to resist knowledge ascriptions by declaiming belief:

—C'mon, you know (deep down) that the Dodgers are the better baseball team this year.

—No: I really don't believe that.

When you believe, you wager your doxastic commitment. Nothing wagered, nothing gained.

D. *Belief and assertion are more alike than not*. Both take propositions as their objects. Both are species of commitment, either doxastic or assertoric. Both are only fully proper when based on the right sort of epistemic standing. Both are responsive to questions—whether questions on an individual's research agenda or the QUDs operative in a conversational context. As Williamson puts it, "[O]ccurently believing *p* stands to asserting *p* as the inner stands to the outer" (Williamson 2000, 255–56).

Some philosophers attempt to explain either belief or assertion (at least partially) in terms of the other. So, some think that (sincere) assertion *expresses* belief, whereas Dummett suggests, in contrast, that "Judgment ...is the *interiorization* of the external act of assertion" (Dummett 1973, 362, emphasis mine). I take the question of how belief and assertion relate to be a matter for careful theorization. I take the claim *that* there are deep parallels between the two as a datum.

E. *We have (limited but substantial) agency over our beliefs.* Of course, we don't have anything like total volitional control over our beliefs. We are subject to implicit bias (for instance) whether we like it or not, and this can affect our beliefs before we even have a chance to consciously reflect on them. Nor can we simply choose our beliefs at the conscious level: I can't believe just by trying that there are an even number of planets for any sum of money, no matter how large.

But the degree to which this is different from action is sometimes exaggerated. Our actions are subject to subconscious influences too, and we can't just do any old

action for arbitrarily large sums of money, either: I can't *fly* for money, nor could I maim someone I love.

Gregory Kavka ends *The Toxin Puzzle* with a comparison between belief and intention:

> One cannot intend whatever one wants to intend any more than one can believe whatever one wants to believe. As our beliefs are constrained by our evidence, so our intentions are constrained by our reasons for action (Kavka 1983, 36).

Kavka is arguing that our intentions are just as constrained as our beliefs. But it is just as useful to look at things the other way around: the mere fact that our beliefs are constrained by our (evidential) reasons does not force us to give up on the idea that our beliefs are, in a limited yet substantial way, under our control. For our intentions are certainly as under our agential control as anything is, and they share the same feature of being constrained by our reasons.[1]

We can't believe by sheer force of will what we (take ourselves to) have no reason to believe; and we can't intend by sheer force of will what we (take ourselves to) have no reason to do. Our ability to exercise agency through volition,[2] in intention and belief, is activated when our reasons (at least as experienced by us) do not force our hand. Kavka says that intentions are only "partly volitional" (1983, 36); the same, I suggest, is true of belief. That isn't as much control over our beliefs as some have hoped for, but it is enough.

3. Themes

---

[1] I set aside free will skepticism, although I don't presume to say what it means that we have free will in some sense.

[2] I don't insist that this is the only way we can exercise our agency.

Finally, I want to briefly highlight a few of the emerging themes that tie the project together.

A. *The position to know is an important epistemic concept.* In the picture I develop, the position to know is not a mere waystation on the road to knowledge: It isn't just assumed that if one is in a position to know that *p*, one would do best to actually know that *p*. Rather, the position to know represents an interesting point of decision in the epistemic life: from it, one can choose to know or one can choose to continue inquiry in the hopes of more than mere knowledge, to commit asertorically to a proposition or to remain silent. Which path an agent chooses tells us a lot about their epistemic ambitions.

B. *The study of inquiry belongs to epistemology proper*. Inquiry is just as central to epistemology as knowledge is. *How* we inquire and *what* questions we inquire into are properly epistemic considerations. These factors may not affect gnoseology, the study of what we know.[3] But they affect the broader question of what we should believe.

C. *Valuable ignorance reveals epistemic conflict.* In many ways, the central tension of this project is between the idea that knowledge is epistemically special and that knowledge is not the only thing of ultimate epistemic value. Ignorance means (at least temporarily) giving up on knowledge. But that doesn't mean stopping progress toward other epistemic goals. Thinking about valuable ignorance is a way to explore those cases in which one relinquishes knowledge in the pursuit of some other epistemic good. Like fault lines, the presence of valuable ignorance indicates tension beneath the surface. This project is thus an exercise in epistemological seismology—recording tensions that reveal the shape of conflict between competing epistemic values.

---

[3] I borrow the term "gnoseology" from Sosa (forthcoming).

There is only one good, knowledge, and only one evil, ignorance.
—Socrates

VALUABLE IGNORANCE: DELAYED EPISTEMIC GRATIFICATION

Abstract: Sometimes, it is epistemically better to ignorantly inquire into Q than it is to know the answer to Q—even when knowledge is readily available. The argument for this claim proceeds in two steps. First, I argue that it's sometimes better to continue inquiry into a question than to knowledgeably close it: doing so enables us to remain sensitive to future evidence bearing on the question. Then I argue that (usually) when it is better to continue than to close inquiry, it is also better to inquire out of genuine ignorance: it is epistemically incoherent to ask and answer a question at the same time. A guiding insight of the paper is that in epistemology, no less than in other domains, delayed gratification is often the best strategy. One shouldn't always settle for immediate knowledge: one can continue inquiry in hopes of even greater epistemic accomplishments (like knowing better, understanding, or being certain).

1. Introduction

I spent much of the morning searching for a runcible spoon. Searching was a physically active process: my eyes swiveled; my hands rummaged. In the end, I found what I was searching for: I then *had* the runcible spoon. I stopped searching.

While my eyes were swiveling and my hands rummaging, my mind was doing something too: it was inquiring. I was inquiring into the question "Where is my runcible spoon?" I *wondered* where it was, I was *curious* about its location, and I was *sensitive* to new information about where it might be.[1] In the end, I came to know the answer to the question I was looking for. It was *there*, wrapped in a five-pound note. I stopped inquiring.

Why did I stop searching for the runcible spoon once I had found it? It'd be awfully silly to search for something that I already had in my (secure enough) possession.

---

[1] Cf. Friedman (2017)

There is a norm, it seems, governing searching and having: Don't search for things you already have.

Once I realize that I have violated this norm—once I discover I've been searching and having the same thing at the same time—I could return to compliance with the norm in one of two ways: I could stop searching for the thing or I could stop having it in my possession. Here's a funny fact: even though the norm just says, "don't do both," I always resolve this tension in the same way. I always choose to stop searching and never choose to stop having. I never chuck my keys into the distance so that my having them does not normatively conflict with my searching for them.[2] Thus the adage that you always find things in the last place you think to look.

Why the asymmetry? Because aims are asymmetrical. Because one searches for things *in order* to have them. Once I have something, there isn't a reason to search for it anymore. If I've got a choice between searching and having, I choose having.

Plausibly, the same sorts of things that we've said about the physical activity of searching apply to the mental activity of inquiring. Just as there is a norm not to search and have things at the same time, there's also this (epistemic) norm: don't inquire into questions you already know the answer to.[3] We occasionally stumble into doing this, as when we inquire about a friend's plans for the day only to recall that we'd asked before and easily would have remembered if we had bothered to reflect. But it's epistemically embarrassing to have done so.[4]

---

[2] A few cases don't fit this pattern so clearly. When I hide a toy for my dog, she brings it back to me to hide it for her again. She's more interested in the game of searching for the toy than having the toy itself. We set these aside for now.

[3] I defend a version of this norm, building on work in Friedman (2017), in §4.2.

[4] But for an argument that the relevant kind of embarrassment shouldn't be construed as norm violation, see Archer (2018, 601–604).

And once again, it seems that there's an important asymmetry. One inquires into

questions (often, at least) *in order* to know their answers.[5,6] Once I know something,

there isn't a reason to inquire into questions that it is the answer to anymore. If I've got a

choice between inquiring into a question and knowing the answer, I want to choose

knowledge. Here's a principle that captures this normative preference:

> **Better to Know:** Relative to the aim(s) of inquiry on a single, given question, it's epistemically better that the agent knows the answer than that they do not.

It's better to know the answer to a question than to be ignorant of it. At least with respect

to the aims of that very inquiry.

Of course, it's not flat-footedly true that *everything* is better to know all things

considered. There's a lot of reasons not to know. Some things are secrets, and a few of

those things should be. Some things are trivial clutter,[7] unworthy of spending our limited

epistemic resources on. Some things are morally bad to know, like how to build an

atomic bomb or what it's like to skin a cat. Some things might even be bad to know for

epistemic reasons: knowing that a given experiment is statistically unlikely to produce

significant, repeatable results might stop a researcher from achieving the wealth of

knowledge that would in fact result from running it. And sometimes it's just better not to

pry.

---

[5] Perhaps we actually want something stronger: to know answers *as* answers to their corresponding questions. For instance, suppose I am wondering whether Batman is at the party but I know that Bruce Wayne is at the party. (For the uninitiated, Batman=Bruce Wayne). In some sense, <Bruce Wayne is the party> *is* the answer to <Is Batman at the party?>, but if someone doesn't know that Bruce Wayne is Batman, if they are not in a position to recognize <Bruce Wayne is at the party> *as* the answer to <Is Batman at the party?>, then one doesn't seem to be doing anything wrong by inquiring into Q and knowing the answer to Q at the same time. Maybe it's just a blameless violation. Or maybe the norm is really this: Don't inquire into Q and know the answer to Q *as* the answer to Q simultaneously.
Having raised this distinction, I'm going to proceed to ignore it for the rest of the paper. From here on out, read "knowing the answer to Q" as "knowing the answer to Q" or as "knowing the answer to Q *as* the answer to Q" in accordance with your favored interpretation. See Schaffer (2007) for related discussion.
[6] For a defense of the claim that the aim of inquiry is knowledge, see Kelp (2014).
[7] See Friedman (2018) and Harman (1986).

Notice that **Better to Know** is carefully crafted to avoid such cases. It's compatible with there being moral, practical, and prudential reasons to prefer ignorance to knowledge. And because the evaluation is relativized to assessing how an agent has done with respect to a single question, epistemic trade-offs between knowing the answers to different questions won't arise.[8] Indeed, **Better to Know** probably strikes us not just as true but as obviously so. Isn't it obvious that, in suitably tailored epistemic contexts, knowledge is better than ignorance?

But **Better to Know** is false. Or so I'll argue. It is sometimes epistemically better, from the vantage point of inquiry into *Q*, to be ignorant about *Q* than it is to know *Q*.[9] In particular, it's often better for agents to be inquiringly ignorant than knowledgeable.

If we are going to find cases that belie **Better to Know**, we had better be looking at cases in which knowledge is a genuine option for the agent. No one will be impressed by the idea that we are better off ignorant when we can't help it. So, to focus our dialectical attention, let's consider only cases in which agents are in a position to know the answer to a question and, if this is something more,[10] *would* come to know the answer if they formed a belief about the answer to that question.[11] This guarantees that knowledge is a *fully available* option for an agent. Inquiring ignorance in such cases is dearly bought, for the (opportunity) cost is knowledge.

---

[8] We thus avoid the concerns about epistemic trade-offs that Firth (1981) and Berker (2013a, 2013b) raise for epistemic consequentialisms.

[9] I use the expression "knowing *Q*" as shorthand for "knowing (the proposition that is) the (complete enough) answer to *Q*."

[10] On the account of the position to know that I develop in Willard-Kyle (2020), it is something more.

[11] This is what Sosa calls a *strong* position to know (forthcoming: chapter 6, §3)

An agent in this fortunate epistemic situation could then proceed in any of the following four ways, depending on whether they *believe* or not and whether they *inquire* or not:

A	B
Knowing the Answer to Q	Knowing the Answer to Q
Continuing Inquiry into Q	Not Continuing Inquiry into Q
C	**D**
Not Knowing the Answer to Q	Not Knowing the Answer to Q
Continuing Inquiry into Q	Not Continuing Inquiry into Q

The agent *believes* and so, given our dialectical assumptions, *knows* the answer to the relevant question in the top row. In the bottom row, the agent does not believe an answer to the question and, since knowledge entails belief,[12] they don't know it either. The agent continues inquiry in the left column but not in the right. The final alternative, D, is to *abandon* inquiry into a question, neither committing to resolving the inquiry (through belief in an answer) nor pursuing it. For the most part, we will be professionally uninterested in this alternative. It might be epistemically permissible to abandon inquiry much of the time.[13] But we'd still like to know what we ought to do epistemically given that we *are* invested in a question.

My task is to show that agents should sometimes land in box C rather than boxes A or B. If that's right, then sometimes it is epistemically better to ignorantly inquire into

---

[12] I assume the orthodox view that knowledge entails belief, but see Black (1971) for a critique of the orthodoxy. If knowledge doesn't entail belief, then agents might count as knowing in the bottom row. I will ultimately argue that sometimes agents ought to end up in box C. If agents know the answer to the question in box C (at least in the cases I consider) then my argument won't establish the conclusion that it is sometimes better for agents to be ignorant than it is for them to know. But parts of the argument could be adapted to defend an interesting (albeit different) conclusion: that it is sometimes better for agents not to believe what they know.

[13] Nelson (2010) argues that we have no positive epistemic duties to believe any particular proposition in part because no proposition has (according to Nelson) a purely epistemic demand on our attention. For similar reasons, one might suppose that no *question* has a purely epistemic demand on our attention.

Q than it is to know the answer to *Q*. My argument proceeds in two steps. First, I argue that it's sometimes better to be on the left side of the diagram. Then I show that in all (or at least many) of those same cases it's better to be on the bottom half of the diagram. Result: these are cases in which we should be in box C—inquiring but not knowing.

Let's sketch the argument again, in a little more detail. First, I'll argue that sometimes, continuing inquiry into *Q* is epistemically better (with respect to that very inquiry) than knowing the answer to *Q* while also foregoing further inquiry into *Q*. Here is an example of the sort of case I have in mind:

> *The Philosophy Student*
>
> A philosophy student is studying Descartes' *cogito* and trying to make sense of what the structure of the argument is. They have read the relevant sections of the *Meditations* several times and listened closely to their instructor's interpretation of the argument. They are a good reader of the text, and their philosophy instructor did indeed present an accurate representation of Descartes' argument. If on an exam they were asked the question, "What is the structure of the *cogito*?" they would give a knowledgeable response. Nevertheless, the philosophy student is deeply interested in the way that the *cogito* works, and they remain curious about the argument's structure. They are not yet satisfied with their inquiry, and this is partly explained by their (perhaps unarticulated) interest in really *understanding* the argument in a way that enables them to explain how they know (to themselves or to someone else). Instead of committing to a belief now, they continue wondering and make a note to look up the *SEP* article on Descartes later that evening.

The fact that the philosophy student could know now shouldn't stop them (I will argue) from inquiring. After all, they are aiming for something *better* than mere knowledge.

I haven't said yet whether the philosophy student (or others like them) should hold onto knowledge while they inquire. In the second part of my argument, I answer *no*. There's something incoherent about asking and answering a question at the same time. Doing so (in the relevant way) is either psychologically impossible or epistemically impermissible and often introduces confirmation bias where it is least wanted. In the

relevant cases, it is epistemically better (with respect to that very inquiry) to inquire into
Q without knowing the answer to *Q* than it is to inquire into *Q with* knowledge of its
answer.

Returning to our diagram, the first step in the argument says it is sometimes better
to (inquiringly) be on the left-hand side. The second step says that (in the same sort of
cases), box A is defective. It's (at least usually) better to be in box C. The conclusion:
sometimes it is epistemically better (with respect to one's inquiry into *Q*) to inquire into
Q without knowing the answer to *Q* than it is to know the answer to *Q*.

A	B
Knowing the Answer to *Q* Continuing Inquiry into *Q*	Knowing the Answer to *Q* Not Continuing Inquiry into *Q*
C	**D**
Not Knowing the Answer to *Q* Continuing Inquiry into *Q*	Not Knowing the Answer to *Q* Not Continuing Inquiry into *Q*

A guiding insight of the paper is that in epistemology, no less than in other domains,
delayed gratification is often the best strategy. One shouldn't always settle for immediate
knowledge: one can continue inquiry in hopes of even greater epistemic accomplishments
(like knowing better, understanding, or being certain).

It turns out that ignorance has an instrumentally valuable role to play in our
epistemic lives: it enables inquiry. This is surprising—to call someone ignorant is to
insult them! But if this paper is right, it can be even worse to be a know-it-all.[14]

2. Characterizing Inquiry

---

[14] This is not to deny that certain kinds of ignorance are Very Bad Indeed. The epistemic, social, and moral
dangers of white ignorance, for instance, have been both well-documented and well-theorized (see, e.g.,
Mills 2007), and it's important to flag that the thesis of this paper does not undermine the dangers of such
ignorance. Rather, the focus will be on how ignorance can be epistemically valuable when it creates space
for positive kinds of inquiry.

Ignorance can be valuable, I will argue, because it enables inquiry. But what is inquiry? And what is it for an inquiry to be open or closed? In this section, I clear the terminological ground.

2.1 Wondering

To inquire involves having an inquiring state of mind that is goal-directed, question-directed, and characterized by sensitivity to new information (Friedman 2017, 307–308).[15] That isn't necessarily *all* that inquiry is: perhaps inquiry also involves *doing* something with the aim of resolving the inquiry—not just having the curiosity itch but trying to scratch it.[16] But all genuine inquiry at least involves a real, inquiring attitude. Prototypical inquiring attitudes include *wondering* or *being curious* about a question. Inquiring states are question-directed in that they take questions as objects: one wonders *whether* it will snow tonight or *why* the kids are being so quiet. One doesn't wonder *that* it will snow tonight.[17] One can also inquire into a subject matter, but these can be thought of (for our purposes at least) as sets of questions. Questions are goal-directed: to have a question is to be on a quest for an answer. This paper will explore what *kind* of answer (or epistemic access to that answer) successfully resolves the aim of inquiry in due course. This curiosity and goal-directedness tend to produce sensitivity to future information that bears on the question.

Admittedly, this isn't the only way the English word "inquiry" is used. one could use the English word "inquiry" instead to pick out answering-preparatory activities. For

---

[15] For defense of the claim that inquiry is goal-directed, see Kelp (2014).

[16] I owe the curiosity "itch" metaphor to Miriam Schoenfield.

[17] If one doesn't mind sounding a bit archaic, one can get away with saying things like "I wonder that you've succeeded in your quest!" But here "I wonder" means "I am amazed" and not "I am curious." Or if the speaker does have a question, it isn't *whether* you've succeeded in your quest but (perhaps) *how*.

instance, Ernest Sosa writes that inquiry includes the "things that one does as one attempts to put oneself in position to answer one's question knowledgeably" (Sosa forthcoming, chapter 4, §D7). Although not eschewing curiosity (the answer-preparatory activities are guided by an aim to know), this approach highlights the *active* side of inquiry. Such activities might include gathering evidence or weighing the evidence one already has, for instance.

Let us draw a distinction. There are inquiring *states* (like wondering or being curious) and inquiring *activities* (like gathering or contemplating evidence). I can inquire_STATE without inquiring_ACTIVITY by wondering where my runcible spoon is from the couch. And I can inquire_ACTIVITY without inquiring_STATE by rummaging through items on the table, as though to find the runcible spoon, disinterestedly.[18]

As I use the word, "inquiry" shall always include inquiry_STATE and not just inquiry_ACTIVITY.[19] Inquiry without curiosity isn't inquiry at all. At least, it doesn't qualify as what I mean by "inquiry" in this paper.[20] One might think that the same goes for inquiry_ACTIVITY: Inert inquiry isn't inquiry at all.[21] I'm sympathetic to this thought,

---

[18] I'm especially grateful to conversations with Ernie Sosa and Sam Carter here for help distinguishing between these meanings.

[19] This does not entail that anyone with an inquiring_STATE is thereby inquiring in the sense of this paper. Perhaps an agent is curious about a question but does not have a goal to answer the question. If such a case is possible, the relevant agent would not be an inquirer in our sense. The question would not be on the agent's "research agenda" (cf. Friedman 2017; Olsson and Westlund 2006). Thomas Kelly's (2003) case of an agent who doesn't want to know the end of a movie before having seen it (despite, perhaps, in some sense, being curious about the ending) may fall into this category, for instance.

[20] This is a bit more restrictive than Friedman who allows that some interrogative attitudes constitutive of inquiry may be processes rather than states (2019, 299). Someone could count as inquiring, for Friedman, by having the (process) IA of deliberating without having the (stative) IA of being curious. I have a hard time imagining an agent who is genuinely deliberating without at least a weak kind of curiosity or other stative IA in the background. But these differences needn't detain us. What will ultimately matter for the argument is that inquiring agents "aim to resolve their question" (299) in an epistemically appropriate way. *Having* such an aim is enough to count as a stative IA in my view.

[21] Sapir and van Elswyk (manuscript) say that "inquiry is a goal-directed *activity*" (emphases mine), and even Friedman says that inquiring is "something that we do," (2019, 297), although she also insists that "we should not straightforwardly identify inquiring with acting" (298).

although officially I will remain neutral on it.[22] If someone wants to use the word

"inquiry" differently to include agents who are inquiring$_{\text{ACTIVITY}}$ without inquiring$_{\text{STATE}}$,

then (as a matter of English) they have my blessing. But in this paper (to avoid overuse of

subscripts), "inquiry" shall always denote something that includes inquiry$_{\text{STATE}}$, and I will

refer to any inquiry$_{\text{ACTIVITY}}$ that is unaccompanied by inquiry$_{\text{STATE}}$ only as

"investigation."

     We're interested in the sort of stative inquiry that involves interrogative attitudes

because we are interested in the questions that agents ask themselves.[23] To capture this

idea, Friedman (2017, 308) follows Olsson and Westlund (2006) in employing the

metaphor of a research agenda. A research agenda is to an individual inquirer and belief

(at least something like) what a question under discussion (QUD) is to a community of

inquirers and assertion.[24] For a question to be on one's research agenda needn't mean that

one is currently paying much attention to it. Just as one can have occurrent and non-

occurrent beliefs, one can have occurrent and non-occurrent inquiries.

     In our lexicon, all genuine inquiries are open inquiries. If Q is on your research

agenda, you have an open inquiry into Q. Inquiry into Q is closed for an agent when they

believe a satisfactory answer to Q.[25] Beliefs can be weaker or stronger: the sort of belief

that closes inquiry is full belief. I won't defend a complete theory of full belief, but here

is a characterization: Full belief is the thing typically (but not only) produced by

---

[22] Nevertheless, one might think that the sorts of reasons that apply to inquiry will have (rough) parallels that apply to investigation. Inquiry into a question provides motivation to investigate it, and investigation is more likely to succeed if motivated by genuine inquiry. For more on norms on investigation or evidence-gathering, see Flores and Woodard (manuscript).
[23] Cf. Whitcomb (2017)
[24] See Roberts (2012)
[25] Cf. Friedman (2019b). I use "closing inquiry" and "settling inquiry" interchangeably, although neither of these (in my mouth) entails having closed inquiry successfully.

decisions to affirm after conscious, judgmental deliberation;[26] it is capable of constituting

knowledge when appropriately situated epistemically; it enables us to have certain

affective attitudes toward propositions (like "being happy that *p*");[27] and, it is a

counterpart to flat-out assertion. This concept of belief is perhaps a "theoretical posit"

rather than an ordinary language unpacking of the English words "belief" or "think"

(Hawthorne, Rothschild, and Spectre 2016, 1402).[28] But it's a concept worth positing:[29]

We want something that stands to our research agendas roughly as assertions stand to

QUDs.

Admittedly, we sometimes conclude an inquiry with less than full belief, as when

a gameshow contestant loses curiosity in a question after making their best guess (where

a guess is conceptualized as involving a weaker kind of commitment than full belief).[30]

But I do not think such cases should be thought of as *closing* inquiry, even when making

the guess plays a role in enabling the agent to end an inquiry. This is because it *would* be

sensible for the gameshow contestant to continue wondering whether they got the answer

---

[26] Cf. Sosa (2015)

[27] I owe this suggestion to Matt McGrath. See his (manuscript).

[28] See also Sapir and van Elswyk (manuscript) who usefully use hedged assertions as a tool to defend the permissibility of open inquiry in conjunction with (at least) weaker-than-full belief. Whitcomb (2017, 158–59) and Avery (2018, 600) make similar observations.

[29] This posit is not an invention of recent analytic philosophy. Sextus employed a similar (if not quite the same) distinction:

> When we say that sceptics [cognate to "inquirers"] do not hold beliefs, we do not take 'belief' in the sense in which some say, quite generally, that belief is acquiescing in something; for Sceptics assent to the feelings forced on them by appearances—for example, they would not say, when heated or chilled, "I *think* I am not heated (or: chilled)'. Rather, we say that they do not hold belief in the sense in which some say that belief is *assent* to some unclear object in the sciences (Sextus 2000, 6, emphases mine).

[30] I'm grateful to Matt McGrath and especially Ernie Sosa for stimulating conversation about guessing and the gameshow contestant, which I can only gesture at here. For an alternative viewpoint that sees guessing as a variety of inquiry-closing judgment rather than a kind of assent that plays no inquiry-closing function, see Sosa (forthcoming). In the background of my favored view is the thought in the spirit of Friedman (2019b) that a distinguishing feature of full belief against other doxastic attitudes is precisely its inquiry-closing function.

right after having made their guess whereas (I will argue in §4) it would not be sensible to continue inquiring into a question that one fully believes an answer to.

In any case, from here on out, we will restrict our attention to full, outright, and occurrent beliefs (and occurrent inquiries) unless otherwise noted. While occurrent, full beliefs commit the agent to an answer to a question, mere dispositions to believe do not. Dispositional beliefs (as opposed to dispositions to believe)[31] may fall somewhere in the middle, and it's unclear whether agents with dispositional beliefs in the answer to a question should take the inquiry to be resolved. Likewise, for tacit, fragmented,[32] implicit, or otherwise "offline" beliefs. I propose that, methodologically, we focus on the clear case of occurrent and outright belief, leaving open how our findings may apply to less robust or less immediate varieties.

What counts as a satisfactory answer—an answer that, when believed, effectively closes inquiry? This is a question we will return to in §4. For now, it's enough to note that all complete (that is, exhaustive) answers count as satisfactory answers, but not every satisfactory answer is complete. To see how an incomplete answer might nevertheless be satisfactory in our still-to-be-considered sense, consider questions like this:

1. Where can I buy a cup of coffee?[33]

Here are two satisfactory answers to this question:

2. You can buy coffee at Hidden Grounds.

3. You can buy coffee at Penstock Coffee Roasters.

---

[31] See Audi (1994) for the seminal articulation of this distinction.
[32] See Egan (2008)
[33] Example from Moyer and Syrett (2019, 8) and (forthcoming).

Notice that these answers aren't complete: neither exhaustively includes all the places one could buy coffee (or even those nearby). Suppose you know (2) but not (3). You still count as knowing where I can buy a cup of coffee (at least in most conversational contexts). And we would not infer from your knowing where I can buy a cup of coffee that you know I can buy coffee at Penstock: you might count as knowing just in virtue of knowing that I can buy coffee at Hidden Grounds. Question (1) prompts *mention-some* rather than complete answers in most conversational contexts (Hintikka 1976). Importantly for our purposes, believing (2) as the answer to (1) counts as closing inquiry into (1) in most contexts.[34]

Our definitions for open and closed inquiry have some important consequences. First it isn't definitionally ruled out that a question could be open and closed for an agent at the same time: if it is possible for an agent to have *Q* on their research agenda while also (fully, occurrently) believing an answer to *Q*, then inquiry into *Q* is both open and closed for them. Second, one doesn't have to have inquired into *Q* in the past for inquiry into *Q* to be closed now. If I have always believed that Santa exists, the question *whether Santa exists* is closed but has never been (merely) open for me. Third, not all inquiries that were open before but are not open now are thereby closed now. When I stopped wondering how Santa delivers all those presents in one night, I ended inquiry without closing it, for I did not replace my wondering with a belief in a complete answer to the corresponding question (the inquiry wasn't closed, it just ended). Finally, some questions get stuck (or set) open, but their being set is not another way of being closed. Agnostics

---

[34] One could insist: "No, I really meant where are *all* the places I can get a cup of coffee?" But this isn't the way the question is intended in most contexts.

have not closed inquiry on God's existence even if they are convinced that they are unlikely to ever form an opinion one way or the other.[35]

2.2 Taking Questions Up

I now want to introduce a new notion, that of *taking up* a question.[36] It's one thing to wonder about a question or to have it on one's research agenda. But sometimes one aims to do more: sometimes one tries to answer it. But usually one doesn't simply try to answer a question no matter what. I could try to answer a multiplication question by believing the first answer that comes to mind. But that's not good epistemic behavior, at least not for forming full, outright beliefs.[37] When I take up a question, I try to answer the question *to a certain epistemic standard*, even if the reasons I've picked that question to take up (rather than some other one) are (partly) moral or prudential. Our question now: What is (or should) that standard (be)? Given that I've taken a question up and that my primary aim at the moment is to do well in my inquiry into *Q*, how should I decide whether to close inquiry or continue it?

Here's a plausible answer inspired by **Better to Know**: I should close inquiry just in case I would do so knowledgeably.

---

[35] If this leads the agnostic to lose interest in the question of God's existence, then inquiry into the question of whether God exists will be neither open nor closed for them: inquiry will have expired. But the agnostic might still have an open inquiry into the question of God's existence if they still wonder whether God exists (despite their stipulated conviction that their wondering will never receive an answer).

[36] I borrow this terminology from Sosa (forthcoming) in whose mouth it has a slightly narrower meaning:
> "When you consciously inquire on whether p, you consciously take up your question with a view to alethic affirmation at a minimum, or to the more ambitious aim: to *judging* (to affirming on whether p, to do so alethically, and not only correctly but aptly) (forthcoming, chapter 4, §D5).

[37] I leave open that there may be other kinds of doxastic commitments (guesses, "hedged" or otherwise provisional beliefs, etc.) for which it is epistemically appropriate to commit oneself (in the relevant way) without being in a very good epistemic position.

> **The Knowledge Rule:** If one is taking up the question Q, then one should close inquiry into Q just in case one would thereby know the answer to Q.[38,39]

The "should" has wide scope: one can always comply with **The Knowledge Rule** by not taking a question up. It's helpful to break up the embedded biconditional in **The Knowledge Rule** into two directions:

> LTR: If one is taking up the Question Q, then if one closes inquiry into Q one should thereby know the answer to Q.

This is something like a knowledge norm on closing inquiry.[40] Since we've said that (full) belief closes inquiry, it's a knowledge norm for (full) belief. Or at least belief in the answers to an important class of questions (the questions one takes up). Although not uncontroversial, I find this norm, or something in its vicinity, quite plausible.[41] Nothing that I say here will challenge the left-to-right direction. Now let's look at the other direction:

> RTL: If one is taking up the Question Q, then: one should close inquiry into Q if one would thereby know the answer to Q.

---

[38] Although I'm not sure anyone has defended precisely this knowledge- and question-centric norm, the literature is replete with defences of biconditional norms in the neighborhood, sometimes with other epistemic or semantic properties (e.g. justification or truth) playing a central role, or with the notion of *considering a proposition* instead of *taking up a question*. Some prominent examples are listed below:
Chisholm: "We may assume that every person is subject to a purely intellectual requirement—that of trying his best to bring it about that for every proposition h that he considers, he accepts h if and only if h is true" (1977, 14).
Feldman: "For any proposition p, time t, and person S, if S considers p at t, then S has a duty to have the attitude toward p that fits the evidence S has at t concerning p" (2002, 368).
Sosa: "When one faces *judgmentally* a question whether *p*," one aims "to affirm alethically re <p?> iff one's alethic affirmation would be apt," where such apt affirmation entails knowledge (forthcoming, chapter 3, §A2).
[39] A difficulty for **The Knowledge Rule** that I shall (mostly) pass over is what to say about questions that do not have a uniquely correct and satisfactory answer (as for mention-some questions). What is *the* answer in such contexts? I do, however, make some suggestions bearing on this topic in §4.2.
[40] For a defense of this view, see Kelp (2014).
[41] See Willard-Kyle (2020).

This says once you can have knowledge, you ought to take it. Once you've taken up a question, you aim to come to close inquiry if doing so meets a certain standard. The RTL direction says that knowledge is the standard one should hold inquiry to.

Let's press pause. Some epistemologists will think that I have unfairly privileged knowledge in the discussion so far. Just as there are competing epistemic norms on proper assertion, so there are competing epistemic norms on proper belief. Some will think knowledge is an important threshold, but others would prefer to focus on variants like **The True Belief Rule**, **The Reasonable Belief Rule**, and so on.[42] Expositionally, it's useful to focus on just one such rule, and, as author, I call dibs on picking which one. But I hope that, so long as one's own favored rule does not require absolute certainty, it will be clear enough how what I say about **The Knowledge Rule** applies to variants of the rule. There will be some important (but not maximally good) threshold at which one could successfully close inquiry and a corresponding question about whether we should always close those inquiries that have been taken up when we have the opportunity to do so successfully.[43]

---

[42] As flagged in an earlier footnote, Chisholm (1977) suggests a principle in the neighborhood of **The Truth Rule**. Kvanvig argues that truth is the goal of inquiry, although he at least grants for argument that, in some sense, "successful inquiry yields knowledge" (2003, 9). Feldman defends an evidentialist version of the rule, explicitly noting that Clifford's (1877) evidentialist rule is introduced under the subtitle "The Duty of Inquiry" (Feldman 2002, 363).
Plausibly, there is a tight link between inquiry-closing belief and QUD-closing assertion. Sosa, for instance, suggests that "If knowledge is the norm of assertion, it is plausibly also the norm of affirmation, whether the affirming be private or public," (Sosa 2011, 48) and Williamson says that belief stands to assertion "as inner stands to outer" (Williamson 2000, 255–56). (See Willard-Kyle (2020, 344–46) for recent discussion of these analogies.) If that's right, then it is easy to imagine the literature on norms of assertion mirroring potential answers to the question of what properly closes inquiry: truth for Weiner (2005), reasonability for Lackey (2007), having supportive reasons for McKinnon (2013) and so on. (Although it's noteworthy that Kvanvig (2011) proposes a different norm for assertion—justification—than he does the goal of inquiry—truth. So, we should not assume that all authors will grant the proposed parallelism.)
[43] I'm grateful to Zach Kofi for pressing for greater clarity on this point.

Press play. This principle—that, conditional upon taking up a question, one should close inquiry just in case one would thereby come to know—is attractively simple and cleanly partitions those cases in which we should close inquiry from those we should not. As is often the case, however, the simple answer runs into difficulty upon closer examination. And it isn't just that **The Knowledge Rule** is false in some nitpicky way: it's false in a way that gives us insight into the way that ignorance can play an epistemically useful role in our lives.

3. On Behalf of Continuing Inquiry

In the remainder of this paper, I offer an argument against **The Knowledge Rule** (and, by extension, **Better to Know**). Recall our space of options for agents who would know the answer to a question they are taking up upon closing inquiry with a full belief:

A	B
Knowing the Answer to Q Continuing Inquiry into Q	Knowing the Answer to Q Not Continuing Inquiry into Q
C	**D**
Not Knowing the Answer to Q Continuing Inquiry into Q	Not Knowing the Answer to Q Not Continuing Inquiry into Q

In this section, I'll argue that when we don't want to simply abandon inquiry (option D) it is sometimes epistemically better to land in the left-hand side of this figure. Sometimes, continuing inquiry into Q is epistemically better (with respect to that very inquiry) than knowing the answer to Q while also foregoing further inquiry into Q.

In §4, I'll go on to argue that when we are genuinely inquiring, we should also forego knowledge (putting us in box C). But in this section, I'm only arguing that it's often better to continue inquiry, knowledgeably or ignorantly, than not to, even when one would *knowledgeably* refrain from continuing inquiry.

3.1 The Dual-Threshold Model for Success in Closing Inquiry

Before getting to the cases, let's address a question: What does it take for inquiry to be closed successfully? I contend that this question is ambiguous between two readings and that **The Knowledge Rule** ultimately goes wrong by only being sensitive to one of them.

Here is one thing it might mean for inquiry to be closed successfully: that the agent has met their own (epistemic) goals for the inquiry. Agents might want to achieve knowledge, but that doesn't have to be the thing they care most about. An agent can be driven by a desire to know but they can also be driven by a desire to understand, or to have justified beliefs, or to have certainty, or to know that they know, or merely to have the right opinion. There are lots of good epistemic things in the world, and, at first glance, it isn't obvious why it would be epistemically bad to aim for any of them.

But, in fact, we don't seem to be happy with epistemic agents *anytime* they've achieved their own individual aims. Sometimes, their epistemic sights aren't raised high enough. Someone who comes to a full belief in a true answer to a question but just by guessing has fallen short of full epistemic success. Yes, they got something epistemically valuable (right opinion) and perhaps that is the thing that they (individually) cared about, but there's some other epistemic standard that they have fallen short of, and *that* standard (whatever it is) seems to play a more prominent role in our evaluation of their belief than the agent's internal goal of (merely) having a true opinion. Despite the plurality of valuable epistemic things in the world, at least one of them seems special for belief-evaluation.

I think that the special epistemic standard is knowledge. Others think it is something else. But whatever the special epistemic standard is, we face two puzzling questions: (1) If there are many genuine epistemic values, why is just one of them marked as special for belief formation? (2) How do we make sense of the variety of potential epistemic aims for inquiry in a normatively coherent way? I will say something brief about each of them.

First, why think that one standard for belief is epistemically special? One reason is that we need a standard of *commitment release* for inquiry (Kelp 2014). Kelp asks us to suppose that a detective promises to resolve an inquiry into a central question of a criminal case. We need some agreed upon standard for when the detective has secured the right sort of epistemic access to the answer to count as having successfully completed the inquiry. A default epistemic standard for inquiry gives us a way to do that.[44]

But we don't need to think that this is a norm that only applies to inquiry-driven professionals. Whenever anyone asserts something, they are answering a question operative in the conversational context (a QUD), even if that question is implicit. And similarly, when we form full beliefs, we do so in response to (perhaps implicit) questions on our research agenda. If you assert that *p*, I need to know (at least in a rough way) what sort of epistemic standard you take yourself to have met Otherwise I won't know how to update my beliefs in response to your public affirmations. Having one *special* epistemic standard helps us to coordinate our social responsibilities for asserting and for appropriate uptake in response to testimony.[45] On the assumption that assertion closely parallels full

---

[44] Kelp argues, and I agree, that the relevant standard is knowledge.
[45] I develop this story in more detail in (Willard-Kyle 2020, 333).

belief, that's a good reason to think that belief, too, has a special epistemic standard among the many epistemic goods.

Second, how do we make sense of plural, potential epistemic aims for inquiry in a normatively coherent way? Here, a grading metaphor is instructive.[46] When one takes a class, there's a single standard that separates a passing grade from a failing one. Broadly speaking, a passing grade in a class indicates a (minimally) successful performance. But while some students do aim to merely pass, most aim for something better. Some students aim for an A, or at least a B. To do as well as they can, or better than Brad, and so on. Each of those grades represents a level of performance that is objectively valuable and objectively *more* valuable than merely passing.

Suppose you ask a student whether they performed successfully in their class. Passing the class is the minimum success condition for this kind of performance. But a student might reasonably take themselves to have fallen short of full success if they did not, in addition, get the (objectively valuable) grade that they *wanted*.

If a student aims too low, they can get the grade they want without thereby counting as succeeding. At least not in any robust sense. Imagine a student who says, "Hey, I got a 40%, which is even better than the 35% I was aiming for!" "Not good enough," we respond. Even though 40% does represent work of *some degree* of objective worth (more than 35% does, for instance), it doesn't represent enough worth to qualify as a success according to our special standard of a passing grade. So, there is a kind of asymmetry. A student fully succeeds when they get the grade that they want *only if the grade they want reaches a certain level* (at least a passing grade).

---

[46] I'm grateful to Carolina Flores for suggesting the grading metaphor.

Let's steer back toward epistemology. As there are a variety of potential academic performances in a class, so there are a variety of potential epistemological performances in response to an inquiry, many of which are objectively epistemically valuable (justified belief, knowledge, understanding, etc.). Certain epistemic performances don't get a passing grade: they don't meet the special standard of success that is relevant to belief formation. I've suggested that the relevant standard is *knowledge*, but our main concern in this section is that there *is* a minimum threshold for success not *what* it is.

But the *minimum* threshold is not the *only* threshold relevant to successful performance. It also matters what threshold the agent was aiming for. Like a student who aims not just to pass but to get an A+, it's laudable to seek the higher epistemic goods. It's good to seek not just knowledge but *knowledge+*.[47]

What is knowledge+? Knowledge+ is the collection of epistemic goods that (1) entail knowledge and (2) are better than mere knowledge. It's good to know, but it's even better to know that one knows (KK): when you know that you know, you don't just know the answer, you know that you got it right. KK entails knowledge. So, KK is one variety of knowledge+. Alternatively, knowing requires a certain degree of justification. But one can have knowledge with even more justification than knowledge requires.[48] Knowledge-with-extra-justification entails and is better than knowledge: knowledge-with-extra-

---

[47] Thanks to Matt McGrath for suggesting the umbrella term. See McGrath (manuscript) for a distinct but related way of identifying a dual-criteria for success in belief-formation.

[48] This, or something like it, is plausibly true even on a knowledge-first epistemology according to which one's evidence is one's knowledge. Suppose I know that $p \supset q$ and that p and (knowledgeably) infer q. On Williamson's (2000) picture, my evidence for q now has an evidential probability of 1. If I then also learn that $r \wedge (r \supset q)$, I seemingly acquire new evidence bearing on q but my evidential probability for q doesn't go up. (After all, it was already 1!). But this doesn't mean there isn't something better about my new epistemic position with respect to q—my knowledge that q is more robust and secure since it is supported by two separate arguments that entail it. It's just that this betterness cannot be cashed out in terms of evidential probabilities.

justification is thus another variety of knowledge+. If you understand well enough to explain how you know that *p,* then you also know that *p*. But it's better to understand well enough to explain how you know that *p* than merely to know that *p*, so that too is a kind of knowledge+. The same is also plausibly true for other states like knowing full well[49] or being rationally certain.

Since knowledge+ entails knowledge, aiming for knowledge+ meets the minimum threshold for epistemic success for closing inquiry. So, like the student who aims for an A (and unlike the student who aims for a 35%), it seems there is nothing epistemically inappropriate about agents aiming for knowledge+. We typically admire people who want not just to know but to understand, who want their (knowledgeable) beliefs to be especially justified, to know that they know, and so on.

At least, there's not anything *inherently* inappropriate about agents aiming for knowledge+. Perhaps some varieties of knowledge+ are too lofty for us to properly aspire to on certain questions. Call this the Icarus objection: on this view, seeking knowledge+, at least of a certain kind, is epistemic hubris. Suppose, for instance, that I am walking down the street and see a parked car. I am in a position to know that it is a parked car. But I *could* try to aim for rational certainty, a species of knowledge+. I could go kick the tires to make sure it isn't a cleverly designed car façade or ask passersby if they *also* see a car there to rule out that I'm hallucinating it, or work through the dream arguments of Al-Ghazali and Descartes. This sort of behavior may strike us as bizarre, epistemically fetishistic even. I might achieve a kind of knowledge+ by testing additional evidence for the car's existence, but the sort of agent who cannot resist such inquiries seems destined

---

[49] See Sosa (2011)

for a bad end. Not aiming higher at knowledge would keep us safe from this kind of epistemic tragedy. [50]

I think there's room to take a hard line here: although such behavior is strange and impractical, one could argue that there isn't anything *epistemically* wrong with it. In the end, the agent really is raising the quality of their epistemic position with respect to whether there is a car there. We even sometimes *encourage* our students to try on this doubt-inducing epistemic lens, despite the danger of skepticism—at least from the safety of the epistemology room. But the hard line is not required. Here is what is important for my argument: for at least some questions, there is a gap between the minimum success condition for closing inquiry and the highest epistemic status that an agent could reasonably aim for. Perhaps we shouldn't aim for anything like rational certainty for perceptual beliefs, for instance. But we can reasonably aim for extra-justified (i.e. more than is required for knowledge) knowledge. And we can plausibly aim much higher than that when contemplating mathematics, for instance. In §3.2, I will present several cases in which it seems legitimate for agents to aim for at least some species of knowledge+.

We now return to our original question: what does it take for inquiry to be closed successfully? We've seen that there are two things this could mean: (1) meeting the minimum success condition (knowledge), or (2) meeting one's own, individual aim with respect to an inquiry. For the second category, not just any individual aim will do. The target must be genuinely epistemically valuable (e.g., no aiming for false belief) and at least as the minimum success condition, which I've suggested is knowledge. When one's individual aim is a species of knowledge+, goals (1) and (2) are both (at least potentially)

---

[50] I'm grateful to Carolina Flores, Matt McGrath, and Ernie Sosa for conversation on this point.

legitimate and yet come apart in interesting ways. There are, then, two relevant thresholds defining success for closing inquiry.

The dual-threshold model for successful inquiry holds these two things in tension: the thesis that knowledge is special, and a commitment to epistemic value pluralism. We need a socially available, epistemic standard for when an inquiry has been answered well enough for belief and testimony. I think that standard is knowledge. But there are a million epistemically good things besides knowledge, and agents can appropriately aim for many of them (those that count as knowledge+) while still upholding the knowledge requirement. We can respect the centrality of knowledge without stripping agents of the epistemic right to aim for even better things.

Like an A+, knowledge+ signifies that one has earned more than the bare minimum required for success. This isn't to disparage those who stop at knowledge. Much of the epistemic life can be run efficiently on the pass/fail model. But thinking about knowledge+ allows us to recognize a wider range of achievements.

We've found a reason to be suspicious of **The Knowledge Rule**: it is insensitive to the agent's epistemic goals during an inquiry. It says that knowledge is the standard an agent should hold inquiry to but makes no mention of the (perhaps higher) individual aims an inquirer may bring to an inquiry. We will now consider some cases in which agents aim for knowledge+.

3.2 Cases

Sometimes, an agent is in a good enough position to know now but has reason to hope that they may be in a position to have knowledge+ later if they continue inquiry. Consider the following cases:

The Detective

A detective is working through several bins of evidence to determine the innocence or guilt of a suspect. After looking through the evidence in 19 of 20 bins, the detective is extremely confident, and rightly so, that the suspect is innocent. Accordingly, she is also reasonably quite confident that the evidence in the 20th bin will also support, or at least not overturn, what the detective takes to be a compelling case for the suspect's innocence. And indeed, it will turn out that the 20th bin does not contain any defeaters for her belief, the suspect *is* innocent, and there aren't any Gettier-traps or other knowledge-defusing devices nearby. Nevertheless, not yet having looked in the 20th bin, she reasons that it would be epistemically irresponsible of her to close inquiry now. After all, there is new, easily obtainable, significant evidence that is available to her. And she chooses not to close inquiry until she has considered the rest of the easily obtainable evidence in the final bin. After all, she wants to be *extra* sure she gets this one right.

The Scientist

A team of scientists has performed and logged the results of 100 tests of a certain chemical reaction to inductively determine whether their hypothesis is correct. All but two of the tests confirmed the hypothesis and, in the circumstances of the experiment, it would be reasonable (and in fact correct) to write off the two disconfirming results as testing errors. On the basis of the experiment, the scientists would be justified in believing the hypothesis, which, it will turn out, is true and knowable. The team of scientists meets to decide whether or not to close their inquiry and write their paper on the basis of their results so far. Despite their very good evidence, the scientists decide not to close inquiry yet. Rather, they continue inquiry and decide to run one more round of 100 tests. After all, they already have the equipment set up, and it won't take as long the second time. And they are proud of their exacting epistemic standards in the laboratory.

The Philosophy Student

A philosophy student is studying Descartes' *cogito* and trying to make sense of what the structure of the argument is. They have read the relevant sections of the *Meditations* several times and listened closely to their instructor's interpretation of the argument. They are a good reader of the text, and their philosophy instructor did indeed present an accurate representation of Descartes' argument. If on an exam they were asked the question, "What is the structure of the *cogito*?" they would give a knowledgeable response.[51] Nevertheless, the philosophy student is deeply interested in the way that the *cogito* works, and they remain

---

[51] Some philosophers think we are not often in a position to know substantive philosophical theses. Notice, however, that the philosophy student is answering a question about the structure of Descartes' argument, not a substantive question of, e.g., epistemology or metaphysics. So, I don't think we need to be overly concerned that knowledge is available to the student in this case.

curious about the argument's structure. They are not yet satisfied with their inquiry, and this is partly explained by their (perhaps unarticulated) interest in really *understanding* the argument in a way that enables them to explain how they know (to themselves or to someone else). Instead of committing to a belief now, they continue wondering and make a note to look up the *SEP* article on Descartes later that evening.[52]

In all three cases, it's extremely intuitive that it's both permissible and indeed better for the subject to continue inquiry until they've finished gathering the specified evidence. Maybe it's not *wrong* of them to stop inquiry right now. After all, their evidential situations are already quite good. But we think the ones who continue inquiry have done even better, epistemically speaking. They have been more thorough, more rigorous, more epistemically responsible. And, plausibly, part of the reason they have been more rigorous is that they were (permissibly) aiming for more than mere knowledge. They wanted to be sure or to understand or to be extra justified: they wanted knowledge+.

We are trying to show that it is sometimes epistemically permissible and better, relative to the goals of inquiry on a specific question, to continue inquiry even though one could knowledgeably end it. These cases suggest that this is, indeed, true. I want to say more about why, on reflection, this shouldn't be so surprising theoretically, but first, it will be helpful to ward off some potential misconceptions.

---

[52] Similar cases appear elsewhere in the literature. For instance, Joseph Raz (1975, 37–38) considers cases in which an agent waits to decide until they will be in a better position to judge because they will be in better cognitive condition (i.e. more alert) and have more time to evaluate evidence they already have. Schroeder (2012, 471–72) considers cases in which you know that there is incoming evidence and argues that this can give you a "right-kind" of reason to refrain from belief. Compare also to discussion in McGrath (forthcoming, §1) and (manuscript), who notes both "future-comparative" and "goal-related" reasons that may bear on suspension, and to the reasons cited to continue investigation throughout Flores and Woodard (manuscript). One notable difference between the cases in the literature and mine is that these cases focus on the ambition of the inquirers to achieve more than mere knowledge.

First, it's compatible with it being *better* for the detective and scientist to continue inquiry that it would also be *permissible* for them to stop (or, indeed, for them to give up on the investigation altogether). The claim isn't that it'd be wrong to know, only that it'd be even better to continue inquiring. Second, it's compatible with its being better for our protagonists to continue inquiry that, at a certain point, it would be counterproductive for them to look at more evidence. Suppose that the scientist has looked at 1,352 of 4,000 test tubes, and each of the first 1,352 has confirmed the intended hypothesis. Perhaps in this case, the returns on inquisitive investment have diminished to the point that it's no longer better for the scientist to continue inquiring. Maybe, maybe not. But in any case, there's no incoherence in thinking that it's better for the scientist to look at samples 101–200 even if it's not better for the scientist to look at samples 1,353–4,000.

One might be worried that something about the *professional* norms of detectives or scientists or philosophers is interfering with our judgments about what is *epistemically* better. Sure, it's better for (e.g.) detectives to be extra thorough and continue inquiry even when they could knowledgeably close inquiry, but (the objector insists) that is because the professional norms of detectives require unusually stringent standards beyond the norms of everyday epistemic agents.[53]

But I don't think the professional standards are actually interfering with our judgments of the case. For one thing, detectives and scientists (and—one dares to hope—

---

[53] Feldman (2002) is representative of this view:
"[W]here there are reasons to gather evidence, they are nonepistemological or else there is no duty to gather evidence at all. We may have moral or prudential duties to learn about topics involving our health or important political issues. Philosophers and other academicians may have duties arising out of their positions to learn about the latest information in their fields. These are highly intellectual matters, but the duties arise not from some general intellectual requirement we all have, but out of moral or prudential considerations. Thus, where there are duties to gather evidence, they are not epistemological duties" (2002, 382).

philosophers!) are supposed to be professionally good *epistemic* agents. If detectives and scientists do better to continue inquiry, it's presumably because there is something *epistemically* better about continuing inquiry. It would be shocking to discover that the generic, professional norms of detectives and scientists require them to act in epistemically *worse* ways than in professions in which the aims are not explicitly intellectual.

That doesn't mean there is no difference at all between the lay investigator and the detective. Perhaps the detective has a professional duty that we do not normally have—a duty to look at all the relevant, available evidence before concluding inquiry. But (if so) that's a duty for the detective to do something that would be epistemically praiseworthy (though perhaps not required) for *any* inquirer properly situated to do. Note that changing the story so that the detective or scientists is an amateur—and so not beholden to professional norms—doesn't obviously change our intuitions about the case. What matters, ultimately, isn't whether the agents are professionals but whether the agents are *curious* in the question in the sort of way that leads them to aim for knowledge+. And curiosity does not, by itself, smuggle in professional or otherwise non-epistemic norms. Being beholden to professional norms may be one way to acquire (or have reason to acquire) the relevant kind of curiosity, but it isn't the only way to have it. Curiosity is democratic: it visits the lay and the learned alike.

One might also be worried that high stakes are interfering with our judgments. Perhaps in high stakes contexts, we don't get to count as knowing unless we have especially good epistemic standing that requires (or tends to require) more rigorous

inquiry, such that the agents in our story don't have knowledge yet.[54] But the cases don't seem to depend on the question involving high stakes. Perhaps there are high stakes involved in the detective's inquiry, but the scientist might be studying samples of mere theoretical interest, and there certainly isn't much riding on whether the student gets Descartes' argument right (we can suppose she has safely earned her A by this point in the course). The stakes of a situation might affect what epistemic goal an agent is (or should be) aspiring to (mere knowledge or knowledge+), but it is the goals themselves and not the stakes that are operative in these cases.

One could argue that although these are clearly cases in which the agents should continue to *investigate* (that is, to inquire_ACTIVITY) it isn't obvious that these are cases in which the agents should continue to inquire in the sense of this paper that entails having an inquiring attitude like wondering.[55] But I think this is an odd way to read the cases. Imagine a superintendent telling the detective: "Now, it's alright for you to continue to *investigate* who committed the crime—just make sure you don't wonder who did it!" And it's easier to imagine the philosophy student planning to read and investigate more about Descartes' argument *because* they are genuinely curious, not despite their lack of curiosity. In future sections, I'll argue that there are significant question-relevant costs for these agents not to inquire in a way that involves genuine curiosity (§3.3) or to close inquiry by forming a belief before ending their investigation (§4.3)—as a rule, we tend to be better investigators when we are also open inquirers. For now, it's enough to note that

---

[54] See Fantl and McGrath (2002) and DeRose (2002) for classic pragmatic encroachment and contextualist interpretations of high stakes situations.

[55] This is perhaps the view defended by Hall and Johnson (1998), at least if they take the relevant kind of acceptance to conflict with genuine inquiry:

"[W]e think you should follow an anti-sceptical synchronic epistemic duty and accept (now) any proposition which your present evidence supports. What [our principle] says is that you should also continue gathering evidence" (1998, 133).

the cases are most naturally read as ones in which the agents do not just plan to *investigate* in the future but in a curious state of *inquiry* now.

Let's take stock. We seem to have cases that suggest that, sometimes, even though they are in a position to know, it's epistemically better for people to continue inquiry than to knowledgeably close it. The relevant cases are those in which agents aim for knowledge+. But cases by themselves may not convince us if we do not see how they fit a broader pattern. Now I'll try to provide some broader motivation for why agents should sometimes continue inquiry instead of settling for immediate knowledge.

3.3 Wine ought to end inquiry

Knowledge is great. But it isn't the best thing epistemically. One can know, but one can also *know better*, or *know that one knows*, or *understand*, or be *rationally certain*. Keeping inquiry open enables agents to aim for these higher epistemic goods. Agents may not *have* to raise their epistemic standards in this way, but it's to their credit if they do. Inquiry can have a wide number of epistemic aims. It might be that all of the appropriate ways to close inquiry entail having knowledge. But that doesn't mean that mere knowledge is the only appropriate goal. It's laudable to aim for knowledge+.

Stopping inquiry too early comes with an opportunity cost: one could know right now, but one might thereby miss out on knowledge+ in the future. It's easy to overlook opportunity costs. It's also intuitive that (within our financial means) we should always accept bets with a positive expected outcome, even though we really shouldn't do so if it precludes making a better bet later. But opportunity costs are real costs, and we ignore them to our disadvantage in epistemology as in life generally.

I find it helpful to think about this problem for inquiry in relation to Pollock's (1983) aging wine paradox. In the aging wine paradox, we are asked when to open a bottle of wine that ages well in perpetuity: for any time t, it is true that the wine will taste better if opened at $t+1$ than if it is opened at t. The puzzle: it seems like it would always be wrong (from the perspective of practical rationality) to drink the wine. After all, it would always be better to wait longer. But if it's always wrong, then one should never drink the wine, which is worse! I accept the following (Pollockian) judgment about the aging wine paradox: it's always permissible to drink the wine, but it is still better to wait. For any time t, it's permissible to open the wine at t, but also: for any time t, it's better (practically speaking) to open the wine at $t+1$ than at t. Waiting to open the wine is thus never required but always supererogatory.

There's an analogy between the wine paradox and our decision about when to end inquiry. Just as it is initially puzzling when I should open the wine, so it is a puzzling question, "When should I close inquiry?" Like a fine wine, inquiry ages well—inquiries of a certain vintage, anyway. Inquiry allows us to be sensitive to new evidence. And for many propositions, new, significant evidence is constantly incoming. We'll have more information tomorrow (or next year, or next decade) about the fundamental laws of physics, about whether my friend will show up for coffee, about the latest arguments for the free will debate. And we also have more opportunity to reflect further on information we already have.

Of course, this isn't true for every proposition. I briefly acknowledge the frequent exceptions: sometimes, we have reason to think our epistemic situation will get worse over time. At the moment my friend shows up for the party, I probably have as good of

evidence for their being there as I'll ever have. Tomorrow, I'll be relying on my memory, and my memories of everyday events like that tend to fade over time.[56] Or I could know that my epistemic faculties will be compromised later so that I'm in as good a position now to judge on a question as I ever will be. These are cases of a bad epistemic terrain, where I know that additional inquiry will not lead to improved epistemic position.[57]

In other cases, inquiring well might lead to an improved epistemic position, but I could know that I am unlikely to inquire well given the circumstances. Perhaps I know that if I start inquiring into the cause of my symptoms now before my checkup, I won't be able to stop myself from reading horror stories on WebMD that will be not only psychologically but also epistemically counterproductive. In other cases, I might need to "buy in" to a subject matter or undergo a certain perspective-shift by believing certain things before I can productively inquire into related questions.

Finally, there are numerous cases in which there are morally or pragmatically inflected (and perhaps also epistemic) reasons not to continue inquiry. It could be obvious that I've reached the point of marginal returns on future inquiry, such that the (epistemic or practical) cost of keeping inquiry open is too great to justify continuing. Systemic racism, sexism or other forms of prejudice may make continued inquiry unsafe for members of oppressed groups to pursue. And there might be moral or practical reasons that I simply have to form a judgment on a question now: it's no good waiting until the

---

[56] Thanks to conversation with Matt McGrath and Ernest Sosa here.
[57] McGrath (forthcoming) considers such cases under the guise of "epistemic worsening" and Friedman (2019a) considers patterns of reopening inquiry whereby an agent's epistemic position deteriorates over time.

cyclist hits me to stop inquiring into whether we're on a collision course, no matter how good my future evidence will be.[58]

I emphasize: we must be careful not to overgeneralize the extent of cases in which continued inquiry is advantageous. We must always remain sensitive to the local (yet not thereby unsystematic) epistemic, social, and psychological factors that bear on the expected epistemic payoff of inquiring further.

Nevertheless, it often happens that we can reasonably expect to be in a better epistemic position to judge if we wait to judge until later than if we judge now. Such cases are not universal, but neither are they modally remote. The point isn't that inquiring ignorance is always the right tool for the job, it's that inquiring ignorance is a valuable tool to have in our epistemic toolkit.

In Pollock's discussion of the wine paradox, he describes the puzzle-making feature of the case as involving "some type of act which can only be done once, but the longer it is postponed the greater the expectation value for doing it" (Pollock 1983, 417). Now, it's not quite true that one can only close inquiry once. One can close inquiry and reopen it later only to close it again a second time.

But that doesn't mean there aren't real costs to closing inquiry early. There is an opportunity cost in lost sensitivity to future evidence from the time of closing until (if ever) it is reopened. Inquiry into Q gives agents extra sensitivity to new evidence bearing on Q. If I'm wondering whether the train has come yet, I'm more likely to hear the train whistle, even from a distance. And I'm more likely to correctly use that information to

---

make a judgment about whether the train has come. Inquiry keeps you question-sensitive and evidence-hungry.[59]

The aging wine paradox also gives us another way to think about the dual-threshold model for inquiry. Here are two questions: (1) When is the wine good enough to drink? (2) When should I open the bottle? These questions are obviously related. One shouldn't open the bottle if the wine isn't good enough to drink. But that doesn't mean one *should* open it as soon as the wine is good enough to drink (although one *may*). It's reasonable to hold out for more than mere drinkability. In the same way, it's reasonable to hold out for knowledge+ even when knowledge is already within our reach.

4. Against Closing Inquiry

So far, we've argued that it's sometimes better to continue inquiry than to close it, even if one could close it knowledgeably. But this doesn't automatically mean that we should choose ignorance over knowledge in such situations. Knowledge is great, and so is continuing inquiry. Why not just do both? The detective could simultaneously know that the suspect is innocent and still inquire into whether they are, turning out the contents of the final box of evidence. The scientists could know their hypothesis to be true while still inquiring into whether it is, checking the final results of the experiment.

Indeed, it may initially strike us as hard to understand the prolonged inquiries of our protagonists as anything but knowledgeable. We stipulated that the detective (for example) was in a position to know that the defendant was innocent but continued inquiry anyway. Perhaps the detective recognized that she was in this situation and thought to herself, "Even though I am in a position to know that the defendant is innocent, I will

---

[59] I owe the "evidence-hungry" metaphor to Ezra Rubenstein.

choose to continue inquiry." But if *that* is what the detective is thinking, it's especially hard to resist the intuition that she is continuing inquiry *knowledgeably*. If we *know* we are in a position to know that *p*, it's psychologically difficult for us to avoid coming to know *p* itself. And failures to make this obvious inference seem like an epistemic failing, not a success story.

But I don't think this is how we need to understand the psychology of our band of inquirers. Although we the audience (having peeked inside the box and the final test tubes or having already read the SEP article) can see from our privileged, third-person vantage point that the agents are in a position to know, nothing in our story demands that the detective or philosopher or scientists themselves know that they are in a position to know the answer to their continuing inquiries.

But how should we understand the agency of our protagonists? What *is* the thought process that would lead them to continue inquiry *sans* knowledge even though knowledge is within their grasp? The answer, I think, is that they aren't thinking about knowledge at all. They are thinking about whether they are *satisfied* with the progress of their inquiry so far. What it takes to *satisfy* inquiry has an open texture: it depends on what the agent's epistemic aims (explicit or implicit) in that context are. And, as we've noted, those aims could be more stringent than mere knowledge. It's this diversity of higher epistemic aims that allows agents to overlook the fact (even when it's available to them) that they are in a position to successfully close inquiry with knowledge. And this is proper. After all, if the agent isn't aiming at (mere) knowledge but at (some variety of) knowledge+, then they should be more sensitive to whether or not their inquiry has

progressed to the point that can give knowledge+ than to whether or not their inquiry has

progressed to the point that can give mere knowledge.[60]

This explains why agents don't as a matter of fact simply stop inquiry as soon as

they are able to do so knowledgeably: sometimes they are sensitive to other epistemic

standards. But we might still wonder whether it wouldn't be better for agents to continue

inquiry knowledgeably. Knowledge is a good thing even if it wasn't the good thing the

agent was looking for.

Attractive though knowledgeable inquiry may seem, it is not epistemically proper.

Inquiry into a question should be accompanied by genuine ignorance constituting an open

stance as to its answer. Here is the basic thought: it's incoherent to ask and answer a

question at the same time. And since inquiring is a way of asking and knowing is a way

of answering, one shouldn't do both. In what follows, I articulate three strategies for

making good on this idea that it's incoherent to (inquiringly) ask and (knowledgeably)

answer a question simultaneously.

4.1 Incoherence as (Psychological) Impossibility

The first strategy for making good on the intuition that it's incoherent to ask a

question one knows the answer to is to argue that doing so in the relevant sense is

psychologically impossible. There are lots of everyday cases that make this strategy look

implausible at first glance—I ask about a friend's plans even though (I now recall) I

already knew they would be at the party. But we've restricted the scope of our argument

---

[60] Some philosophers view the sort of agency that aims for such lofty cognitive achievements with
skepticism. Thus, Kvanvig writes, "To inform someone that she had found the truth but had fallen short of
[even mere!] knowledge would be met, I submit, with relative indifference" (2003, 148). This is no doubt
true for some agents. But I think that the existence of cases in which agents diligently pursue long inquiries
that achieve a high degree of epistemic success suggests that agents are often enough aiming (or at least
sensitive to) a much higher goal than mere truth.

to *occurrent* beliefs and *occurrent* inquiries. When I asked my friend about their plans even though I knew they would be at the party, I was able to do that because my mind was, in some sense, not fully attentive to what I knew. If both my wondering and my believing were at the forefront of my mind, could I really wonder about a question that I also know the answer to? Could I have the thought, "*p*, but is it really the case that *p*?" and retain my full belief that *p* through the end of the sentence?

Here is a reason to think not: it's hard to wonder about things we know on demand. Try right now to wonder whether you have just read this sentence. I find that it's hard for me to do so just by trying. I can't just *generate* curiosity in the question, and a plausible explanation for why I can't be curious in the question is that I also can't help but believe (and also know) the answer. On this way of thinking, curiosity needs doubt like fire needs air.

One strategy by which I might be able to get myself to wonder about something I know is to attend to various skeptical arguments, perhaps working through the Pyrrhonian modes, for instance.[61] But this method only works by inducing doubt and so removing outright belief.[62] In contrast, it's easy to induce at least *some* curiosity in questions about which I have no opinion just by raising them: I can easily wonder how many stars there are or what dogs dream about. Maybe our minds just aren't structured to (occurrently) wonder about things we (occurrently) believe answers to.[63] Just as it's very

---

[61] See Sextus (2000) edited by Annas and Barnes for the definitive translation of the Pyrrhonian modes in *Outlines of Scepticism*.

[62] For reflections on the relationship between doubt and belief, see Moon (2018).

[63] Things get trickier when the relevant beliefs are not obviously of the knowledge-producing kind. It isn't obviously impossible to believe the answer to a substantive philosophical thesis while inquiring into it, for instance. Nor is it obviously impossible to believe religious propositions like "God exists" while still inquiring into whether or not God does. Notably, these are also cases in which it's unclear whether agents really have full beliefs in the relevant sense. Fleisher (2017, forthcoming) argues that philosophical "beliefs" are often better understood as *endorsement*, and certain (though not all) varieties of religious

difficult (and perhaps impossible) to believe a proposition that we have no evidence in just by trying, so it's very difficult (and perhaps impossible) to wonder about a question that is fully closed for us. Knowledge is curiosity averse.

4.2 Incoherence as (Epistemic) Impermissibility

Although the argument above is suggestive, we won't rest the weight of the paper on it. Alice thought one couldn't believe impossible things. The Queen claimed to believe six impossible things before breakfast.[64] And philosophers have yet to decide who was right, even when the relevant impossibilities are obviously incoherent upon full reflection. The Queen's philosophers won't be much impressed by our first argument. After all, if we can manage to believe the blatantly impossible, we can surely also manage to ask questions we know the answer to, perhaps even occurrently.[65]

Lest we underestimate the human capacity for accepting cognitive conflict, we hasten to a second way of articulating the incoherence between inquiring and knowing: inquiring while knowing violates an epistemic norm. Friedman (2017) and Whitcomb (2017) are among the groundbreakers for this approach:

---

commitment may be better understood as faith than outright belief. I don't at all mean to insist that we *don't* have philosophical or religious beliefs. Some of us certainly do. But I think it's telling that the sorts of cases in which there is unclarity about whether we can (or can permissibly) inquire into a question that has an answer we are doxastically committed to are *also* cases in which there's unclarity about whether the doxastic commitment at issue is best classified as outright belief.

[64] See chapter 5 of Lewis Carrol's (1899) *Through the Looking Glass*.

[65] This echoes Friedman (2017), who is among the Queen's philosophers:

"It is commonly thought that subjects can have conflicting beliefs at a time, ...that a subject can both believe p and believe ¬p at a single time. Certainly, everyday experience taken on its face makes this sort of thing look not only possible but commonplace. There are typically (typically unnoticed) conflicts across our vast doxastic databases. But once we admit that doxastic conflict is possible, why shouldn't we admit that possibility in full generality? Just as a subject can be conflicted with respect to her beliefs on some matter, she should be able to be conflicted with respect to her beliefs and suspendings on some matter" (2017, 305).

Ignorance Norm for Inquiring: Necessarily, if one knows Q at t, then one ought not to be inquiring into Q at t.[66,67,68]

This section defends the **Ignorance Norm for Inquiring** and offers a new response to the charge that the motivating linguistic data for the norm can be otherwise explained in the same way that Moorean paradoxes can. But first: Why think that this norm is true?

One reason is that it sounds odd to claim to know and wonder about something at the same time. As Friedman notes, utterances like

1. ??I know the Yankees are winning the game, but are the Yankees winning the game?

or

2. ??I know the Yankees are winning the game, but I wonder whether the Yankees are winning the game.

are deeply puzzling.[69] Why is the speaker *asking* whether the Yankees are winning if they already *know* the answer to their own question? Normally, you answer someone's questions in order to *inform* them. But you wouldn't be informing your friend by answering them: after all, they already know. Of course, not all questions express a (genuine) desire to be informed. There are leading questions, rhetorical questions, and

---

[66] This is modeled after Friedman's original version, which reads thus: **Ignorance Norm for the IAs:** Necessarily, if one knows Q at t, then one ought not have an IA towards Q at t. (Friedman 2017, 311) IAs, in Friedman's nomenclature, are "interrogative attitudes," attitudes like *wondering* and *being curious about*. Since I've defined inquiry into a Q so as to entail having an IA toward that Q, I have reframed the norm as applying directly to inquiry.

[67] Whitcomb (2017) defends a related norm: **The Ignorance Norm**: Inquiring is constitutively governed by the norm *inquire as to what Q's answer is only if you don't know Q's answer.* (Whitcomb 2017, 152) So do Sapir and van Elswyk (manuscript): "One ought not inquire into Q at t and know p^Q at t." And Hawthorne (2004) suggests something similar for the speech act that corresponds to inquiring: "[A]s knowledge is the norm of assertion, ignorance is the norm of questioning" (Hawthorne 2004, 24).

[68] Friedman (2019b) also argues—and I agree—that there is a normative conflict between *full belief* and inquiry. See Whitcomb (2017) and Sapir and van Elswyk (manuscript) for the contrary view.

[69] A similar construction shows up in Kvanvig (2003) who writes: "It makes no sense to say, 'I know that it is raining, but I believe further inquiry is warranted'" (149) and also in Whitcomb (2017, 149).

riddles. But those phenomena don't seem descriptive of our case—your friend really wants to know. But they already know! So, what response could you make that would satisfy them?

I think the most plausible way to interpret the sentence is that, although the speaker knows the answer to their question, they aren't entirely psychologically at ease about it: they need some extra (re)affirmation. (Yes, the Yankees really are winning—it's ok.) This, at least, is comprehensible. But it doesn't feel "epistemically happy" (Friedman 2017, 310). You might overlook it if your friend needs extra assurances once or twice, but the situation quickly gets ridiculous when iterated:

> Speaker: I know the Yankees are winning the game, but are they winning the game?
>
> You: Yes, it's ok. The Yankees are winning.
>
> Speaker: Ok, good. I'm glad to know that: but are the Yankees winning the game?
>
> You: Yes…
>
> Speaker: Ok, good: I'm glad to know that: but are the Yankees winning the game?
>
> You: ...

If there's nothing epistemically unhappy about the original case, it's hard to explain why things get progressively worse in the iterated version.

But perhaps even though it's problematic to *say* (and also to *believe*) that one knows the answer to Q and is inquiring into Q, that doesn't give us a strong reason to think there is a norm against knowing and wondering together. Compare "Friedman sentences" to Moorean paradoxes:

3. ??The Yankees are winning, but I don't know that the Yankees are winning.

Both Friedman sentences and Moorean paradoxes sound strange to say despite being semantically consistent[70] and describing possible states of affairs—it's *possible* that the Yankees are winning while I don't know that they are; it's *possible* that I know the Yankees are winning while I also wonder whether they are.[71] Indeed, Whitcomb goes so far as to call related constructions "Moore-paradoxical questions" (Whitcomb 2017, 149).[72]

But Moorean paradoxes teach us an important lesson: we shouldn't infer too quickly from a thing's being bad to *say* that the thing said expresses a normatively defective state. After all, there isn't automatically anything normatively defective about the thing expressed by a Moorean Paradox being true. One can't infer from the truth of <The Yankees are winning, but I don't know that the Yankees are winning> that I have done something wrong. I'm not required to know every truth. I'm not even required to know the truth of every proposition I'm attentive to. Perhaps I have misleading evidence—the boisterous cheering of Red Sox fans nearby—that (misleadingly) prompts me to think that the Yankees are losing. If so, the Moorean paradox would be true, but I would be epistemically innocent.

Maybe Friedman sentences work the same way. Maybe even though it's bad to *say* (to the world or to oneself), "I know the answer to Q, but Q?" it's not automatically

---

[70] Although see Gillies (2001) for a defense of the unorthodox view that Moorean paradoxes are semantically *in*consistent

[71] We set aside, for this section, the arguments in §4.1.

[72] Although Whitcomb discusses both forms, he spends more time focusing on variants of the form, "*p*, but is it the case that *p*?" I'll focus, instead, on the ways that the linguistic profiles of Moorean paradoxes differ from "I know that *p*, but is it the case that *p*?" Whitcomb presents an interesting argument especially focused on the earlier form to an ignorance norm on inquiry, an argument that begins with the assumption that knowledge is the norm of assertion. One dialectical advantage of focusing on the latter form (what I call "Friedman sentences") is that, since knowledge is built into the first conjunct, arguing for an ignorance norm on inquiry doesn't require a corresponding view about the relationship between knowledge and assertion.

bad to both know the answer to Q and inquire into Q, which is what the **Ignorance Norm for Inquiring** requires. So, if Friedman sentences are to support the **Ignorance Norm for Inquiring**, we must be able to distinguish the kind of awkwardness in Friedman sentences from the kind of awkwardness in Moorean paradoxes.

Friedman considers the relevance of Moorean paradoxes, and her way of distinguishing them from her sentences rests on her thesis that having an inquiring attitude entails suspending judgment. Since, in Friedman's picture, inquirers into Q have thereby suspended judgment on Q, agents who both inquire into Q while knowing (and so believing) the answer to Q find themselves in doxastic conflict.[73] Just as it's wrong to believe p and also to believe $\neg p$, so it's wrong to suspend judgment on Q while believing p as the answer to Q. In the former case, the conflict arises because agents are committed to contradictory answers to a question. In the latter case, the conflict arises because belief commits the agent to an answer whereas suspension of judgment involves a commitment to neutrality: belief requires partiality and suspension impartiality.

Having argued that those who inquire into Q while believing an answer to Q are mired in doxastic conflict, Friedman now has a distinction to drive a wedge between her sentences and Moorean paradoxes. For Friedman thinks that "[c]lear-eyed assertions of first-personal doxastic conflict" are "worse by a mile" than third-personal reports of doxastic conflict but that this is "not the pattern we find in ...Moore-paradoxical assertions" (Friedman 2017, 312).

But there are two reasons to hope for a different explanation than Friedman gives us. The first is dialectical: Friedman's thesis that to inquire (or at least to have an IA) is to

---

[73] Although see McGrath (forthcoming) for an argument that suspension of judgment is not, strictly speaking, a *doxastic* state

suspend judgment is controversial.[74] One might think, instead, that suspension of

judgment is putting judgment off (McGrath forthcoming) or intentionally omitting

affirmation (or denial) because of an insufficiently high confidence toward a proposition

(Sosa forthcoming). We won't settle this debate here. But it would be nice to have an

explanation of how Moorean paradoxes and Friedman sentences differ that does not

depend on Friedman's account of suspension.

The second reason is this: it's unclear whether Friedman sentences (or Moorean

paradoxes) exhibit the cross-personal pattern that Friedman suggests. Moorean paradoxes

do sound worse by a mile when expressed in the first person, and Friedman sentences

don't obviously improve when taken out of it.

In fact, I think the best way to distinguish Moorean paradoxes from Friedman

sentences goes in the opposite direction of what Friedman (2017) suggests. That

Moorean paradoxes dissolve when taken out of the first-person present gives us a reason

to think that there isn't an underlying norm prohibiting the thing expressed (p's being true

and S's not knowing that p is true): the air of paradox only arises when an agent commits

to that being true of themselves right now. Instead, there's something wrong with

committing to (through saying, to oneself or to the world) a Moorean proposition, not

with the proposition being true. There are different stories about why this is so—one is

that asserting a sentence commits the agent to being able to know it and (first-person

present) Moorean paradoxes are unknowable.[75] In contrast, the dysfunction of *Friedman*

sentences, I shall argue, is robust enough to survive third-personal and suppositional

---

[74] Avery Archer (2018) explicitly argues against this aspect of Friedman's picture, for instance.
[75] See my (2020, 336–40), which builds on chapter 11 of Williamson (2000), for this explanation of why Moorean paradoxes are not properly assertable.

contexts.[76] This suggests that there is really something wrong with the state of affairs underwriting Friedman sentences rather than (merely) something wrong with committedly expressing them.

First, consider how Moorean paradoxes dissolve in the third person:

4. The Yankees are winning, but they don't know the Yankees are winning.

But we still feel that something is amiss when Friedman sentences are put in the third person:

5. ??They know that the Yankees are winning, but are the Yankees winning?

or

6. ??They know that the Yankees are winning, but they are wondering: are the Yankees winning?

Even if (6) is possibly true, we strongly feel that the agent is not behaving well epistemically, evincing the presence of a violated epistemic norm. We see a similar pattern for the past tense, where the awkwardness of Moorean paradoxes dissolves but the awkwardness of Friedman sentences merely gets stirred around:

7. The Yankees were winning, but I didn't know that the Yankees were winning.

8. ??The Yankees were winning, but were the Yankees winning?

Similarly, the oddity of Moorean paradoxes dissolves when embedded under suppose:

9. Suppose that the Yankees are winning but I don't know that the Yankees are winning.

Again, this sounds totally fine! There is no intuition that (in the suppositional context) I have automatically done something wrong epistemically (recall the misleading evidence

---

[76] Friedman does briefly consider sentences of this sort in an intriguing footnote in her (2019b, fn. 18). I agree with much of what she says there.

from the boisterous Red Sox fans). But embedding Friedman sentences under "suppose" does not remove our intuition that bad epistemic things are happening (in the suppositional context):

> 10. ??Suppose that I know the Yankees are winning but I am wondering whether the Yankees are winning.

There's nothing wrong with *supposing* this, of course. But it's clear that, in the world of the supposition, I've done something awkward. Moreover, it's confusing to give advice to both know and inquire at the same time.

> 11. ??You should both know that the Yankees are winning and wonder whether the Yankees are winning.

But if knowing and wondering simultaneously were permissible, we'd expect (11) to, at least occasionally, be good advice to give.

The analogy between Moorean paradoxes and Friedman sentences has broken down. The awkwardness of Friedman sentences is more robust, capable of surviving third-person, past tense, and suppositional contexts. An epistemic norm has the right kind of normative profile to explain our robust disapproval of agents who know and inquire simultaneously across such contexts. In each case, the ignorance norm on inquiring—the norm that one shouldn't inquire and know simultaneously—explains the awkwardness of Friedman sentences.[77]

But let's take a step back. It's one thing to notice that the awkwardness of certain sentences gives us evidence that a given norm must be true. It's another to give an

---

[77] For this section on Moorean paradoxes and Friedman sentences, I'm heavily indebted to conversations with Caley Howland and Thony Gillies.

explanation of why that norm holds. Why are knowledge and inquiry normatively allergic to each other?

Knowledge and inquiry are normatively related because questions and answers are normatively related. It is good for questions and answers to come in pairs: Questions *want* answering. And answers aren't answers at all unless they are answers to questions. One reason to think that questions and answers are supposed to come in pairs is this: When a question does not have an answer, it is a defective question. And at least if we know a question is defective, we shouldn't ask it. The question "Why does the sun rise in the West?" is defective: the sun, if it rises at all, rises in the East. There is no true answer to that question. And because it has no true answer, it's a question we don't and shouldn't ask.

That's a reason to think questions and answers should come in pairs and not alone. Here's a reason to think that three's a crowd: you should never believe different complete answers to the same question. If you believe <Colonel Mustard is in the library> as the complete answer to <Where is Colonel Mustard?>, you shouldn't also believe <Colonel Mustard is in the billiards room> as this would also completely answer the same question. When you have a complete answer to a question, one is barred epistemically from believing any competing answers to that same question, at least without giving up one's first belief. Complete answers *saturate* questions: questions only have room for one.

But what about *incomplete* answers that are, nevertheless, satisfactory? Recall that, intuitively, "I can buy a cup of coffee at Hidden Grounds" is a satisfactory answer to "Where can I buy a cup of coffee?" But surely that doesn't prohibit me from also

believing "I can buy a cup of coffee at Penstock Coffee Roasters!" So, shouldn't I be able to inquire into the question "Where can I buy a cup of coffee?" even while believing one (but not all) of the satisfactory answers to that question?

That's a tempting thought, but—intriguingly—the ban on inquiring into questions we know satisfactory answers to holds up even when those satisfactory answers are incomplete. Notice that Friedman sentences maintain their awkwardness for satisfactory-but-incomplete answers:

10. ??I know I can buy coffee at Hidden Grounds, but where can I buy a cup of coffee?

11. ??I know I can get to Penstock Coffee Roasters by taking a right on Third, but how can I get to Penstock?

Perhaps these don't sound *quite* as bad as when we embed complete answer and question pairs in Friedman sentences.[78] One reason that they might not is that expressing (10) or (11) might indicate that the speaker was *really* asking for an exhaustive answer with a sentence that (as it so happens) prompts a mention-some answer request in most contexts. So, in (10), for example, the speaker might be saying: "I know I can buy coffee at Hidden Grounds, but where *(are all of the nearby places)* where I can buy a cup of coffee?" But I think that (10) and (11) are clearly awkward if we hold fixed that the speakers are asking mention-some questions. Notice how much the sentence improves when we slightly alter the question by eliminating the known answer from the question with "else":

12. I know I can buy coffee at Hidden Grounds, but where *else* (i.e. other than at Hidden Grounds) can I buy a cup of coffee?

---

[78] I'm indebted here to Moyer and Syrett (forthcoming) and especially to conversation with Morgan Moyer.

13. I know I can get to Penstock Coffee Roasters by taking a right on Third, but how *else* can I get to Penstock?

Much better! Here's a suggestion: When we believe a complete answer to a question on our research agenda, we tend to remove the question from our research agenda to cohere with our answer. But when we believe a satisfactory-yet-incomplete answer to a question we once inquired into, we sometimes *modify* our question instead to exclude known answers from the set of possible answers to our new question. By changing the question, we accord with the **Ignorance Norm for Inquiring**. By following the norm, we also significantly improve the sentences: this further evinces the truth of the norm.

We are now in a position to make good on our promise to say what answers count as satisfactory. Satisfactory answers are all and only those to which the **Ignorance Norm for Inquiring** applies. A mark of satisfactory answers is whether or not they sound awkward when embedded in Friedman sentences.

Questions and satisfactory answers are complementary. Questions want satisfactory answers: answers aim to satisfy questions. This is why they come in pairs. And that's why it's both impermissible and incoherent to ask and answer the same question at the same time.

4.3 Incoherence as (Inquisitorial) Inefficiency

Let us return to our chart. There are four ways an agent who has been inquiring into Q could proceed, given that they would know if they formed a belief:

A	B
Knowing the Answer to Q	Knowing the Answer to Q
Continuing Inquiry into Q	Not Continuing Inquiry into Q

C	D
Not Knowing the Answer to Q	Not Knowing the Answer to Q
Continuing Inquiry into Q	Not Continuing Inquiry into Q

In §3, I argued that it is sometimes better to be on the left-hand side of the chart: continuing inquiry allows us to aim for more than mere knowledge, and agents can reasonably aim for more than mere knowledge even when knowledge is available. In this section, I've argued that, if one is on the left-hand side, it's better to be in box C than in box A. Inquiry is best done from a place of ignorance.

There is something incoherent about being in box A. Perhaps it's *impossible* to (occurrently) know the answer to Q while (occurrently) inquiring into it. But I've placed more weight on the idea that even if it is possible, there is a norm that makes knowing the answer to Q and inquiring into Q simultaneously *epistemically impermissible*. In this subsection, I articulate a third strategy for making good on the idea that it's incoherent to ask and answer a question at the same time: Confirmation bias makes us *inefficient* inquirers once we have already committed to an answer.

Closing inquiry by forming a belief subjects agents to confirmation bias, making them less likely to respond appropriately to future—especially disconfirming—evidence. In this section I'll argue that we should be sensitive to the costs that confirmation bias imposes on closing inquiry by forming a belief, even when that belief would constitute knowledge. We incur these costs whenever we close inquiry, so unlike the arguments in §4.1 and §4.2, this argument gives us a reason not to close inquiry whether we are inquiring or not (i.e. a reason to avoid being in either box A or B). Recall our protagonists from §3—the detective, the team of scientists, and the philosophy student. They decided that there was an opportunity cost to not continuing inquiry—to taking the

relevant question off their research agenda. Without continuing inquiry, they would be less likely to be sensitive to future evidence in a way that would give rise to knowledge+. Those same agents should also think that their chances of getting knowledge+ take a hit if they choose to close inquiry by forming a belief. Forming a belief will activate confirmation bias and make them less likely to interpret future evidence appropriately, making future knowledge+ less likely in turn.

But what is confirmation bias? Although there is widespread agreement that confirmation bias exists, there are many distinct phenomena that fly under its banner, and there is disagreement about how functionally unified the discrete phenomena are. Admittedly, some of them do not apply only when we have beliefs or favored hypotheses but when we do any kind of inquiry. In particular, Klayman and Ha (1987) show that humans tend to prefer "+testing" (roughly "looking for the presence of what you expect, as opposed to looking for what you do not expect" [Klayman 1995, 386])[79] even when this is not the optimal procedure for inquiry. This particular variety of confirmation bias is not obviously exacerbated by closing inquiry since the relevant hypothesis does not have to be preferred for the bias to have an effect. Nor are all forms of confirmation bias (at least, obviously) normatively dysfunctional. Koehler (1993) is quick to note that an observed tendency to rate disconfirming evidence as lower in quality than confirming evidence may be consistent with Bayesian approaches to rationality since they explicitly allow an agent's prior beliefs to factor into the evaluation of future evidence. Despite these caveats, there remains significant scientific evidence that agents tend to

[79] Klayman (1995, 386) seems to suggest that +testing might not be best classified as a species of confirmation bias at all.

systematically ignore or otherwise improperly assess counterevidence once they have formed beliefs.[80] I briefly recall some of the most compelling studies below.

Anderson, Lepper, and Ross (1980) show that "even after the initial evidential basis for their beliefs has been totally discredited, people fail to make appropriate revisions in those beliefs" (1980, 1042). In their study, students given fictitious evidence for a hypothesis maintained their raised levels of confidence in the hypothesis *even after* they had been informed that the evidence was made up. The failure to fully erase the weight of discredited evidence was exacerbated by, although not limited to, cases in which agents had developed a theory that "explained" the truth of their prior belief.

In another study, Lord, Ross, and Lepper (1979) show that, at least for socially sensitive topics, people become more polarized, not less, when confronted with ambivalent sets of evidence. Instead of having a moderating effect, exposure to mixed evidence tends to make people more polarized than they had been before. Interestingly, it wasn't necessarily the case that participants completely ignored the study that constituted counterevidence their previous belief; rather, they were more willing to accept criticisms of studies they disagreed with and so "rebound" to prior levels of confidence than they were for confirmatory studies (Lord, Ross, and Lepper 1979, 2105).

A resulting vast literature on confirmation bias shows that there are real costs to closing inquiry too early: it causes us to ignore or improperly weigh future evidence. If it

---

[80] The psychological literature is admittedly not always very clear about what kind of doxastic state is the minimum one for instigating the relevant kinds of confirmation bias. It's quite plausible that various kinds of doxastic commitment that fall short of the full, occurrent belief of traditional epistemology also instigate confirmation bias. My argument here only requires that full belief is *sufficient* for instigating certain varieties of confirmation bias.

is important enough to us to avoid these epistemic costs, it can be rational for us to forego knowledge, at least temporarily.[81]

But one might accept all this and insist that it has little bearing on epistemic normativity. What we ought to do is *both* (1) close inquiry now and (2) vigilantly resist the effects of confirmation bias. And we can fight the effects of confirmation bias—to a degree. Hernandez and Preston (2013) show, for instance, that increasing disfluency during one's evidential evaluation (e.g. by reading information written in hard-to-read font) can reduce, although not eliminate, the effects of confirmation bias. So, there's hope that we can resist confirmation bias in certain cases.

Indeed, I think that sometimes the right thing to do is to close inquiry while resisting confirmation bias as well as one can. But often the better strategy will be to not yet close inquiry. First, effortful resistance mitigates the risk of confirmation bias but, as Hernandez and Preston (2013) note, the bias does not disappear. So, there are still costs to closing inquiry even if they are not as high when accompanied by effortful resistance as they might have been. But second, we often just don't have the cognitive capacity available for the resistance strategy to be successful for one inquiry, let alone to pursue resistance as a general strategy for engaging the many inquiries we pursue in the course of an ordinary day. Life doesn't give us that kind of cognitive space. Moreover, Hernandez and Preston show that when we are under significant cognitive load, the

---

[81] Levi (1991) succinctly anticipates this general thought when discussing "uncoerced contraction" of belief:

"[C]ontraction does not import error from the inquirer's point of view. It incurs a loss of information. But if this loss is accompanied by the opportunity to undertake an inquiry to settle the question as to the truth of the rival theories without begging the question, there may be a benefit in doing so to compensate for the initial loss of information. If this is right, a good reason for implementing an uncoerced contraction would be that it allows a promising theory incompatible with current doctrine to be examined without prejudice" (Levi 1991, 153).

mitigating effects of disfluency on confirmation bias evaporate (2013, 181). So, *always* resisting confirmation bias—and incurring the cognitive load that such resistance brings with it—can't be a good general strategy for handling the costs of such bias.

One might suspect that by bringing in factors about psychological bias, a non-epistemic kind of normativity has been smuggled in. But I think this is exactly the kind of real-world normativity that epistemologists should care about. It's useful to compare this kind of epistemic situation to a case in ethics. Suppose you ought to apologize to your friend for something you did, and it would really be best if you apologized sooner rather than later. But right now you are still upset—you are bearing a high emotional load—and you know that if you try to go apologize now, you're likely to end up saying something hurtful that you'll regret, whereas if you take some time to cool off, you will be more likely to offer a carefully-worded apology. Should you apologize now or later?[82] There's a clear sense in which you should apologize later. This despite the fact that it would, in some sense, be better if you could apologize now while also vigilantly avoiding saying anything hurtful. But since you *know* you will likely say something hurtful if you apologize now, it often makes sense to wait. Apologizing well now isn't a genuine option for you. In any case, I'd rather be friends with the sort of person who makes the delayed but well-offered apology.

Similarly, given that I know I am the sort of person who is subject to confirmation bias, I would rather be the sort of person who (at least sometimes) waits to close inquiry

---

[82] This is, perhaps, a trickier question than it first appears and one that ethicists have spent considerable time wrestling with. If it helps, one can read the cases as ones in which *apologizing well now* or (in the epistemic case) *believing and yet successfully resisting the effects of confirmation bias* aren't genuine options for the agent. They are beyond the agent's ability and so beyond their obligations. But see, among others, Goldman (1978) for critical discussion of related issues.

when there's a high chance of doing better epistemically in the future. Perhaps ideal

epistemic agents, unencumbered by cognitive biases or with an unlimited cognitive

capacity to handle them, have no reason to continue inquiry in the sorts of cases we have

been discussing. Ideal epistemic agents can close and reopen inquiry at will without

thereby changing their sensitivity to new evidence. Maybe we can even approximate this

ideal with special effort in limited domains for limited periods of time. But for those of us

who are not currently operating at an ideal epistemic level, it can be better to keep inquiry

open in than to knowledgeably close it.

4.4 All Together Now

In §3, I argued that it's sometimes more important to continue inquiry than to

have knowledge. In such cases, it's better to be on the left side of our chart than the right.

A	B
Knowing the Answer to Q Continuing Inquiry into Q	Knowing the Answer to Q Not Continuing Inquiry into Q
C	D
Not Knowing the Answer to Q Continuing Inquiry into Q	Not Knowing the Answer to Q Not Continuing Inquiry into Q

In this section, I've argued that there's something epistemically defective about being in

box A. So, if one is on the left side of the chart, it's much better to be in box C than it is

to be in box A.

Our first strategy suggested that it's not even obvious that it's psychologically

possible for us to be in box A once we've limited the scope of our claim to occurrrent

beliefs and inquiries. If that's right, then C is the only box one can inhabit on the left side

of the chart. So, if it's better to be on the left side of the chart than the right (if it's better

to inquire than to know), then it's better to inquire in ignorance than it is to know. Sometimes, it's better to be (inquiringly) ignorant than (un-inquiringly) knowledgeable.

Our second strategy was to argue that it's epistemically impermissible to know the answer to a question one is inquiring into. Don't ask and answer a question at the same time. If this is right, then (if it's better to be on the left side) it is at least usually better to be in box C than in box A.

Why only *usually*? Well, most norms can be overridden or outweighed by competing values. So, it's possible that special circumstances conspire to make it right for us to violate the epistemic prohibition on knowing while inquiring (perhaps we have special reason to know the answer to Q and also special reason to remain as sensitive as possible to new evidence bearing on Q). But the ultimate conclusion of this project is modest enough to allow for such special circumstances. The thesis: *sometimes* it is epistemically better (with respect to one's inquiry into Q) to inquire into Q without knowing the answer to Q than it is to know the answer to Q. This allows that the norm may be overridden in certain cases. But given that there is a stable norm prohibiting knowing the answer to Q while inquiring into Q, it would be surprising if a great many of the cases in which it is better to be on the left side of the chart were not also going to be ones in which it is better to be in box C than in box A.

The third strategy was to show that inquirers into Q who know the answer to Q (and thereby believe an answer to Q) incur the costs of certain kinds of confirmation bias. The mere fact that there are costs does not by itself mean that we should never pay them. But it seems that these costs are going to matter most for the same sort of agents we considered in §3. If an agent is aiming for knowledge+ (but only knowledge is presently

available to them), they have reason to remain sensitive to future evidence. One way to do that is to continue inquiry; another is not to close it. Quite plausibly then, many of those agents who opted to continue inquiry rather than end it in §3 should also opt to avoid the costs of confirmation bias by foregoing immediate knowledge.

If the thought of foregoing knowledge is unsettling, it may be a comfort to note that our protagonists are not giving up *everything* good about their current epistemic standing by continuing inquiry. Importantly, if the inquiry goes well, they are very likely to remain in a position to know the answer to the question they are inquiring into. In these circumstances, the inquirer maintains equal epistemic *access* to knowledge even though they genuinely forego *having* it. And this may make the sacrifice seem less reckless.[83,84]

When it's better to inquire than not to, it's often better to inquire from a place of beliefless ignorance than from knowledge. Beliefless ignorance preserves a genuinely open stance to the ongoing inquiry.

5. Two Ways of Furthering Inquiry

I've argued that sometimes, even when one is in a strong position to know, it can be better to further inquiry without belief (and so without knowledge), thereby increasing

---

[83] I'm grateful to Susanna Schellenberg for conversation on this point.

[84] It's also compatible with my view that agents hold onto some doxastic state weaker than full belief while continuing inquiry (a high confidence, for instance). I don't have much to say here about the epistemic norms that may govern doxastic commitments that are less than full belief. But I think there's room for happy collaboration with, for instance, Will Fleisher's (2018) and (forthcoming) work on the doxastic attitude of *endorsement*, which is explicitly tailored for contexts in which (perhaps for social epistemic reasons) researchers need to take some kind of a (doxastic) stand on an issue even though knowledge is inappropriate. Fleisher focuses on cases in which knowledge is inappropriate because agents do not have knowledge available to them (because of entrenched peer disagreement, for instance); in the context of my project, knowledge might sometimes be inappropriate for researchers because what they ought to do (or at any rate *are* doing) is to inquire.

Space precludes a full discussion. But to those who worry that foregoing knowledgeable belief is too high a price to pay for inquiry, I offer that Fleisher's work may provide the key to allowing inquirers to maintain a kind of immediate (albeit not knowledgeable) doxastic success. (Thanks to Andy Egan for conversation on this topic.)

the likelihood of permissibly achieving knowledge+ in the future, rather than to settle for immediate knowledge. An important piece of the argument (§4) has been that furthering inquiry in the relevant way while also knowing the answer is epistemically inappropriate.

But (I will now suggest) it can be permissible to know the answer to a question and yet further inquiry in *some* sense. In this section, I distinguish between two ways of furthering inquiry: *continuing* inquiry and *advancing* inquiry. I concede that it is sometimes epistemically better to knowledgeably advance inquiry than to ignorantly continue it. But not always: there remains an important role for ignorantly continuing inquiry in the healthy epistemic life.

I've been focusing on the species of furthering inquiry that I have been calling *continuing* inquiry. To continue inquiry into Q is for Q to remain on one's research agenda: that is, to continue to have a goal-directed interrogative attitude toward a question that provokes sensitivity to new evidence.[85] But that doesn't describe all cases of furthering inquiry. Indeed, furthering inquiry in this way can look quite odd in certain cases.

Suppose I am talking to my friend over coffee and they tell me that they have quit their job. This is surprising to me, since they had always expressed that they were quite happy with their position, a post they had securely held for a long time. I remain deeply curious about my friend's situation. I don't merely want to know about their decision to quit, I want to understand it: I want, that is, a kind of knowledge+. And I immediately want to ask them more about quitting their job. But it would be extremely strange for me to withhold belief that my friend has quit their job while I (inquiringly) try to understand.

---

[85] Recall discussion in §2.1 and in Friedman (2017, 307–308).

Instead, I fully believe that my friend has quit their job, and my further inquiry will take the form of follow-up questions: *Why* did you quit your job? *When* did you do so? Did something unusual happen at work recently?

In fact, it's hard to see how I could even coherently ask some of these questions if I did not already fully believe that my friend had quit their job. Asking *when* they quit, for instance, *presupposes* that they already have. And while I could ask if something strange happened at work recently whether or not I believe they have quit, that question is newly salient to me because I am looking for a potential *explanation* for why they have quit.

This series of follow-up questions constitutes a way of furthering inquiry on the subject of my friend's quitting their job. But it isn't *continuing* inquiry as we've defined it. Let us say instead that such follow-up questions *advance* inquiry. Advancing inquiry doesn't continue inquiry into the same question, but it furthers inquiry into a subject, where a subject is understood as a set of questions that are related or unified in the right way.[86] More carefully, an agent advances inquiry only if (i) Q and Q' are questions that belong to the same subject and (ii) at least in part because the agent has closed inquiry into Q, they open inquiry into Q'.

So far, I've just said that there's a way of furthering inquiry that doesn't amount to continuing inquiry. That by itself doesn't make trouble for my argument. *Unless*, that is, every time that I've argued it would be best to *continue* inquiry it would actually be better to *advance* inquiry. For those who advance inquiry, unlike those who continue it, come to believe and so (in the good case) know the answer to their initial question.

---

[86] For our purposes, we can be fairly open about what it takes for questions belonging to a common subject to be related "in the right way."

And, indeed, in many cases, advancing inquiry can recoup some of the advantages I have been claiming for continuing it. For one way to improve our epistemic status with respect to a proposition is to learn more about related questions. Suppose that I know that my friend has quit their job on the basis of their testimony and I then learn (on the basis of further testimony in response to a follow-up question) that they had been harboring doubts about the leadership of the company for the last year. This gives me additional evidence for the proposition that they have really quit (more than I needed for knowledge, since their initial testimony was sufficient). So, I now have a variety of knowledge+ that my friend has quit their job. And I have achieved this knowledge+ without giving up knowledge during the interval between my friend's initial testimony and their answer to my follow-up question. Isn't this a better strategy?

This problem may seem particularly acute when the follow-up question is explicitly epistemic in character—for instance, if Q is *whether p* and Q' is *whether I know that p*.[87] Suppose, for instance, that an agent is curious about p and is not inclined to be satisfied with inquiry into whether p until they (are in a position to) know that they know that p. Here are two ways that they could get there:

> **The Method of Delayed Gratification**: S is inquiring into the question of whether p. At t_1, S arrives in a strong position to know that p is true. But S withholds judgment about whether p. Later, at t_2, S arrives in a strong position to know that they know that p. They knowledgeably close inquiry, coming to believe that p. Simultaneously (or soon after), they may form the additional belief that they know that p is true.

> **The Scaffolding Method**: S is inquiring into the question of whether p. At t_1, S arrives in a strong position to know that p is true. S knowledgeably closes inquiry, coming to believe that p. S immediately (or soon after) opens inquiry into a new question: whether S knows whether p. Later, at t_2, S arrives in a strong position to know that they know that p. They knowledgeably close inquiry, coming to believe that they know that p is true.

---

[87] I'm grateful to Ernie Sosa for helpful discussion on this and related points.

Note that, as in the conversation with the friend who quit their job, **The Scaffolding Method** has an important advantage over **The Method of Delayed Gratification**. Only by following **The Scaffolding Method** does S get to know the answer to whether p from the interval t_1–t_2. And both methods produce similar epistemic situations at t_2. In the end of both stories, S can know both p and that they know that p. Moreover, by employing **The Scaffolding Method**, S preserves an interrogative attitude toward a p-involving question. So, it's at least arguable that S will reap many of the inquiry-activated benefits that I've claimed for **The Method of Delayed Gratification**. So, why give up knowledge during the interim? Why delay gratification when one can scaffold instead?

My response is partly concessive. I think that **The Scaffolding Method** is sometimes the best strategy. Note that those who follow **The Scaffolding Method** respect the **Ignorance Norm for Inquiry**, that one shouldn't both know an answer and inquire into *that very same* question simultaneously. So, any reason to employ **The Method of Delayed Gratification** over and against **The Scaffolding Method** will arise from the decision-theoretic reasons considered in §3 (loss of sensitivity to new evidence bearing on the question) and §4.3 (the costs of confirmation bias). These decision-theoretic reasons will, of course, be quite sensitive to local factors about the particular agent's epistemic and psychological situation.

Nevertheless, I think that there are often strong reasons to prefer **The Method of Delayed Gratification** to **The Scaffolding Method**. Agents who employ one method rather than the other are asking and have answered different questions, and this has the potential to shape the intermediary inquiry between t_1 and t_2 in important ways.

When following **The Method of Delayed Gratification**, S is inquiring into

<whether p> (Q) and not (at least until the very end of the story) <whether I (S) know

that p> (Q'). Although both questions are at least loosely about p, Q' has a distinctively

epistemic character and is likely to make different evidence salient. For instance, whether

I believe that p and what the nature of knowledge is are relevant considerations that bear

on Q' but not (in the typical case) on Q.[88] But if what I'm *really* interested in is whether p

and not whether I *know* that p, then I should want to remain most sensitive to evidence

without the epistemic character made salient by Q'.

As discussed in §4.3, S may also activate different varieties of confirmation bias

by utilizing **The Scaffolding Method**. Between t_1 and t_2, S already believes that p. So,

it's not just that by using this method S is prompted to be sensitive to evidence (of an

epistemic character) that is irrelevant to Q, S is also introducing confirmation bias to

evidence that bears directly on the first-order question Q.

The advantage to **The Scaffolding Method** is that one gets to know the answer to

one's initial question earlier. In some cases, this will be worth it. And when the evidence

is clear-cut enough, we may not even be able to help proceeding in this way. But the

Scaffolding Method brings with it (a) sensitivity to evidence bearing on epistemically-

inflected higher-order content, which may distract from the agent's guiding interest in the

subject, and (b) confirmation bias on the first-order question. When those costs are high

enough given an agent's goals in inquiry, they should still prefer **The Method of**

Delayed Gratification.

---

[88] Cf. Alston (1980).

Finally, I want to reflect on the different phenomenology of following one method rather than the other. There's an important difference between (a) being disposed to remain unsatisfied with an inquiry until one (e.g.) is in a position to know that one knows the answer and (b) being curious about (not just whether p but also) whether I know that p. Sometimes I am deeply interested in a question and have high standards for answering it. When I am deeply interested in a question in this way, I don't typically *lose* curiosity in that first-order question midway through the inquiry and gain curiosity in a closely related but distinct higher-order question about my epistemic standing. Rather, I preserve my curiosity in the first-order question throughout the entire inquiry. This phenomenology fits much better with **The Method of Delayed Gratification** than **The Scaffolding Method**.

Recall, for instance, the detective from earlier in this paper. When they look at the 20th bin of evidence, here are two questions whose curiosity could be motivating their search: (1) Is the suspect innocent? Or (2) Given that the suspect is innocent, do I know that the suspect is innocent? Unless the detective is of a surprisingly epistemological bent, I think it's much more natural to suppose that they are inquiring into the first question. They are curious about *the suspect's innocence* not their own epistemic accomplishments. And that's the question they want to retain evidential sensitivity toward as they examine the contents of the final bin.

Again, the goal is not to insist that *all* inquiry must fit this mold. It's very natural when discussing my friend's job situation over coffee to *stop* wondering whether they have quit their job and to become curious about the why and the how instead. Or in the epistemology room, it's natural to spend a lot of time thinking about whether we *know*

that we have hands and not just whether we have them. But it's also natural for the protagonists of this paper—the detective, the team of scientists, the philosophy student—to maintain curiosity in their initial questions throughout the inquiry. **The Method of Delayed Gratification** and **The Scaffolding Method** represent two modes of inquiry that are both more demanding than merely knowing the answer to a single question. Which method it's reasonable for an agent to employ may depend on the shape of their curiosity as much as anything else. When an agent's curiosity is *fixated* on a single question, prompts *high standards* for satisfaction, and when the agent is particularly motivated to remain sensitive to evidence bearing on the *initial*, first-order question, it is often more sensible to adopt **The Method of Delayed Gratification** than its competitor.

6. Knowledge-Hoarding and Its Limits

I want to think once more about our protagonists from earlier in the paper (the detective, the scientists, and the philosophy student) and to work out with the full argument in view why, at the moment we interrupt their lives, it's better for them to be ignorant than knowledgeable on the question they are taking up. Then I want to step back and think about the kind of epistemic life someone who takes the advice of this paper might be inclined to live.

Our philosophy student didn't just want to know what the structure of Descartes' argument in the *cogito* is—they were intensely interested in the question and had greater ambitions than mere knowledge. They wanted to understand the structure of the argument, to be extra justified in their beliefs about the argument, and to know the structure of the argument well enough to be able to articulate it in a clear way when asked: they wanted knowledge+. It's not necessarily that they cognized these aims in this

way, but they weren't disposed to feel satisfied with inquiry that didn't end with this kind of epistemic achievement. The student has read the text very well and would know (but not know+ in the relevant way) what the structure of Descartes' argument in the *cogito* is if they formed a belief about it. But they haven't yet. The student could proceed in any of the following four ways:

A	B
Knowing the Answer to Q Continuing Inquiry into Q	Knowing the Answer to Q Not Continuing Inquiry into Q
C	**D**
Not Knowing the Answer to Q Continuing Inquiry into Q	Not Knowing the Answer to Q Not Continuing Inquiry into Q

They don't follow path D. That would be to abandon inquiry into a question that they remain deeply interested in. They don't follow path B, for they want to continue inquiry! After all, they still want knowledge+ and all that's available to them right now (we've stipulated) is knowledge. They are more likely to get knowledge+ if they continue inquiry. Continuing inquiry will keep them sensitive to new evidence. What's more, choosing to know the answer would introduce a kind of confirmation bias, and that would (often) make it less likely that they would get knowledge+ in the future. They could, nevertheless, follow path A, committing to both continuing inquiry into their question and knowledgeably believing the answer. Or at least *maybe* they could—this is a difficult psychological trick to pull off if one keeps both things at the front of one's mind. But even if they could pull it off, this is an impermissible epistemic state to be in. It's incoherent to ask and answer the same question at the same time. And one still has to reckon the costs of confirmation bias if one takes this route. Better to avoid path A as

well. That leaves path C: continuing inquiry while withholding belief and so remaining ignorant of the answer.

Ignorance is the price of inquiry—at least inquiry of a certain kind. One must forego knowledge if one wants to permissibly inquire into a question while retaining heightened sensitivity to new evidence. But the price is reasonable if knowledge+ is both the agent's (desired enough) aim and (likely enough) reward.

Although it's initially surprising that ignorance might play such a prominent role in the life of virtuous epistemic agents, I think that the vision of epistemic life that accompanies it is attractive. It requires a certain humility to sit with ignorance during an inquiry. Humble inquirers want the answer, but they don't need to be the most knowledgeable person in the room: they aren't know-it-alls.[89]

But humility is not meekness. On the contrary, agents are encouraged to be epistemically ambitious! Knowledge isn't the best there is, nor is all curiosity sated by it. Have courage to aim for the higher things. It may be that knowledge is often good enough. But it would be sad, I think, for an epistemic agent never to try for more. Such an agent lacks a full capacity for wonder.

Ambition requires patience: the higher things do not often come easily. And those who have the appetite for knowledge+ must have the maturity and self-control to exercise delayed epistemic gratification. They must resist believing too soon. The idea that restraint is central to the well-lived epistemic life is deeply Cartesian:

> If, however, I simply refrain from making a judgement in cases where I do not perceive the truth with sufficient clarity and distinctness, then it is clear that I am behaving correctly and avoiding error. …[T]he perception of the intellect should always precede the determination of the will. In this incorrect use of free will may be found the privation which constitutes the essence of error (Descartes 1984, 41).

---

[89] I'm grateful to Susanna Schellenberg for conversation on this point.

Good inquirers (Descartes says) keep their wills from interfering with the perception of the intellect. Descartes is not merely concerned with the error of being wrong but the error of having arrived at the truth "by pure chance" (41): not just the error of being wrong but the error of not being epistemically successful "with sufficient clarity and distinctness." We need not have Cartesian standards about what counts as "sufficient clarity and distinctness" or a Cartesian theory of the will to appreciate the Cartesian insight that the restraint to withhold our judgment until our epistemic goals have been achieved is an important epistemic skill.

In contrast, the knowledge-hoarder grasps at knowledge as soon as they can. Lacking a full capacity for wonder—the capacity to be motivated in their inquiry by a desire for more than mere knowledge—the knowledge-hoarder thereby lacks epistemic ambition. Or if they have ambition, they are immature, unable to practice delayed epistemic gratification, the restraint necessary to achieve the greater epistemic goods. Or perhaps they are overly risk-averse, unable to stomach the risk of foregoing knowledge in a venture for knowledge+. Too proud to sit with the ignorance that proper inquiry requires. The knowledge-hoarder may know more, but those who can accept ignorance make better inquirers.[90]

Unlike that of the knowledge-hoarder, this project presents a picture of epistemology that is inquiry-positive. Inquiry is not just a dutiful means in the generation of knowledge, but an epistemically valuable and wonder-filled state in its own right in which we remain sensitive to evidence, attentive to questions, and from which we are able to pursue a wide range of epistemic aims. And not just inquiry but ignorance too

---

[90] Thanks to Carolina Flores for conversation about the dangers of knowledge-hoarding.

turns out to have an important role to play in the epistemic life: ignorance can be epistemically valuable. Although knowledge is allergic to inquiry, ignorance gives us the space to be good inquirers. Sometimes we inquire with the aim of achieving more than mere knowledge, and sustained ignorance helps us get there. It's often especially appropriate for us to do this if we have reason to expect new evidence to trickle in, or if we are willing to dedicate future time to interpreting the evidence we have, or if we are about to acquire a new cognitive skill. Or if we are filled with a deep enough sense of wonder. In epistemology, as in life, delayed gratification is often the best strategy.[91]

AGAINST THE DOCTRINE OF INFALLIBILITY

Abstract: According to the doctrine of infallibility, one is permitted to believe *p* if one knows that necessarily, one would be right if one believed that *p*. This plausible principle—made famous in Descartes' *cogito*—is false. There are some self-fulfilling, higher-order propositions one can't be wrong about but shouldn't believe anyway: believing them would immediately make one's overall doxastic state worse.

1. The Doctrine of Infallibility

A proposition is epistemically infallible for an agent just in case that it's

impossible for that agent to falsely believe it:

> **Infallibility**: A proposition *p* is infallible for S iff it's impossible that S falsely believes that *p*.

This definition closely resembles the following:

> **Infallibility***: A proposition *p* is infallible for S iff it's necessary that if S believes that *p* then S truly believes that *p*.

If, necessarily, all beliefs are either just true or just false—as I believe they are—then

infallibility and infallibility* are equivalent. I shall suppose that they are equivalent

throughout the rest of the paper, although there is interesting territory to explore for those

who believe that some propositions are neither or both true and false.

Since one can't be mistaken about one's infallible beliefs, it's tempting to think

we should always believe them. When there's no risk of false belief, why not? At least, if

we're in a position to recognize that a proposition is infallible for us, surely then we may

believe it. As Alston says, "one could hardly have a stronger (epistemic) justification for

holding a certain belief than the logical impossibility of the belief's being mistaken"

(Alston 1971, 229). Consider, then, the doctrine of infallibility:

> **Doctrine of Infallibility**: If S knows <necessarily, if she herself were now to believe that *p* then she would truly believe that *p*> then it is thereby (rationally) permissible for S now to believe that *p*.

In slogan form: You're always permitted to believe (known) guaranteed truths.

I argue, however, that not all propositions known to be infallible may be believed. In fact, some propositions that are known to be infallible should be disbelieved!

The doctrine of infallibility emerges as a battleground between two otherwise attractive philosophical theses: one the one hand, *veritism*, the thesis that accuracy is the fundamental epistemic good, and on the other hand what I shall call *reflectivism*, the thesis that one's reflective attitudes about one's first-order beliefs make at least some difference to the epistemic quality of those first-order beliefs or one's belief system. Although the thesis of this paper is that the doctrine of infallibility is false, a recurring theme is that reflectivism is in tension with veritism.

2. Clarifying "Infallibility"

Some clarifications are in order. First, infallibility is sometimes ascribed to agents in relation to a proposition or subject matter (e.g. "The Pope is infallible about matters of faith when he speaks *ex cathedra*"). Thus, Alston writes: "one can be said to be infallible *vis-à-vis* a certain subject matter provided one cannot be mistaken in any beliefs he forms concerning that subject matter" (Alston 1971, 229). Talking about infallible *agents* might suggest to the reader some kind of super-knower or an agent with extra-special epistemic access to some domain (like a Pope, prophet, or supercomputer).

Terminologically, I prefer to speak in the opposite way: Infallibility is a property that *propositions* have in relation to *agents*. This serves to emphasize that infallibility (as used in this paper) requires no special competence on the part of the agent for whom a

proposition is infallible. All necessary mathematical truths are infallible for infants, for instance, though of course infants do not *know* that those truths are infallible for them.

Focusing on propositions rather than "subject matters" also allows us to be more fine-grained in targeting particular beliefs. That I exist is infallible for me (in my sense), but it's at least not obvious that I couldn't be mistaken about the subject. Suppose I believe that I am a fiction, for instance, and that strictly speaking I don't exist.[1] Or that there are no selves at all. I'd then be mistaken about the question (or subject) of whether I exist even though, necessarily, if I believed <I exist> I'd be right. One can be mistaken about a subject matter even if one can't be mistaken in believing a particular proposition that is an answer to a question belonging to that subject matter.

My definition of "infallibility" also differs from certain other uses in the literature. For instance, Jeremy Fantl and Matthew McGrath suggest that an agent knows (and perhaps also believes) that *p* infallibly when there is no epistemic chance for them that not-*p* (2009, 11). This definition is in some ways more and in some ways less restrictive than the one used here. It is less restrictive because it counts as infallible those propositions for which an agent has maximal justification: there's no *epistemic* chance that not-*p* for such agents. But it is often still possible (though not *epistemically* possible, perhaps) that agents could be wrong about a proposition for which they have perfect justification: they could have had different evidence, for instance. Fantl and McGrath's definition is, in other respects, *more* restrictive than mine because they leave open that a belief in a necessary proposition may be fallible (e.g. when a math student believes an

---

[1] See Lebens (2015) for an intriguing articulation of a view in this neighborhood.

axiom on the basis of testimony rather than by working out the proof [cf. 2009, 8]) even though, necessarily, if they believe the axiom, they will be right.

Nor does my use of "infallibility" perfectly overlap with the notion employed in what Jessica Brown calls "probability 1 infallibilism" (Brown 2013, 626), the sort of infallibility one has toward a proposition when it has probability 1 conditional on one's evidence.[2] A proposition could have less than probability 1 and yet still be infallible for me. Necessary truths that I can't conceptualize may not have probability 1 for me (because I may not have *any* evidence bearing on them) and yet be such that, necessarily, if I believe them, I will be right.

Second, what I mean by "infallibility" is related to but distinct from what some authors mean by "incorrigibility." It's useful to briefly disentangle them. Frank Jackson discusses the view that "it is logically impossible to be mistaken about certain of one's current mental states" under the guise of an "incorrigibility thesis" (Jackson 1973, 51). This is an obviously related notion, but one that comes apart from my definition of infallibility. Nothing in my proposal centers on an agent's own mental states, and, as for Alston, there is be a difference between one's ability to be mistaken *about* a proposition (by believing *either* that it is true *or* that it is false) and one's inability to be mistaken in believing of a particular proposition *that it is true*.

Sydney Shoemaker's definition of incorrigibility is closer in that it is conditional upon a kind of affirmation of the truth of a proposition: "If a person sincerely asserts such a statement it does not make sense to suppose, and nothing could be accepted as showing, that he is mistaken" (Shoemaker 1963, 215). Shoemakerian incorrigibility is a kind of

---

[2] Brown (2013) has Williamson's E=K thesis particularly in mind.

public infallibility. If a person believes an infallible proposition, their (sincerely believed) assertions of the same will be incorrigible in Shoemaker's sense—after all, no proposition that is true can be shown to be mistaken. But not all (Shoemakerian) incorrigible propositions are infallible, for a proposition might be incapable of being disproved but false, nonetheless.

There may, however, be a sense in which my infallibility is a kind of incorrigibility. Incorrigibility is (according to one way of thinking) the inability to be corrected. An agent might be uncorrectable with respect to a proposition because (a) they are unable to change their mind (or, at any rate, to have their mind persuasively changed from the outside) or (b) any change of mind would not constitute a correction, i.e. because the belief in question was already true. Any belief in an infallible proposition is incorrigible in the second sense.

3. On Behalf of the Doctrine of Infallibility

But before defending the surprising conclusion that the doctrine of infallibility is false, let's give the doctrine its due. The doctrine is deeply intuitive and for good reason. First, the doctrine has historical pedigree. It is to infallibility that Descartes appeals at the climax of the *cogito*:

> [Suppose] there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me. …[L]et him deceive me as much as he can, he will never bring it about that I am nothing so long as I think that I am something. …I must finally conclude that this proposition, I am, I exist, *is necessarily true whenever it is put forward by me*[3] or conceived in my mind. (Descartes 1984, 7)

---

[3] I take *believing* a proposition to be at least one way of *putting it forward*.

It's the infallibility of "I exist" that makes the wheels of the *cogito* turn and—on at least one way of reading Descartes[4]—makes it the first item to fully survive the scrutiny of Descartes' method of doubt.

The *cogito* illustrates the benefit of having a principle whereby guaranteed truth is *sufficient* all by itself to license belief. If the doctrine of infallibility were false, that would mean that Descartes had not yet done enough (at this point in the *Meditations*, at least) to show that we are permitted to believe that we exist! Descartes is often accused of making the standard for defeating skepticism too high: rarely has he been accused of making the standard too low.

Contemporary epistemologists continue to appeal to Cartesian infallibility at crucial junctures. Ernest Sosa, for instance, appeals to infallibility to ward off dream skepticism. Having argued that in dreams "we do not really believe; we only make-believe" (Sosa 2007, 8), Sosa claims that if we really believe that we are dreaming we must not be dreaming: If we were dreaming, our "belief" wouldn't really be a belief at all, but only a make-belief. We should affirm (and neither suspend nor deny) that <I am awake> "since only about that option [i.e. affirming] is it obvious to me now that *if I take it I will be right*" (Sosa 2007, 19).[5]

A second reason for the doctrine of infallibility emerges from a particular picture of epistemic value. "Believe truth! Shun error!" says Williams James (1907, 18).

---

[4] There's an important exegetical question about how Descartes' notion of infallibility employed in the *cogito* relates to the clarity and distinctness criterion that takes center stage in Meditation III. I am grateful to Christopher Frugé and Ram Neta for excellent discussion on alternative interpretations of Descartes, which unfortunately cannot be captured in a footnote.

[5] Cf. Wittgenstein: "The argument 'I may be dreaming' is senseless for this reason: if I am dreaming, this remark is being dreamed as well—and indeed it is also being dreamed that these words have any meaning." (1969, 383). An important difference is that, for Wittgenstein, the belief "I am dreaming" can only be true or senseless whereas for Sosa it can only be true or not a belief at all.

Epistemically, "these are our first and great commandments" (17),[6] and—a more

ambitiously reductive epistemologist might have added—the only ones.

James's aphorisms are suggestive of one of the main characters in our dialectic:

epistemic veritism (cf. Goldman 1999, 5). According to veritism (as used here), the only

fundamental epistemic value is accuracy.[7] Whatever else can be said for and against this

view, it is attractively simple. James's aphorisms are telic: believing truth and shunning

falsehood are goals that episequic agents ought to promote.[8] James notes that an agent

might emphasize one of these goals more than the other. A cautious believer might be

hesitant to believe when there is even a small chance of error (cf. Kelly, 2014). Infallible

beliefs, however, are such that there is some chance they will lead to true belief and no

chance that they will lead to false belief. If the only—or, at any rate, the most

fundamental—epistemic goods are believing truly and avoiding false belief, then there's

seemingly always a good decision-theoretic reason to take a chance on infallible beliefs:

There's an opportunity (indeed a certainty!) to gain an epistemic good without any risk of

epistemic harm.

One could accept the doctrine of infallibility without being a veritist of this stripe,

but veritists should be especially attracted to the doctrine of infallibility. Adding a true

belief *always* improves accuracy, at least if one can do so without adding any false beliefs

or removing true ones.[9] And there's no obvious reason why believing an infallible

---

[6] Intriguingly, James himself fluctuates between *believing* the truth (1907, 18) and *knowing* the truth (17) as the positive epistemic commandment. In this paper, I shall represent James as endorsing the commandment to believe the truth, although this is a simplifying, historical fiction.

[7] Cf. Goldman: "[T]rue belief is the ultimate value in the epistemic sphere" (2001, 32).

[8] For a non-telic, or at least anti-consequentialist, version of veritism that falls outside the sort targeted in this paper, see Sylvan (2018).

[9] Berker (2013a), building on Firth (1981), argues against veritism on the grounds that it permits wrongheaded tradeoffs, allowing agents to believe obvious falsehoods to gain true beliefs downstream: veritists ignore "the epistemic separateness of propositions" (2013a, 365). Since the cases in this paper that

proposition would require one to also acquire false or abandon true beliefs. Indeed, the very plausibility of the doctrine of infallibility is likely to be seen as an argument for veritism. If guaranteed truth is by itself sufficient to license belief, a good explanation is that accuracy is what matters most in epistemology.

Third, the doctrine of infallibility, or at least a principle that entails it, explains the right verdict in certain tricky cases that reverse the normal causal direction of fit between mind and world. Note that the doctrine of infallibility is a weak version of a family of principles that permit agents to believe when they know that their belief would have some truth-oriented property or other if believed. In particular, the

> **Doctrine of Infallibility**: If S knows that <necessarily, if she herself were now to believe that *p* then she would truly believe that *p*> then it is thereby (rationally) permissible for S now to believe that *p*.

is entailed by the

> **Doctrine of Truth:** If S knows that <if she herself were now to believe that *p* then she would truly believe that *p*> then it is thereby (rationally) permissible for S now to believe that *p*.

For if S is permitted to believe *p* in virtue of knowing that she *wouldn't* be wrong about *p*, then she is surely permitted to believe *p* in virtue of knowing that she *couldn't* be wrong about *p*.

The doctrine of truth echoes Velleman's (1989a/2000) articulation of epistemic freedom, in which he argues that one is "entitled to say," and, Velleman suggests elsewhere, *believe*, what "wouldn't be false if [one] said it" (Velleman 2000, 40).[10] It turns out that the doctrine of truth is extremely useful in explaining why we are permitted

---

might cause trouble for veritism don't depend on tradeoffs of the relevant sort, I will stay neutral on whether veritists are committed to permitting them.

[10] Elsewhere (2000, 40), Velleman suggests that the agent must have *evidence* that they wouldn't be wrong about *p*, bringing Velleman's principle even closer to my formulation of the doctrine of truth.

to believe certain propositions when the ordinary direction of fit between mind and world is reversed. Indeed, several authors including Velleman (2000, 40, 44), Reisner (2013, §2), Kopec (2015, 404), Raleigh (2017, 332–333), Drake (2017, 4901) and Dahlback (forthcoming) appeal, at least implicitly, to something like the doctrine of truth in order to make sense of such cases.

Here's a representative case from Dahlback (forthcoming). A friendly demon guarantees that the result of a coin flip will match your belief about whether it is heads or tails. Dahlback reasons that so long as we know that we are in such a situation, we are permitted either to believe that the coin will land heads or that it will land tails (and so not heads) since we'd know that our belief was correct. In this case, the doctrine of truth seems to yield the proper result. It does seem permissible to believe either that the coin will land heads or that it will land tails.

Importantly, this seems right even though the evidence favors neither the thesis that the coin will land heads nor that it will land tails: the favoring evidence runs out. By "favoring evidence," I mean evidence that favors one attitude or contiguous range of attitudes over its competitors. Normally when the evidence favors neither p nor not-p we are required to suspend judgment. But in this case, to suspend judgment would to "willingly cast aside the promise of truth" (Drake 2017, 4902). One can't *follow* the evidence to the conclusion that the coin will land heads or tails—one simply believes and thereby makes oneself right. That the antecedent evidence favors neither thesis is important: It makes clear the role that the doctrine of truth plays (or seems to) in explaining the permissibility of believing either that the coin will land heads or that it will land tails.

As Dahlback notes, an important feature of the case is that the ordinary causal direction of fit between mind and world is reversed: the belief is a kind of self-fulfilling prophecy. We are guaranteed to be right not because our mind is tracking the world but because the world is tracking our mind. That the doctrine of truth can deliver the right verdict in such cases—when the favoring evidence seems to run out—is a strong point in its favor. Considered broadly, such cases fit neatly with the reductive gloss on James sketched earlier. True and false beliefs are what matter. So long as you knew you'd be right, who cares (our imagined Jamesian shrugs) whether how you got there was by following the evidence? After all, one's self-fulfilling beliefs and one's evidence remain "subjunctively linked" insofar as the belief "creates adequate evidence for it[self]" (Foley 1991, 102). And if the goal of following the evidence is to find the truth, then we may ask with Velleman (1989b, 63): "Why would rules [to follow the evidence] designed to help one arrive at the truth forbid one to form a belief that would be true?" What's important isn't how you get there but that you knew you'd be right at journey's end.

I've suggested that the ability of the doctrine of infallibility to explain why we are permitted to believe that the coin will land heads is a point in its favor. But some disagree: some think that we are *not* epistemically entitled to believe that the coin will land heads (although we may, for instance, be *pragmatically* entitled to *form* the belief that the coin will land heads). There's an important debate to be had here about the right and wrong kinds of reasons for belief (see Kavka 1983; Resiner 2009; Schroeder 2012) and whether we aim sometimes to have true beliefs or only to believe what is true (Antill 2020). Intuitions on this subject can be hard to leverage. For instance, Kopec (2015), Raleigh (2017), Drake (2017) and Dahlback (forthcoming) use cases of self-fulfilling

belief to argue for a robust permissivism whereas Antill (2020) argues against interpretations like theirs (in part) precisely *because* they lead to robust forms of permissivism.

Philosophically, there's much left to resolve. Dialectically, however, we may sidestep this issue. Those who insist that we don't properly respect our evidence (because we don't *follow* it) in cases of self-fulfilling beliefs will *already* be skeptical of the doctrines of truth and infallibility. For the doctrine of infallibility does not require that agents have beliefs that are arrived at by following the evidence: it allows beliefs that they merely know they are guaranteed not to be wrong about. And, in cases of self-fulfilling belief, these criteria can come apart.

I think, however, that the defender of the doctrine of infallibility is right to take our intuitive permission to believe either way in certain cases of known-to-be self-fulfilling beliefs as evidence for their view. One further point in favor of this interpretation, articulated in Raleigh (2017, 338–339), is that the corresponding principle according to which we are *not* permitted to have beliefs that are guaranteed to be *false* is also highly plausible. For instance, no one should believe <*p*, but I don't believe that *p*> (cf. Raleigh 2017, 329; Reisner 2014, 482). One doesn't (necessarily) decide not to believe this proposition by weighing the evidence for it: one can simply decide to reject it on the grounds that one can only believe it falsely.

In any case, I will ultimately argue that the doctrine of infallibility is false because there are *particular* (known-to-be) self-fulfilling judgments that we are not epistemically

entitled to make. But my argument won't depend on any qualms about the propriety of self-fulfilling beliefs in general.[11]

4. The Problem of Easy Downgrade

So, there's much to be said for the doctrine of infallibility. It anchors prominent anti-skeptical arguments, encourages us to take smart epistemic bets, and helps explain some otherwise tricky cases when the evidence follows our beliefs rather than our beliefs the evidence.

Nevertheless, the doctrine of infallibility is false. Although we usually make our overall epistemic state better by believing in such a way that we couldn't be wrong, we can also make our overall epistemic state worse.

The doctrine of infallibility is false because it makes it too easy to permissibly acquire defeaters for our beliefs (or at least too easy to rationally downgrade them). A belief is defeated by another belief, in the stipulative sense of this paper, when the second belief makes the first lose some positive epistemic status. Suppose I believe it is noon but then learn that the clock I had based my belief on runs an hour late. My belief that the clock is running an hour late defeats the justification for my belief that it is noon. Losing justification is one way of losing a positive epistemic status; so, the belief has been (in our sense) defeated. (Note that our sense of "defeat" is intentionally broader than those that require loss of some *particular* epistemic quality like justification or knowledge.)

It's controversial just when defeat happens. To get a grip on the problem for the doctrine of infallibility, let's begin by considering a very permissive defeat principle.

---

[11] See Berker (2013b, 376–377) for an argument against veritism that *does* rest on qualms about the propriety of self-fulfilling beliefs in general. Notably, however, Berker explicitly restricts his case to beliefs that are *not* known to be self-fulfilling.

Although many (the author included) will find this first-pass defeat principle ultimately unconvincing, it will help us to identify a recipe for finding infallible propositions that ought not be believed. This recipe will give us a strategy for cooking up counterexamples to the doctrine of infallibility that can be adjusted for philosophical taste. Here is the first principle:

> **Easy Irrationality**: Necessarily, it is irrational for S to believe that *p* if S believes that it is irrational for S to believe that *p*.

The principle **Easy Irrationality** has some plausibility. Suppose S believes that it is irrational for her to believe some proposition, but she believes it anyway. The agent apparently displays a lack of appropriate epistemic reflection. She believes she has no good reason for believing *p* and believes it anyway. Something seems to have gone wrong epistemically.

Easy Irrationality is thus one avatar (though not the only one) of our second character: *reflectivism*. It's one way of expressing the intuition that one's reflective attitudes about one's first-order beliefs make at least some difference to the epistemic quality of those first-order beliefs.

Let's not worry too much about whether this principle survives scrutiny.[12] What I want to argue here is that *if* **Easy Irrationality** is true then there are certain infallible propositions that should not be believed.

Suppose an agent knows both *p* and **Easy Irrationality**. They consider whether they may believe, in addition, that it is irrational for them to believe that *p*. They ask themselves, "Suppose I were to believe that it is irrational for me to believe that *p*. Would that belief be true?"

---

[12] See Coates (2012) for a critique.

Absolutely! For according to **Easy Irrationality**, believing that it's irrational to believe that p is enough to make believing that p irrational. Simply having the belief makes it so. More than that, since **Easy Irrationality** is a necessary truth, it's impossible that the belief could be false. The agent knows that the belief <it is irrational for me to believe that p> is infallible for them. The agent knows that, necessarily, if she herself were to believe <it is irrational for me to believe that p> then she would truly believe <it is irrational for me to believe that p>.

But obviously, it's wrong to believe that a belief is irrational just because **Easy Irrationality** makes that higher-order belief infallible. This would lead an agent to have a worse set of beliefs overall if p was otherwise rational to believe (and if there wasn't independent reason to doubt that it was rational to believe that p). The first-order belief could become needlessly irrational.

This is most clear in the case in which, antecedent to forming the easily-irrationalizing belief, the agent had been in a position to know and rationally believe both <p> and <it is *rational* for me to believe that p>. The agent then has (at least) two choices:

A. Believe <it is rational for me to believe that p> and believe <p>.

B. Believe <It is irrational for me to believe p> and believe <p>.

In this situation, the agent should clearly choose A over B. For only by choosing A will the agent emerge with two beliefs that are both true and rational. For if **Easy Irrationality** is true, then choosing option B will result in an irrational belief that p.

But even if the agent hadn't been in a position to know that their first-order belief was rational, it seems wrong to believe <it is irrational for me to believe that p> *merely*

because one would be right. Doing so takes too cavalier an attitude toward the possibility of downgrading the rationality of one's first-order beliefs. One is making it irrational to believe something without any evidence that one's epistemic situation forces this undesirable outcome. Something seems to have gone wrong.

Maybe what's gone wrong is **Easy Irrationality**. After all, many epistemologists think defeat is hard to come by. Lasonen-Aarnio (2010), Coates (2012), Williamson (2014), and Weatherson (manuscript), for instance, all argue against **Easy Irrationality** or analogues of it. Nevertheless, thinking about **Easy Irrationality** was valuable, for it has given us a template for thinking about how certain higher-order beliefs could in principle be infallible. Consider the following infallibility recipe:

> **Infallibility Recipe**: Necessarily, if S believes that it is F for S to believe that p then it is F for S to believe that p.

Easy Irrationality is true just in case we can plug in "irrational" as the ingredient for "F." As noted, if defeat is hard then simply believing that a belief is irrational might not be enough to make the first-order belief irrational. But it still seems that negative higher-order epistemic appraisals make their corresponding first-order beliefs worse in some way, even if it doesn't always make them irrational. Our task is to find some F that captures whatever way it is that first-order beliefs become worse upon receiving negative higher-order appraisals.

There are four broad ways that one might argue that no negative, epistemic property satisfies the infallibility template. First, one might endorse *extreme level-splitting*,[13] the view that the quality of our belief systems is not at all impacted by the relationship between our lower- and higher-order beliefs.

---

[13] I borrow this term from Horowitz (2014), although my use is more restrictive.

Extreme level-splitting seems unduly strong. If we think it's a total disaster to believe that p but believe p anyway, surely that lowers the quality of our first-order belief or our belief system in *some* way. If extreme level-splitting is true, then we can completely ignore our higher-order beliefs when evaluating their first-order counterparts. That stretches credulity. Surely there are better ways to argue that nothing satisfies the infallibility template. In other words, endorsing extreme level-splitting violates the intuitive thesis we've called epistemic *reflectivism*: our reflective beliefs about first-order beliefs have at least *some* impact on the quality of our corresponding first-order beliefs or belief system.

I take the falsity of extreme level-splitting as a datum. But even those who embrace a degree of level-splitting don't explicitly endorse the extreme thesis that second-order beliefs have *no* effect on the quality of first-order beliefs of belief systems. Coates (2012) argues that agents can rationally believe that their belief that p is *irrational* while, nevertheless, *rationally* believing that p. This view doesn't entail extreme level-splitting though, since it's consistent with all this that believing that it's irrational to believe p makes one's belief that p worse in some way even if it doesn't make it flat-out irrational. Similar observations apply to Weatherson's claim that "what we should believe can come apart from what we should believe that we should believe" (manuscript) and Williamson's (2014) view that one can know that p while knowing it is improbable that one knows that p.

These views are suggestive, however, of a second strategy for arguing that the infallibility template is never truly instantiated: One could argue that although negative higher order beliefs affect first-order beliefs in some way, no negative property is such

that believing a belief has that property automatically makes the corresponding first-order belief bad *in that very way*. After all, we can be mistaken—even rationally mistaken—in our first-order beliefs. Why should our second-order beliefs be any different? On this picture, believing that one's beliefs are bad in an F-ish way does indeed make them worse (extreme level-splitting is false), but one's first-order beliefs may be made worse in a G-ish way rather than an F-ish way.

This strategy is consonant with the view of defeat articulated by Lasonen-Aarnio (2010). Lasonen-Aarnio argues persuasively that we shouldn't confuse our evidence being such that it's *unlikely* that we know that *p* with our actually not knowing that *p* (2010, 10). An agent might have evidence that their visual capacities are misfiring, but if their visual capacities are operating well, and if the agent bases their belief solely on their visual capacities, it might be that "being stubborn pays off" (2). If this is right, agents who stubbornly believe what they are unlikely (on their evidence) to know may nonetheless emerge with full knowledge.

But Lasonen-Aarnio is equally emphatic that we can, nevertheless, genuinely criticize agents who believe against the evidence—they are being *unreasonable*. After all, they are risking a lot (epistemically-speaking) by forming a belief that, according to their evidence, is very unlikely to constitute knowledge. It's just that they might not be criticizable in the way that we first thought. They still get to count as knowers, but *unreasonable* knowers.

Indeed, if Lasonen-Aarnio could not explain why agents who believed (and thereby came to know) in the face of significant (but ignored) counter-evidence were criticizable in *some way or other*, that would speak against her account. It adds

significantly to the plausibility of the overall picture that she does not give the unreasonable knower uniformly positive marks. We need to be cautious about saying *how* higher-order criticisms negatively impact first-order attitudes—but that doesn't give us reason to doubt *that* they do so.

It's important that the most plausible views according to which defeat is difficult nevertheless preserve ways to criticize agents whose first-order beliefs are in tension with their higher-order ones. For it will allow us to fill in the infallibility recipe by going general. Our second strategy for avoiding the problem of easy defeat said that negative higher-order epistemic appraisals make their first-order counterparts worse in some way, just not automatically in the way we believed them to be worse. To counter this strategy, we can move to a principle that uses a sufficiently general negative, higher-order appraisal. Consider:

> **Easy Problems**: Necessarily, if S believes that it is problematic for S to believe that *p* then it is problematic for S to believe that *p*.[14]

"Problematic" is such a general, negative term that *any* way of making a first-order belief worse counts as problematic. So, it seems that **Easy Problems** is true even if more specific principles like **Easy Irrationality** are not.

Unless our third strategy for resisting the infallibility recipe succeeds. Our second strategy was to argue that although believing that one's beliefs are bad in an F-ish way makes them worse, it always makes them worse in a G-ish way and not an F-ish way. Our third strategy admits that there is some sufficiently general F such that believing a belief to be F makes it worse in an F-ish way. But it denies that the first-order belief must be bad enough to make it F outright. So, for instance, the defender of the third strategy

---

[14] I'm grateful to Laura Callahan and Ernie Sosa for conversation that led to this version of the principle.

insists that although believing that it is problematic to believe *p* entails that it is *more* problematic to believe that *p* than it might have been otherwise, it doesn't make believing *p* problematic outright. It treats "problematic"—or any F that might otherwise satisfy the infallibility template—as a threshold term, such that being more F does not entail being F (just as being *taller* than something doesn't entail being *tall*).

It's unclear whether this strategy is properly motivated—it's far from obvious that "problematic" is relevantly like "tall." But instead of pressing this line, we'll look for a term that avoids the objection altogether: an F such that being more F (than whatever) entails its being F outright.

We can do this by stipulating new terminology that does not operate with this sort of threshold. Let's introduce the term "besmirched": An agent's belief is *besmirched* just in case a belief is (epistemically) problematic *to any degree*. "Besmirched" maintains the generality of "problematic" but, by stipulation, is not a threshold term.

One need not be terribly worried to discover that one has a besmirched belief—in certain epistemic circumstances, one probably *should* believe that one has besmirched beliefs (as may also be true for "irrational" and "problematic"). Nevertheless, there is something unfortunate about such beliefs. The best of the best beliefs are unaccompanied by *any* negative higher-order epistemic appraisals—even slight ones. And believing that a belief is besmirched is one way to have such a negative higher-order epistemic appraisal. Accordingly, the following thesis is true:

> **Easy Besmirchment**: Necessarily, if S believes that it is besmirched for S to believe that *p* then it is besmirched for S to believe that *p*.

The slightest stain besmirches: being more besmirched (than whatever) entails being besmirched outright. So, **Easy Besmirchment** escapes our third objection.

Suppose an agent knows that *p* and knows that **Easy Besmirchment** is true. They consider whether to believe, in addition, that it's besmirched for them to believe that *p*. They ask themselves, "Suppose I were to believe that it is besmirched for me to believe that *p*. Would that belief be true?"

Absolutely! For according to **Easy Besmirchment**, believing that it's besmirched to believe that *p* is enough to make believing that *p* besmirched. Simply having the belief makes it so. And since **Easy Besmirchment** is a necessary truth, it's impossible that the belief could be false. The agent knows that, necessarily, if she herself were to believe <it is besmirched for me to believe that *p*> then she would believe so truly.

The doctrine of infallibility faces a problem: it insists that we are permitted to believe that our beliefs are besmirched just because believing it would make it so. But taking this path makes our total epistemic state worse even if we acquire true beliefs in the process. We're *not* rationally permitted to needlessly downgrade our beliefs in this way—not even lightly.

Let's summarize where we've come so far. Given that the quality of our first-order beliefs is at least somewhat impacted by our higher-order beliefs about them, it's hard to resist the conclusion that there are some infallible propositions such that (a) we can know that they are infallible for us, and yet (b) if we believe them, they will needlessly damage our (potential) first-order beliefs in a way that is not epistemically permissible.

One could, however, preserve the reflectivist intuition that the relationship between our beliefs and our higher-order assessments of those beliefs *matters* epistemically while denying that the way that it matters affects the quality of our first-

order *beliefs* themselves. One could insist that what is impacted by negative, higher-order epistemic appraisals is not (necessarily) the corresponding first-order belief but a *network* of beliefs, including at least the first-order belief along with the negative, higher-order assessment.

This is, in many ways, an attractive position. What's bad about believing both <p> and <it's problematic for me to believe that *p*>? Part of the answer is that the beliefs do not mesh together as well as they might. There's tension. When the agent has very good reason to believe both, the best choice might be to live with that tension, but there's tension just the same. The tension metaphor suggests that the problem is with how beliefs (or potential beliefs) fit together and not, in the first instance, with the beliefs themselves. This is at least suggestive of the view above that the thing damaged in easy downgrade cases is a belief system and not necessarily any particular belief. The damage done is wholly holistic.[15]

This seems to make trouble for the recipe for finding infallible propositions. Recall:

> **Infallibility Recipe**: Necessarily, if S believes that it is F for S to believe that *p* then it is F for S to believe that *p*.

The **Infallibility Recipe** prompts us to look for negative properties to ascribe to believing a single proposition. But if the sort of tension involved in the cases discussed so far is not a property of believing any single proposition but a property of a belief *network*, then the **Infallibility Recipe** has us looking in the wrong place.

---

[15] Matt McGrath and Ernie Sosa both helpfully pressed the importance of this alternative in conversation. This option evokes, for instance, the view expressed by Worsnip: "Coherence requirements are widescope, and do not speak in favour of individual attitudes simpliciter but rather against particular combinations of attitudes" (2018, 36–37).

Before jumping in to plug the hole, let's take a step back. We've already seen that we are not permitted to believe infallible propositions if so believing does needless doxastic damage to another (potential) belief. We are permitted—required, even—to forego the guarantee of truth when doing so protects the quality of our other beliefs in certain ways.

If the goal of safeguarding the epistemic quality of particular beliefs can require us not to believe certain infallible propositions, the goal of safeguarding the epistemic quality of our belief networks should serve just as well. Once we see that there are (at least) two potentially competitive goals in play—truth and maintaining other qualities (rationality, coherence, unproblematicness, etc.) of our beliefs or belief system—the idea that we can be automatically licensed to believe in virtue of being guaranteed to attain just one of them starts to sound suspicious.

But let's turn to a specific example. Suppose I form the following belief: <it is problematic for my belief system to include p>. On the holistic picture, so believing would *make* it so that there would be tension within my belief system if I also believed that p. That tension would not necessarily be a problem for my (potential) belief that p itself, but it *would* be problematic for my belief *network*. But of course, that is exactly what I believed in the first place: that it is problematic for my belief system to include p. It seems, then, that the following principle is true:

> **Easy Systemic Problems**: Necessarily, if S believes that it is problematic for S's belief system to include p then it is problematic for S's belief system to include p.

And, of course, **Easy Systemic Besmirchment** is at the ready if concerns about thresholds rise again.

Here, in short, is the problem: If reflectivism is true, then acquiring certain higher-order beliefs can make either our first-order beliefs or else our belief systems worse in some way (without making them false). So, we *shouldn't* form such beliefs when doing so can be easily avoided. But the doctrine of infallibility only cares about accuracy. It wrongly predicts that we *are* permitted to form certain problematic higher-order beliefs anyway, just because we're guaranteed to be right about them.

One could be complacent about this result. Beliefs and belief systems are easily besmirched. And we're far from ideal agents. It might turn out that the vast majority of our beliefs are already besmirched whether we believe that they are or not. And if so, why worry that the doctrine of infallibility permits us to believe that our beliefs are stained in precisely the way they already are?

But such complacency is unmerited. First, if the doctrine of infallibility were true, then even perfect knowers—oracles, supercomputers, gods—would be permitted to believe (truly, once believed) that their beliefs (or belief systems) were besmirched. But surely such powerful agents would not have antecedently besmirched beliefs, even if we mere mortals often do. They certainly shouldn't downgrade their beliefs so needlessly.

But second, when one believes, e.g., that <it is problematic to believe that *p*> solely because that proposition is infallible, the proposition is not made permissible *thereby*. The doctrine of infallibility is supposed to be an explanatory thesis: S's knowing that <necessarily, if she herself were to believe that *p* then she would truly believe that *p*> explains why S is permitted to believe *p*. Guaranteed accuracy explains permissibility. But if the doctrine of infallibility only comes out true because we are permitted to believe some infallible propositions for reasons unrelated to their guaranteed truth—because we

epistemically frail creatures have antecedently besmirched beliefs—then infallibility does not play the explanatory role that we first thought. The problem is not that we shouldn't believe our beliefs are besmirched (for all I've said, epistemic humility nearly always demands this of us!) but that we clearly shouldn't do so *just because* we'd be guaranteed to be right. Either way, the chain linking guaranteed truth with permission is severed.

5. Direction of Fit Solutions

The problem of easy downgrade shows that the doctrine of infallibility is flawed. Given how unassailable the doctrine appeared at the start, this itself is a significant conclusion. But those sympathetic to the doctrine may hope that its flaws can be mended—or at least safely ignored—in most contexts. In particular, it's noteworthy that the counterexample involves a reversal of the ordinary direction of fit between mind and world. Our aim in believing is (at least in part) to form a representation in our mind that appropriately matches the outside world—to tailor our minds to fit the world. Ordinarily, the world cares little how our mind represents it. But not always. Sometimes the world tailors itself to fit our representation. And indeed, if any of the defeat principles proposed are true, the quality of our epistemic states depends in some part on how we represent those states to ourselves. Perhaps, then, even though the doctrine of infallibility is false, we can treat the doctrine as true when there's no reversal of the ordinary direction of fit between mind and world—when the way that the world is (in the domain of our proposed belief) does not depend on our beliefs themselves. We replace the original doctrine with this revised principle:

> Doctrine of Infallibility, Dependence Edition: If the truth-value of S's belief that p would (if formed) not depend on S's belief that p itself, *then* if S knows that <necessarily, if she herself were to believe that p then she would truly believe that p>, then it is thereby (rationally) permissible for S now to believe that p.

The ambition behind this strategy is a sensible one. It seems that there is something right about Descartes's appeal to infallibility in the *Meditations*, and given that the doctrine of infallibility is false, epistemologists should be eager to find a more restrictive principle that allows the Cartesian inference while avoiding the problem of easy downgrade. And since the problematic cases of easy downgrade *are* ones in which the downgrading beliefs explain their own truth, the revised principle is a tempting tactical retreat.

But the problems with the doctrine cannot be excised merely by restricting its domain to cases in which beliefs play no role in their own truth. For we seem to need the doctrine of infallibility (or something that entails it) when dealing with other cases with a reversal of dependence. The revised principle does not neatly divide the good cases from the bad.

Consider again the *cogito*.[16] Before Descartes concludes <I exist>, he concludes <I am thinking>. But <I am thinking> is the sort of proposition whose truth is (at least partially) grounded in the belief itself. If I believe <I am thinking>, my very belief grounds the truth of the believed proposition. Moreover, it is permissible to believe <I am thinking> even if <I am thinking> is the only thought one is having at the moment, so the belief could even be the full grounds for the truth of the proposition. The belief <I am thinking> *depends*—at least partially and potentially fully—on itself for its truth. But <I am thinking> is a paradigmatically good infallible proposition! The dependence edition

---

[16] I'm grateful to Ezra Rubenstein for suggesting that I inquire into whether a restriction on beliefs that ground their own truth would solve the problem and to Ernie Sosa for pointing out the relevance of Descartes' <I am thinking>.

of the doctrine of infallibility is thus overly restrictive: it does not succeed in neatly

distinguishing the good infallible propositions from the bad.

Perhaps, the objector rejoins, this is because we were operating with too wide a

notion of dependence. After all, in the case of <I am thinking>, the belief <I am

thinking> *grounds* (or perhaps *constitutes*) the truth of the believed content. But

grounding isn't the only kind of dependence out there. Perhaps *causal* dependence is the

problematic kind:

> Doctrine of Infallibility, Causal Dependence Edition: If the truth-value of S's
> belief that *p* would (if formed) not causally depend on S's belief that *p* itself, *then*
> if S knows that <necessarily, if she herself were to believe that *p* then she would
> truly believe that *p*>, then it is thereby (rationally) permissible for S now to
> believe that *p*.

I think, however, that we should be suspicious of moves to restrict the epistemically

relevant sort of dependence in this way. It's clear why *guaranteeing* the truth is

epistemically relevant. But it's not clear why I should care, epistemically speaking, about

whether the source of that guarantee is causal or constitutive. A guarantee is a guarantee.

Neither kind is less likely to be true than the other.

Moreover, some of the cases that motivate the doctrine of infallibility in the first

place involve causal dependence on a belief. Recall the coin-flip case. In these cases, the

agent knows that they will be right whether they believe that the coin will land heads or

that the coin will land tails, and the fact that the agent knows this seems to license belief

in either proposition. As noted earlier, the most straightforward explanation of this seems

to be the doctrine of truth, which entails the doctrine of infallibility. But the coin-flip

cases themselves have a reversed causal dependence! The agent's belief—through the

demon's intervention—causes the world to match the belief. The correct judgment in the

coin-flip case is not, of course, uncontestable.[17] But giving up the permissibility of believing either way in such cases does sap one of the main arguments in behalf of a doctrine of infallibility of its strength.

And even if one goes in for the version of the Doctrine of Infallibility that excludes cases in which the truth of one's belief causally depends on the belief itself, it's unclear that this avoids the problem. When I believe <it is besmirched for me to believe that p>, does this *cause* it to be besmirched for me to believe that p? Certainly, it's nothing like how the throw of a rock causes the breaking of a window. It's not a relation between events. Nor is it like the coin-flip case, in which a demon causally interferes in the world to guarantee a certain outcome. Rather, there is a conceptual link between my believing <it is besmirched for me to believe that p> and the belief's being true: it is a consequence of, among other things, the way "besmirched" was defined.

I do not take myself to have shown that there is no possible variety of dependence that can divide the good infallible propositions from the bad—that can count as licit the inferences in the *cogito* and the coin-flip case while excluding the inferences central to the problem of easy defeat. But I hope I have cast doubt on the idea that the doctrine can be easily bandaged by invoking a simple distinction between beliefs that do and do not explain their own truth.

The objector hoped to show that there is something odd or deviant about beliefs that depend on themselves for their truth. It's not hard to enter this frame of mind. Since Anscombe (1957), philosophers have thought that one of the distinguishing features of belief as a mental state is that it has a world-to-mind direction of fit. That there is coffee

---

[17] See, for instance, discussion in Antill (2020)

in front of me (in the world) is a reason for me to *believe* (in my mind) that there is coffee in front of me, whereas my *desire* (in my mind) that there is coffee is a reason for me to make it true that there is a cup of coffee in front of me (in the world). Platt would later argue that "beliefs should be changed to fit with the world, not vice versa" (1979, 257).[18] Certainly, it'd be very strange to form the belief that there is coffee in front of me and then, because I have that belief, to make a cup of coffee so as to make that belief true.

But I don't think we should extend our suspicion of this strange behavior to all cases in which the relationship between a belief and the truth of its content are entangled. Even in the coin-flip case, we can say that, were the result of the coin-flip different than I had believed—because, say, the demon had misread my mind—my belief would be *mistaken*. From my perspective at least (things may be different for the demon who *desires* that the world match my beliefs), the mistake is with my belief and not the world (cf. Anscombe 1957, 56). Similarly, Humberstone writes that even self-fulfilling prophecies "involve beliefs with the same direction of fit as any other beliefs, being appraised for correctness ...in terms of how well their content matches how things are with their subject matter" (Humberstone 1992, 71). So, although self-fulfilling beliefs have a different *causal* direction of fit than typical beliefs, their metaphysical role as representations of how the world is and their normative success conditions that depend on how the world is in fact remain unchanged. Once we see that beliefs whose truth-value is metaphysically entangled with being held can, nevertheless, bear the ordinary direction of fit between mind and world as other beliefs in *this* sense, the case to restrict the scope of the doctrine of infallibility to "ordinary" cases loses some of its urgency.

---

[18] Cf. Williams: "[A] man's word, and his beliefs, should *reflect* things as they are" (1966, 20, emphasis mine).

6. The Disbelief Solution

Perhaps we've been focusing too much on aiming for the good outcome of a true belief and not enough on the bad outcome of believing a falsehood. With this thought in mind, we recall that the *cogito* has not just one but two things going for it. First, as we've noted, if one believes that <I exist> then one is guaranteed to be right. But equally, if one believes its negation, <I don't exist>, then one is guaranteed to be wrong (Sosa 2007, 18; Shah 2009, 189).

Moreover, this distinguishes the *cogito* from the easy downgrade propositions of this paper. Notice that <it is besmirched for me to believe that *p*> is *not* such that if I *disbelieve* it, I am guaranteed to be wrong.

This suggests a new doctrine:

Doctrine of Infallibility, Disbelief Edition: If S knows that <necessarily, if she herself were to believe that *p* then she would truly believe that *p*> *and* S knows that <necessarily, if she herself were to believe that ¬*p*, then she would falsely believe that ¬*p*, then it is thereby (rationally) permissible for S now to believe that *p*.

This successfully saves the *cogito* without endorsing besmirching propositions. But much like the attempted fix by fit, it leaves unexplained why it's permissible to believe in cases when one knows one would be right either way. In Dahlback's case, I can believe that the coin will land heads—seemingly, just on the basis that I know I will be right—even though I will also be right if I *disbelieve* that the coin will land heads. Once again, this revised principle cannot neatly divide all cases of good infallible propositions from the bad.[19]

7. Concluding Thoughts

---

[19] I'm grateful to Ernie Sosa for conversation on this and related points.

We've encountered a puzzle. The doctrine of infallibility seems overwhelmingly plausible. It is the basis for the *cogito*, it makes sense of certain self-fulfilling prophecies, and it gives voice to the enticingly straightforward thought that accuracy is what matters most (epistemically) when deciding what to believe.

But the problem of easy defeat shows that the doctrine of infallibility is false. If reflectivism is true—if reflective attitudes about one's first-order beliefs make at least some difference to the epistemic quality of those first-order beliefs or one's belief system—then there are some guaranteed truths one should refrain from believing.

Does the falsity of the doctrine of infallibility lead to skepticism? Perhaps if one comes into the problem in a Cartesian mood. If one can't automatically trust even infallible propositions (the disillusioned Cartesian asks), what can we trust? But most epistemologists have (wisely) not demanded that some of our beliefs must be infallible to count as knowledge.

No, the primary puzzle is not, "How can we really know that we exist or know that the coin will land heads if the doctrine of infallibility is false?" We were (rightly) pre-theoretically confident that these were good judgments, and we need not abandon them just because the principle that we thought explained their permissibility turned out to be false. Rather, the puzzle is how to separate the good infallible propositions from the bad. We're left to wonder: why wasn't the guarantee of truth good enough to license belief? What was so valuable, epistemically, that it was worth foregoing a guaranteed true belief?

I see no easy answer to this question.

If this paper is right, we are left with a broken, false doctrine that had seemed foundational to Cartesian epistemology and left without an obvious way to repair it. Is there anything positive we have learned?

I conclude by suggesting two modest lessons. First, we learn that there's tension between veritism and reflectivism. Perhaps this shouldn't surprise us so much in the end. Veritism (as we are using the term) says that what matters most at bottom is accuracy. Reflectivism says that something else matters too: namely, how our first-order beliefs and higher-order beliefs fit together. Nevertheless, one might have thought that the reason it matters how our different levels of beliefs fit together is because such relationships can help us to become more accurate. But although this might be part of the story, it can't be the whole story. For believing easy-downgrade propositions on the basis of their infallibility guarantees accuracy while damaging reflective fit.

Second, we have found a surprising argument against Cartesian infallibilism. Cartesian infallibility (and the certainty it engenders) is often taken to be too stringent a requirement for either knowledge or proper belief. But if Cartesian infallibility seemed extreme, it at least also seemed like a natural stopping point. What more could one hope once infallibility had been achieved? What greater epistemic assurance? If infallibility seemed too stringent to be necessary for permissible belief, it at least seemed obviously sufficient for it.

But we've learned that infallibility isn't sufficient for permissible belief: sometimes, one shouldn't believe even when one knows one would be right. Fallibilists—who already believed that infallibility was not *necessary* for right belief—may feel justly

emboldened knowing that it isn't *sufficient* for permissible belief either. It isn't the natural cut-off point in epistemic normativity that we've been led to believe.

Infallibility isn't always worth having even in those rare cases when it is there to be had.[20]

BEING IN A POSITION TO KNOW IS THE NORM OF ASSERTION

Abstract: This paper defends a new norm of assertion: Assert that *p* only if you are in a position to know that *p*. We test the norm by judging its performance in explaining three phenomena that appear jointly inexplicable at first: Moorean paradoxes, lottery propositions, and selfless assertions. The norm succeeds by tethering unassertability to unknowability while simultaneously untethering belief from assertion. The PtK-Norm foregrounds the public nature of assertion as a practice that can be other-regarding, allowing asserters to act in the best interests of their audience when psychological pressures would otherwise prevent them from communicating the knowable truth.

1. Introduction

What is the norm of assertion? Many have thought that, whatever it is, it is something epistemic.[1] We assert, at least in part, to communicate *information* about the world: we add a proposition to the common ground (Stalnaker 1978) or produce an attitude or reaction toward a proposition in an audience (Grice 1969). Plausibly, in order for the speaker to be licensed in offering a proposition externally, they must stand in the right relation to the proposition internally. That is, the speaker must meet certain epistemic conditions with respect to the publicly asserted content. I argue that the relevant condition is being in a position to know:

---

[1] For a defense of the view that knowledge has a norm and that it is *something* epistemic, see Goldberg (2015a), especially chapters 1–3. This is not entirely uncontroversial, as this already rules out some proposals, notably Weiner's (2015) defense of the (semantic) truth norm. I take the epistemic character of the norm of assertion as a given, but see §3.6 for an argument that bears on Weiner's Gricean explanation of lottery propositions.

PtK-Norm: One should assert that *p* only if one is in a position to know that *p*.[2,3]

The Position to Know Norm faces competition from two accounts that have shaped the contemporary debate: Timothy Williamson's knowledge norm and Jennifer Lackey's reasonable to believe norm.

K-Norm: One must: assert *p* only if one knows *p*. (Williamson 2000, 243)[4]

RTB-Norm: One should assert that *p* only if it is reasonable for one to believe that *p*. (Lackey 2007, 608)[5]

Williamson and Lackey offer different answers to two, distinguishable questions. Question (1): What epistemic property normatively guides our assertions? Williamson answers knowledge; Lackey reasonability. Question (2): Must the doxastic state of the asserter exemplify this property in order to assert properly? Williamson answers yes; Lackey no. Whereas Williamson claims that an agent ought to actually know the asserted proposition, Lackey claims that an agent may assert a proposition though she does not reasonably believe it, so long as it is reasonable *for* her to believe it.

---

[2] Some philosophers argue that the norm of assertion is a *constitutive* norm—the sort of norm that makes assertion the thing that it is. Certainly, this is Williamson's (2000) thesis. But this constitutive claim is separable from the claim that proper assertions ought to meet an epistemic requirement. In arguing for a norm of assertion, I do not take myself to be committed to the view that the norm is *constitutive*, only that it is *binding*. Of course, those who endorse the K-Norm will think that the PtK-Norm is trivially binding, since knowledge entails being in a position to know. So, properly understood, the thesis of this paper is really that the PtK-Norm is binding *and* that it is the most foundational epistemic requirement on proper assertion. I'm indebted to John Phillips for helping me clarify this point.

[3] Some theorists argue that there is a default norm of assertion that can, nevertheless, be modified across different contexts, as in "conditions of epistemically diminished hope" (Goldberg 2015a, 285) or when hedged (Benton and van Elswyk 2020). My thesis should be understood as the claim that being in a position to know is the *default* norm of assertion.

[4] Cf. Unger (1975), DeRose (2002), Benton (2011, 2012, 2013, 2016), Sosa (2011). Turri's (2011) express knowledge norm and Goldberg's (2015a) nuanced thesis that knowledge is the default norm of assertion are important variants.

[5] Lackey's full formulation of the RTB-norm includes this second condition: "if one asserted that *p*, one would assert that *p* at least in part because it is reasonable for one to believe that *p*" (Lackey 2007, 608). But we'll return to this second condition in §5. Her official account is also augmented by a Not Misleading Norm that will be discussed in §3.

I contend that the *position to know* norm preserves the best features of both proposals. It gets the correct verdict in Lackey's cases of selfless assertion. But it also gets the correct verdict in the lottery case and the Moorean paradoxes that are favorable to Williamson. The PtK-Norm gets the cases right by siding with Williamson on our first question and with Lackey on our second. First, *knowledge* is the property that normatively guides our assertions. But second, the doxastic state of the asserter *need not exemplify* this property (knowledge) for the agent to assert properly. It preserves Williamson's insight that knowledge is central to assertion while accommodating Lackey's insight that "it is a mistake to require proper assertion to pass through the doxastic states of the asserter" (2007, 600).[6]

2. A Provisional Account of the Position to Know

The position to know is something like knowledge minus belief. Beyond this slogan, can we say more precisely what being in a position to know is? Here I offer a provisional account, intended to tell us enough to evaluate the PtK-Norm.

2.1 General Features of a Position to Know

Unlike knowledge and reasonability, the concept of being in a position to know is predominantly an epistemologist's term of art, finding scarce expression in ordinary parlance. We do sometimes say that someone "could have known" or "should have known," but it's unclear how closely these expressions line up with the epistemologist's usage. A parent could and should have known how their child's day was because they could and should have asked (at least if the child would answer!)—but not having asked,

---

[6] In personal conversation, Lackey confirms that—at least on the right interpretation of epistemic position—her norm could be read as "Assert that *p* only if you are in a position to reasonably believe that *p*." This formulation makes especially salient the structural similarities between our accounts.

it's doubtful they are in an epistemic position to know. This does not mean that we have no independent grip on what it takes to be in a position to know or that the concept is philosophically unimportant: it does mean the concept requires explication.

Being in a position to know is a modal notion. One is in a position to know *p* if one *could* know it under the right circumstances. One need not *actually* know to be in a *position* to know. So, in order to determine whether someone is in a position to know *p*, we look across modal space to see if there is a world in which someone knows *p* under the right circumstances. What circumstances are those? Those circumstances in which someone shares the agent's *epistemic position*.

This brings us to our second observation: When discussing what one is in a position to know, we are interested in the epistemic vantage point itself, not what use the agent has or is in fact psychologically capable of seeing from it. This is what it means to say that being in a position to know is knowledge *minus belief*—yet not just belief, but a broad range of facts about the agent's psychology. As a first pass (although important qualifications will be required), let us say that the position to know *depsychologizes* away from facts about the agent's particular mental states.

Suppose, for instance, that a hypochondriac is psychologically incapable of believing that he does not have the flu. Multiple tests indicate that he has no symptoms, and his doctor testifies that he is in full health. In addition to having perfectly good evidence that he does not have the flu, it is true that he does not have the flu, and his epistemic situation is free from chancy or Gettierizing factors that might obstruct knowledge. The agent—because of his pathology—is unable to know that he does not have the flu. He is, nevertheless, *in a position to know* since the locus of the concept's

evaluation is not the agent himself but the epistemic situation in which the agent finds himself. The epistemic position is what we are holding fixed in our evaluation. *Someone* in the agent's epistemic situation could know even if the agent himself could not: someone with a different psychological profile.

With these observations in place, we are now prepared to attempt a first pass at a definition of being in a position to know:

> **PtK**: S is in a position to know that *p* iff S could know that *p* given their actual epistemic position.

We hold fixed the epistemic position of the agent, but what we don't hold fixed is just as important: we are free to vary other facts about the agent's beliefs and psychology. Of course, we're often (physically, psychologically) capable of changing our epistemic position, e.g., by asking someone a question. But one isn't in a position to know by being able to ask a question. Rather, by asking a question, one can try to move from a state of not being in an epistemic position to know the answer to being in an epistemic position to know it.

The proposed definition is useful as far as it goes. It captures the modal (*could* know) and depsychologizing (allowing psychological variation) structure of the concept. But it leaves some central questions unanswered: What is an epistemic position? And what epistemic position does an agent have to be in to be able to know that *p*?

2.2 Constraints on Epistemic Position

Whatever answer we give must satisfy some basic constraints on an adequate theory of epistemic position. At least, any such theory must respect the following four principles:

PtK to Truth: If S is in a position to know that *p* then *p* is true.[7]

Knows to PtK: If S knows that *p* then S is in a position to know that *p*.

PtK to No Further Inquiry: If S is in a position to know that *p* then S does not require further inquiry to come to know that *p*.[8]

Anti-Collapse: For some S and for some *p*, it is possible that S is in a position to know that *p* but does not know that *p*.

The first constraint says that being in a position to know is factive.[9] The second says that knowing that *p* is sufficient for being in a position to know that *p*. Knowledge, like most properties, can manifest without being disposed to manifest: a vase must be *breakable* in order to break, but it need not be *fragile* (disposed to break). This rules out certain dispositionalist accounts of the position to know, whereby to be in a position to know one must be (robustly enough) disposed to know under certain circumstances.[10]

What exactly counts as inquiry can be, for our purposes, left open. But our third condition encodes the idea that agents who are in a position to know already have at their disposal the means to know the target proposition. I'm not in a position to know what the capital of Estonia is merely in virtue of the fact that I could easily inquire by consulting an encyclopedia (or Google).

The final principle says, in effect, that there is no general inference from an agent's not knowing *p* to the agent's not being in a position to know that *p*: it's possible (and almost always true) that what an agent is in a position to know extends beyond what she actually knows.

---

[7] Cf. Williamson: "Thus being in a position to know, like knowing and unlike being physically and psychologically capable of knowing, is factive" (2000, 95).

[8] Thanks to Ernie Sosa and Susanna Schellenberg for suggesting such a principle.

[9] Its factivity guarantees that the PtK-Norm satisfies what Ronald Jager calls *the initial datum*: "to assert something is somehow to incorporate the claim that what is asserted is true" (Jager 1970, 161).

[10] E.g. "Being in a position to know a proposition is to be disposed to acquire the knowledge that the proposition is true, when one entertains it on the right evidential basis" (Stanley 2008, 49).

There might be occasions in which one can infer from *this particular agent's* not knowing *this particular proposition* that the agent is not in a position to know it. A super-knower might know everything she is in a position to know. Propositions that are both self-evident (or near-enough) and attention-demanding (or what Chisholm 1982 called "self-presenting") may always be known when their subjects are in a position to know them. For instance, "I have a sharp headache" or "I have an experience like *this*" might always be known by those in a position to know them.[11] All our second principle says is that there is no general rule licensing this inference for all agents and all propositions. What we are in a position to know outstrips what we actually know.

A rough picture of the view we've sketched so far is that, when evaluating whether an agent is in a position to know that *p*, we hold fixed facts about the world beyond the agent's mind (the truth of the target proposition, whether the knower's context is free of Gettier traps) and vary facts about the agent's psychology (what beliefs they actually hold, any idiosyncratic psychological limitations they are subject to). We give the modal landscape a good shake and see if any worlds in which the agent knows that *p* pop out.

This picture is missing one key ingredient—and it is here that we must qualify our characterization of the position to know as depsychologizing. Although epistemic position largely abstracts away from an agent's psychology, there must be *something* about the agent's mental life that we hold fixed: we must hold fixed whatever it is that fixes the agent's rational access to the world.

---

[11] And then again, maybe not. Williamson's (2000) arguments for anti-luminosity are certainly relevant here.

What does that mean? Answering that question puts us squarely in substantive epistemological territory. Plausible answers include an agent's evidence, an agent's evidence plus her priors or epistemic values, an agent's (epistemic) competencies, or an agent's way of looking at the world. Fortunately, the cases considered in the next section do not require that we take a stand on this substantive question.[12] It's enough to note that a significant change of new, non-inferred evidence bearing on a proposition tends to put an agent in a new epistemic position toward it.

A full theory of the position to know would say more about what exactly grounds our rational access to the world and exactly how and when an agent's psychological state can affect her position to know. Indeed, I take one of the main upshots of this paper to be that epistemologists ought to devote significantly more resources to understanding the position to know. Nevertheless, even without a full theory, we're capable of recognizing clear cases and of articulating some general conditions on the concept. "Reasonability"

---

[12] This is, perhaps, a little quick. In particular, it's difficult to see how Williamson's E=K thesis, combined with an evidentialist interpretation of the position to know, fits this model. Suppose that the facts that make up an agent's epistemic position are exhausted by facts about the agent's evidence. For Williamson, your evidence is all and only what you know. Suppose that I am in a position to learn that p at t1 and then come to believe p and thereby know that p at t2. For Williamson, at t2 I have special evidence for p that I did not have at t1, namely the fact that p. At t2, my epistemic probability for p conditional on my evidence is 1 though it may not have been 1 before.

The crucial thing for our purposes is that, on the E=K thesis, coming to know something changes one's evidence (and so, plausibly, one's epistemic position) in an important way. On this picture, it is backward to try to evaluate whether an agent is in a position to know by assessing whether they could know *holding fixed their prior epistemic position*: coming to know that p is always (or at least usually) to enter a new epistemic position with respect to p.

Here, I am forced to show my hand: I don't think the E=K thesis is true. And part of the reason I think it isn't true is that I think evidence is a kind of support relation that is anti-reflexive (or at least, behaves anti-reflexively in typical cases). (See Brown 2013.) P can't be evidence for itself, and so, even if *only* knowledge is evidence, it can't be that *all* knowledge is evidence bearing on any proposition: my knowledge that p doesn't count as evidence for p, even if it counts as evidence for something else.

Still, we may be able to recover something like my account, even assuming a Williamsonian model. Perhaps we could introduce a notion of epistemically accessible transitions between epistemic positions. If I am in a position to know that p, then there is an epistemically accessible transition from my current epistemic position to an epistemic position that includes my knowledge that p. I'm grateful to an anonymous reviewer for encouraging greater clarity on this point.

and "knowledge" are similarly contested terms, and so, it's no dialectical disadvantage that there is no generally agreed upon theory. We've now said enough about the general structure of the position to know—that it *depsychologizes* and *modalizes*—to begin our assessment of the PtK-Norm of assertion.

3. Three Cases

How should we evaluate the norm of assertion? That depends on the theoretical work we want the norm of assertion to play. Why should we hope that there is a unified, simple, and epistemic norm of assertion in the first place?

There are two main reasons. First, assertions are a kind of affirmation, and as such it bears a close relationship to beliefs or judgments.[13] Insofar as there are epistemic rules that govern our beliefs, we should expect that there are analogous bounds on assertion. I'll say more about the extent and limits of such similarities and why the PtK-Norm is the right sort of analogue in §4. Second, there's a coordination problem for assertion that would be significantly resolved by a simple norm that operates, at least, as a default. There are several valuable epistemic goals that make plausible aims for assertion— truth,[14] likely truth, reasonability,[15] justification,[16] supportive reasons,[17] rational credibility,[18] knowledge,[19] and certainty[20] among others. But if one person asserts with the aim of saying what is probable and another with the aim of saying what is certain, it will be hard to know how to update our opinions upon receiving testimony.

---

[13] Here I draw on Williamson (2000, 255–56) and Sosa (2011, 48).
[14] Weiner (2005)
[15] Lackey (2007)
[16] Kvanvig (2011)
[17] McKinnon (2015)
[18] Douven (2006)
[19] Unger (1975), Williamson (2000), DeRose (2002), Benton (2011), Sosa (2011), a near-variant in Turri (2011) and a sympathetic (though nuanced) account in Goldberg (2015a).
[20] Stanley (2008)

For this coordination problem, it matters very much what actual conventions are already in place, even if those conventions are robustly contingent. If everyone already drives on the right side of the road, then that very practice makes it sensible for there to be a norm to drive on the right side of the road, even if it would have been no worse for the practice to have developed in such a way that the sensible norm was to drive on the left. It is appropriate, therefore, for us to look at the actual linguistic practice of contemporary English speakers in order to determine what the norm is. By doing so, we freely confess that the norm of assertion we develop is only binding for contemporary English speakers: it's possible that other linguistic communities have solved the coordination problem for assertion in a different way.

I propose, therefore, that we test the PtK-Norm by seeing whether it gets the correct verdict in three cases: selfless assertions, Moorean paradoxes, and lotteries. These cases are not chosen at random, for these cases jointly pose a difficult puzzle. Moorean paradoxes and lotteries seem to show that *nothing weaker than knowledge* could be the norm of assertion.[21] In the lottery case, for instance, one's epistemic position could be arbitrarily good short of knowledge and still fail to license the assertion that one's ticket lost. In contrast, selfless assertions seem to show that *nothing as strong as knowledge* could be the norm of assertion: any norm that entails even belief will be too strong. But this leaves no conceptual space for any norm of assertion to succeed: any norm will be too weak or too strong. Perhaps we need to give up our intuitions about one set of cases or the other.

---

[21] Sanford Goldberg summarizes part of Williamson's contribution this way: the "KNA appears to be better positioned than any epistemically *weaker* standard to explain why lottery propositions...are not properly assertable even when the odds of one's winning are arbitrarily small" (Goldberg 2015b).

Enter the PtK-Norm of assertion, which seamlessly delivers the correct verdict in all three cases. There *is* something weaker than knowledge that can play the role of the norm of assertion after all.

I won't say much to defend the intuitions that Lackey is right about selfless assertions and the knowledge camp are right about Moorean paradoxes and lotteries. I'll assume each side is correct about their favored cases (as I believe they are) and show why the PtK-Norm gets all three cases right. (As always, those who disagree with the starting intuitions can read the paper as defending a conditional claim.) But enough introduction—onto the cases.

3.1 Selfless Assertions

Lackey (2007) develops several cases of selfless assertion in which agents can properly assert *p* despite not knowing *p*. In each case, the agent fails to know *p* by failing to believe *p*. Here's one of her examples:[22]

> DISTRAUGHT DOCTOR: Sebastian is [a] ...pediatrician [who] recognizes ...that all of the scientific evidence shows that there is absolutely no connection between vaccines and autism. However, shortly after his ...daughter received one of her vaccines, ...she was soon diagnosed with autism. …[T]he grief and exhaustion brought on by his daughter's recent diagnosis cause him to abandon his previously deeply-held beliefs regarding vaccines. …[W]hile performing a well-baby checkup..., the child's parents ask him about ...the rumors surrounding vaccines and autism. Recognizing both that the current doubt he has towards vaccines was probably brought about through [his] emotional trauma ...and that he has an obligation to his patients to present what is most likely to be true, Sebastian asserts, "There is no connection between vaccines and autism" …[although it's false that] Sebastian himself believes or knows this proposition. (Lackey 2007, 598-99)[23]

---

[22] The distraught doctor case is not chosen from among Lackey's examples at random. I think the distraught doctor and racist juror case are more compelling than the creationist teacher case, in part (I contend) because it's unclear that the creationist teacher really is in a position to know (given their network of background beliefs) that evolution is true.

[23] Requiring that testifiers know in order to transmit knowledge "conflates *reliable knowers* with *reliable testifiers*" (Lackey 1999, 481).

Cases of selfless assertion are problematic for the knowledge norm of assertion, for they are cases in which agents properly assert what they do not know. Sebastian doesn't know that vaccines don't cause autism because he doesn't believe that vaccines don't cause autism.[24] So, how can he properly assert it?

The position to know norm, however, does not encounter the same difficulty. It isn't just rational for Sebastian to believe that vaccines cause autism: Sebastian is in a position to know it! The only thing he's missing is belief: if he *were* to believe it, he would thereby come to know it. The PTK-Norm thus accurately predicts that Sebastian can felicitously assert that vaccines do not cause autism.

There's a natural worry concerning Lackey's distraught doctor case: Given the high moral stakes, isn't it possible that moral intuitions are interfering with our more narrowly assertoric judgments?[25]

We can assuage this worry by developing a parallel case in which the moral stakes are not so high. Consider the following story:

> Santa Claus: Jimmy has recently reached the age when he's too old to believe in Santa Claus. He has good evidence that Santa doesn't exist, and he is in a position to reasonably believe and even know that Santa doesn't exist. But for reasons having entirely to do with Jimmy's particular psychology, doubt persists—he can't fully believe (or thereby know) that there is no Santa. Perhaps he's afraid that if he doesn't believe in Santa, he won't get any presents. His brother, a year

---

[24] It's worth flagging that there is a minority report in epistemology that knowledge does not entail belief (see, for instance, Black 1971). In this paper, I'll simply assume the orthodox position that knowledge entails belief, but the combination of views that includes the K-Norm of assertion and the thesis that knowledge does not entail belief is worthy of exploration.

[25] Thus, Jonathan Kvanvig writes that though Lackey's selfless assertions "are clearly appropriate, that intuition is fairly clearly an intuition about what is, all-things-considered, appropriate" (Kvanvig 2011, 235) and do "not survive scrutiny from an epistemic point of view" (Kvanvig 2011, 235). But the locus of evaluation should not be purely epistemic but assertoric. Cases of selfless assertions just are ones in which the asserter fails as an epistemic agent by not having beliefs that are well-supported for them. The question is not whether selfless asserters perform well *epistemically* but whether they perform well *assertorically* despite their epistemic failings.

younger than Jimmy, asks him whether Santa exists. Jimmy replies: "There is no Santa."

The moral stakes are much lower in this story. In fact, to the extent that there are moral stakes, we are as likely as not to think that Jimmy was in the wrong—he oughtn't to have spoiled Santa for his brother! But even if Jimmy's assertion was morally wrong, it was assertorically felicitous (just as it can be felicitous but morally wrong to be rudely candid). Jimmy did nothing wrong in his duties as an asserter even if he failed his duties as a brother. And once again, Jimmy does not know the asserted proposition, though he is both in a position to know and to reasonably believe it.

Our intuitions in cases of selfless assertion remain consistent even when there is no moral weight in favor of (and perhaps even some against) asserting a proposition. Lackey's cases illustrate how selfless assertions may be especially important in morally loaded situations. But our judgments are not thereby captive to the purely moral elements of the cases.

3.2 Roles to the Rescue?

Prominent in Lackey's stories are the *social roles* (doctor, juror, teacher) of the selfless asserters. Thus, Ernest Sosa suggests this strategy for the K-Norm to explain Lackey's cases:

> I propose to accommodate [Lackey's cases] by means of a distinction between assertion in one's own person, as a human being who communicates with other human beings, and assertion as occupier of a role. ...One may still proceed with epistemic propriety if one is playing one's epistemic role properly. To play one's epistemic role in such contexts... one serves as a mouthpiece for a deeper institutional source of the information conveyed... (Sosa 2011, 47).[26]

---

[26] Likewise, Kvanvig observes of a similar case that the asserter "is being paid to fill a particular social role" (Kvanvig 2011, 235).

As applied to the case of the distraught doctor, this is to say that although Sebastian does not know that there is no connection between vaccines and autism *in his own person*, Sebastian does know it *in his role as a doctor*. The medical profession is the "deeper institutional source of the information conveyed."

Focusing on the role that an agent is in rather than the agent herself depsychologizes away from facts about the particular mental life of the agent. In this way, Sosa's response shares a core feature of the position to know hypothesis and respects Lackey's insight that proper assertions need not pass through the doxastic state of the asserter: proper assertion can go through a role-enabling social source instead. The central difference between Sosa's proposal and mine (along with Lackey's) is this: Sosa's account requires a deeper institutional source—or at any rate an established and recognizable social structure—within which the asserter occupies a role. The PtK-Norm (and the RTB-Norm) requires no such social source or role. Ostensibly in Sosa's favor is the fact that each of Lackey's cases does employ agents whose social role (e.g. doctor) is in focus. Nevertheless, we can develop compelling cases of selfless assertion that are not role- or institution-dependent.

Imagine, for instance, that Sebastian is not at work, but simply walking down the street when he overhears two parents discussing whether it is safe to vaccinate their child. Sebastian, ever eager to be helpful if not also a bit of a busybody, leans over and says, "There is no connection between vaccines and autism," before continuing on his way. Sebastian's assertion is still acceptable, even though (1) he does not know (because he does not believe) that there is no connection between vaccines and autism and (2) he at

no point invokes—explicitly or implicitly—the social role of being a doctor. We might doubt the efficacy of Sebastian's assertion but not its felicity.

Or better yet, imagine that in the story above Sebastian is not a doctor at all, but simply someone who has happened to study the relevant evidence and recognizes its significance. Perhaps there aren't even doctors in Sebastian's community at all—he's just some bloke who has looked at the evidence on the issue carefully. It is still felicitous for Sebastian to assert, "There is no connection between vaccines and autism," even though the role of doctor is entirely absent from the case.

It might be argued that just in virtue of asserting without knowledge, Sebastian creates a role for himself to speak from. He takes up the role of public informant about vaccinations and autism. By asserting something he does not know, he stops playing himself and plays the role of someone who knows what he is in a position to know. But if this is all that playing a role requires, then *all* assertions (or all public assertions, which are the dominant mode of assertions) involve taking on a certain social role (e.g. testifier, questioner, participant in an inquiry) so that it's inadequate to try to handle Lackey's cases of selfless assertion by *bracketing* them as a kind of special case. Assertion in a role—in this thin sense of role—is normal not exceptional, and the norm for assertion should reflect that. If one can always assert from the role of "informant on proposition p," one need not know that p at all. It's sufficient to be in a position to know that p.

Merely appealing to roles cannot handle revised versions of Lackey's cases unless the concept of role is so thin as to make virtually any assertion an instance of assertion from a role: Either roles cannot explain all cases of felicitous, selfless assertion or else roles cannot distinguish cases of selfless assertion from normal assertions. Moreover, it's

a virtue of the PtK-Norm that it can explain cases of selfless assertion without complicating the account by appealing to special roles.[27]

3.3 Moorean Paradoxes

Whereas the knowledge norm had trouble explaining the felicity of selfless assertions, the reasonable to believe norm has trouble explaining the infelicity of Moorean paradoxes. Moorean paradoxes are expressed when an agent makes the infelicitous assertion, "*p*, but I don't know [or believe] that *p*." Lackey explains why the RTB-Norm cannot get the right verdict on its own:

> [I]t is reasonable for Sebastian in DISTRAUGHT DOCTOR to believe both *that there is no connection between vaccines and autism* and *that he neither believes nor knows that this is the case*. Accordingly, his asserting… "There is no connection between vaccines and autism, but I don't know that this is the case" …satisf[ies] the RTBNA (Lackey 2007, 613).

But can't the exact same thing be said about the PtK-Norm? Consider the following argument:

> Sebastian is in a position to know in DISTRAUGHT DOCTOR both *that there is no connection between vaccines and autism* and *that he neither believes nor knows that this is the case*. Accordingly, his asserting "There is no connection between vaccines and autism, but I don't know that this is the case" satisfies the PTK-Norm (adapted from Lackey 2007, 613).

Isn't this a knock-down argument against the PtK-Norm? Or at least that the PtK-Norm, like Lackey's, needs to be supplemented? Consider, in contrast, how easily the K-Norm handles such cases.

If an agent knows that "*p*, but I don't know that *p*," then they know both that *p* and that "I don't know that *p*." But they can only know that "I don't know that *p*" if it's

---

[27] Likewise, McKinnon objects that Sosa's view "seems to require all sorts of specialized norms of assertion, depending on one's institutional role" (McKinnon 2015, 175). For a helpful, critical discussion of Sosa's defense of the K-Norm, see Kelp (2015).

true that they don't know that p. So, they both know and do not know that p—a contradiction! Since such propositions can't be known, the K-Norm elegantly explains why they also cannot be properly asserted.[28]

But note that the argument shows not just that Moorean propositions are unknown but that they are unknowable. This is good news for the PtK-Norm. For the fact that Moorean propositions are unknowable (for the speaker) rather than simply unknown explains why no one is in a position to know them. Consider the following argument:

(7) It is impossible for anyone to know a Moorean proposition.[29]

(8) If it is impossible for anyone to know a proposition, then no one is in a position to know it.

(9) Therefore, no one is in a position to know a Moorean proposition.

If this argument is sound, then the PtK-Norm can explain why Moorean propositions are not properly assertable just as easily as the K-Norm: One cannot properly assert Moorean propositions because one is never in a position to know them. The argument is valid. Premise (7) is just a consequence of Williamson's argument for the unknowability of Moorean propositions. The crucial premise, then, is (8). Is (8) true?

Premise (8) falls straight out of the definition of position to know that we considered earlier:

PtK: S is in a position to know that p iff S could know that p given their actual epistemic position.

---

[28] Cf. Williamson (2000, 253). DeRose notes that this is "one of the most important" arguments for the K-Norm (2002, 180).

[29] Of course, here and in the following lines of the argument, I mean any Moorean proposition *of which one is the subject*. It's impossible for me to know "p, but I don't know that p," even if it is possible for you to know what is arguably the very same proposition (when addressed to me): "p, but you don't know that p." But for me to be in a position to know a proposition, it must be possible for *me* to know it. For more on second-person variants of Moorean paradoxes, see the next chapter.

If a proposition is impossible for a speaker to know, then the speaker could not know it from *any* epistemic position, much less from the epistemic position they are actually in. Therefore, the PtK-norm can indeed explain why Moorean paradoxes are improper to assert.

But what of the argument we adapted from Lackey? Where did it go wrong? The mistake is the apparently innocuous move from being in a position to know each conjunct to being in a position to know the whole conjunction. Sebastian is in a position to know both "that there is no connection between vaccines and autism," and "I [Sebastian] neither believe nor know that this is the case." But this doesn't entail that Sebastian is in a position to know the conjunction "that there is no connection between vaccines and autism *and* I [Sebastian] neither believe nor know that this is the case." Indeed, as we just saw, it is impossible for Sebastian to be in a position to know a proposition of this form. Being in a position to know is not closed under conjunction.[30]

---

[30] In his (2015a) *Assertion*, Goldberg considers (although he does not endorse) the PtK-Norm of assertion and even goes so far as to offer a friendly argument that the PtK-Norm can explain the infelicity of Moorean paradoxes. I believe, however, that his account is mistaken. His argument relies on the claim that "being in a position to know… can be… defeated by being in a position to know that one does not know" (160). So, then if one really were in a position to know <I don't know that *p*> one wouldn't be in a position to know that *p*. According to Goldberg, the agent who is in a position to know <I don't know that *p*> is neither in a position to know that *p* nor the conjunction <*p*, but I don't know that *p*> whereas on my view the agent can be in a position to know each conjunct of the Moorean proposition but not the conjunction. It's easy to see the attraction of the principle that Goldberg proposes. If my total evidence is such that, if I reflected on it, I would not be doxastically justified (and so would not know) that *p*, plausibly I'm not in a position to know that *p* now, even before I've reflected on the relevant evidence. So, perhaps any time I'm in a position to know that I don't know that *p*, I'm not in a position to know *p*.
But the principle loses plausibility when we focus on cases in which the reason one knows that one doesn't know have nothing to do with the epistemic quality of the belief in question. If I am in a position to know—or even know outright—that I do not know that *p* but *only because* I know I can't psychologically get myself to believe *p*, this doesn't exclude me from being in a position to know that *p*. Lackey's case of Sebastian the doctor is useful here. Sebastian might be well aware that he can't get himself to outright believe <there is no connection between vaccines and autism>. So, he knows (and *a fortiori* is in a position to know) that he does not know <there is no connection between vaccines and autism>, since knowledge entails belief. But this in no way prevents Sebastian from being in a position to know <there is no connection between vaccines and autism>. What's so frustrating about Sebastian's psychological position is (among other things) that he can't get himself to believe what he is in a position to know.

But isn't it strange that the proper assertability isn't closed under conjunction? According to the K-Norm, Moorean paradoxes are not properly assertable because one of the two conjuncts is not properly assertable. Either the agent doesn't really know that p or doesn't really know that "I don't know that p." But according to the PtK-Norm, the agent is in a position to know both p and "I don't know that p" but not the conjunction of the two. The knowledge norm, unlike the position to know norm, seemingly maintains that a conjunction is properly assertable iff its conjuncts are each properly assertable.

Indeed, it is intuitive that the proper assertability of conjunctions supervenes on the proper assertability of their conjuncts. The truth of conjunctions supervenes on the truth of their conjuncts, and knowledge—as a truth-seeking state—often follows suit.

Intuitive, but false. Consider the preface paradox. An author believes—and we might add, asserts—each individual proposition in a book she publishes. But in the preface, she humbly adds, "Despite the tremendous help I have received from my colleagues in writing this book, I have surely made some mistakes here and there. All errors are, of course, my own." She believes, on the basis of good meta-inductive evidence, that one of the propositions she believes (and asserts) in the book is false. For each proposition in the book, she believes it. But she does not believe the conjunction of all her beliefs.

Little changes if we add that—miraculously—she in fact *knows* (and, *a fortiori*, is in a position to know) every individual proposition in the book. It is still not rational for her to believe the conjunction of every proposition: the meta-inductive evidence that she's made a mistake somewhere is too strong. And if it isn't rational for her to believe

the entire conjunction, she neither knows nor is in a position to know the conjunction. So, even on the knowledge norm, proper assertability is not closed under conjunction.

Thus, the PtK-Norm and the preface paradox share a common consequence. In both cases, to determine whether an agent is in a position to know a conjunction of propositions, we must look at more than just whether the agent knows each proposition individually but at how the agent's epistemic access to each proposition reacts when combined with others. In the preface paradox, the chance of error of the whole *compounds* when each conjunct is asserted not independently but as a part of a conjunction. In the case of the Moorean paradox, none of the accessible worlds that explain the agent's being in a position to know *p* overlap with the worlds that explain the agent's being in a position to know that "I don't know that *p*." Although there are worlds in which an agent in the speaker's epistemic position knows that *p* and worlds in which an agent in the speaker's epistemic position knows that "I don't know that *p*," there are no worlds in which an agent knows both propositions. In both cases, confusion is avoided by refusing to reduce the assertability of a conjunction down to the assertability of its conjuncts.

Here's a different worry: Sometimes it's not clear when someone has asserted a conjunction instead of merely the separate conjuncts. The mere presence of a conjunction such as "and" or "but" can't be the only factor. It would be just as bad for Sebastian to tell his patient, "There is no connection between vaccines and autism: I don't know that this is the case," although, when spoken instead of written, nothing overtly signifies that the propositions are conjoined. So, in borderline cases, how are we to tell whether Sebastian has asserted a conjunction or merely two discrete propositions?

Plausibly, there are clear cases on either side. Suppose Sebastian says to his patient, "There is no connection between vaccines and autism," and adds in a hushed tone, "but I don't know that this is the case." Merely hushing the second clause is not enough to absolve him of having asserted the whole conjunction rather than two simple propositions. On the other hand, if Sebastian says to his patient, "There is no connection between vaccines and autism," and then rushes across the hall to his therapist's office to lament, "I don't even know that what I said is true!" Sebastian has not asserted the whole conjunction but only two simple propositions in close temporal proximity. But there will be cases in which it's not obvious whether Sebastian has asserted the whole conjunction or just each would-be conjunct. If both his patient and his therapist are in the room and Sebastian says, "There is no connection between vaccines and autism," before sharply turning to his therapist and adding, "I don't know that what I just said is true," has he asserted the whole conjunction or merely two disjoint propositions in near temporal succession? It's not clear, and so it's not clear what the position to know hypothesis predicts is properly assertable. It's also unclear whether or not this assertion (or assertions) is (or are) felicitous.

I answer: It is no knock on a theory that it predicts confusion where confusion is. We feel genuinely confused about whether Sebastian has asserted properly when it's unclear whether his assertion is a conjunction or two discrete statements in close proximity. It's evidence *for* the PtK-Norm and not against it that the cases in which our intuitions are muddled covary with cases in which it's unclear what the PtK-Norm predicts. This is precisely where we should expect to be confused if the PtK-Norm is true:

it's further evidence that we implicitly appeal to the PtK-Norm to determine whether an

assertion is felicitous.

3.4 Lotteries

In addition to Moorean paradoxes, the RTB-Norm struggles to explain

assertions about lotteries. Williamson writes:

> Suppose that you have bought a ticket in a ...lottery. …[Y]our ticket did not
> win, but I have no inside information to that effect. On the merely
> probabilistic grounds that your ticket was only one of very many, I assert to
> you flat-out "Your ticket did not win" ...[M]y grounds are quite inadequate
> for that outright unqualified assertion. ...I was representing myself to you as
> having a kind of authority to make the flat-out assertion which in reality I
> lacked (Williamson 2000, 246).[31]

The knowledge norm explains why it's impermissible to assert "Your ticket lost"

even when the odds are arbitrarily high. Exactly why we cannot know lottery

propositions without special evidence is a matter of debate, but plausibly it's

something like this: Knowledge requires the elimination of (salient) chance. And it's

very salient in lottery cases that there is a chance—however small—that your ticket

will win.[32]

Lottery cases show that one's evidence could make a proposition as probable

as one likes (shy of 1) and still fail to warrant assertion. This puts pressure on the

reasonable to believe norm. Lackey concurs:

> [T]he very good probabilistic grounds for lottery propositions presumably
> render it reasonable to believe in their truth. Hence, so long as a speaker
> asserts a lottery proposition ...because of this reasonableness, these assertions
> are taken to satisfy the RTBNA (Lackey 2007, 613).

---

[31] Williamson later clarifies: "There is a special jocular tone in which it is quite acceptable to say '[Come
off it—] Your ticket didn't win', but the tone signals that the speaker intends not to make a flat-out
assertion. In the imagined example, I do not use that tone" (Williamson 2000, 246).
[32] As Lackey reminds us, "it is not universally accepted that subjects *do* in fact lack knowledge in lottery
propositions" (Lackey 2007, 617). Nevertheless, for the purposes of this paper, I take as given the orthodox
position that we do not know lottery propositions.

But unlike the RTB-Norm, the PtK-Norm gets the intuitively correct verdict. It isn't just that agents *don't* know lottery propositions without special evidence: they *can't* know them. Lotteries aren't the right places to go looking for knowledge: the terrain is too full of luck and salient chances for knowledge to survive without special evidence.

> (10) It is impossible for anyone to know <this ticket lost the lottery> without special evidence (e.g. inside information on the result of the draw).

> (11) If it is impossible for anyone to know <this ticket lost the lottery> without special evidence, then no one without special evidence is in a position to know it.

> (12) Therefore, no one without special evidence is in a position to know <this ticket lost the lottery>.

This argument closely resembles our argument that no one is in a position to know Moorean paradoxes. A notable difference is that lottery propositions are not absolutely impossible to know but only impossible to know *without special evidence*. Consider again PtK:

> **PtK**: S is in a position to know that *p* iff S could know that *p* given their actual epistemic position.

Premise (11) is true if obtaining (new) special evidence changes one's epistemic position. Clearly, obtaining special evidence quite dramatically changes one's epistemic position. Therefore, premise (11) is true. If one *can't* know lottery propositions in normal evidential circumstances, then one isn't in a *position* to know either. Thus, the PtK-Norm, unlike the RTB-Norm, correctly predicts that lottery propositions are not properly assertable in ordinary evidential contexts.

3.5 Why the Not Misleading Norm is not Enough

Lackey is well aware of the challenges Lottery cases and Moorean paradoxes pose for the RTB Norm. To resolve this tension, Lackey augments the RTB norm with the Not Misleading Norm of Assertion. The final version of this principle reads thus:

> NMNA**: S should assert that *p* in context C only if it is not reasonable for S to believe that the assertion that *p* will be misleading in C relative to the purposes of the exchange in question (Lackey 2007, 615).

One source of discomfort for NMNA** is the ease with which we as reasoners can be misled by making irrational yet predictably human inferences. I find that when I focus on how easy it is for an assertion to be misleading, and how onerous such misdirection can be to contain, I lose the intuition that a speaker can never assert that *p* if they reasonably believe that it will be relevantly misleading in some way. Because humans are subject to widespread and deeply ingrained biases, it's often very hard not to be misleading in some way or another. And especially if the reason it's hard not to be misleading is that the *auditor* is being unfair or irrational, it seems overly demanding to require that the speaker eliminate all reasonably anticipated sources of relevant misdirection before asserting. Suppose, for instance, that when her education is being discussed in a certain society, men are disposed to believe that the speaker has gone to secretarial school even if she in fact studied philosophy. It seems overly cumbersome for the norm of assertion to demand that she always add "but I didn't attend secretarial school!" when mentioning her education, even if her education is contextually relevant and the speaker knows full well that her audience is drawing a false conclusion. Her assertion is, through no fault of her own, misleading in a contextually relevant way, but it's not the assertoric responsibility[33]

---

[33] Of course, there may be cases in which we have moral reasons to contain misdirection even when it isn't our narrowly assertoric duty to do so.

of the speaker to correct this particular kind of misdirection.[34] Rather, it's the
responsibility of the audience to manage the influence of background cultural
assumptions that make the assertion misleading.

But there is an even deeper reason to be dissatisfied with the attempt to save the
RTB-Norm by tacking on NMNA**. Lackey owes us an answer to this question: Why is
it that Moorean paradoxes and lottery propositions *are* misleading? We should want, in
our theory of assertion, not just a norm that prohibits being misleading but an explanation
whence the misdirection comes.

If knowledge is the guiding normative property for assertion (as both the K-norm
and PtK-norm maintain), there is a natural answer to the question of why lottery
propositions and Moorean paradoxes are misleading: assertions present their speakers as
(at least) in a position to know the asserted proposition, and lottery propositions (in
normal evidential circumstances) and Moorean paradoxes are unknowable. But if we
don't normally assume that people are in a position to know what they assert, why are we
confused when confronted with lottery propositions and Moorean paradoxes?

At this point, proponents of the RTB-norm might return to the Gricean injunction
to make one's assertions as informative as one can: Don't say things everyone already
believes (cf. Weiner 2005). And everyone already reasonably believes (on Lackey's

---

[34] Another way to interpret this case is as one in which the speaker is not misleading (the speaker is just
telling the truth!) but the audience is being misled by their own preconceptions (or by the systemic injustice
of the society). The speaker isn't misleading her audience, rather she is letting her audience be misled by
their own biases. The misdirection is occasioned but not caused by her speech. But introducing this
distinction raises a challenging question for the NMAA**. For suppose I say that you have lost the lottery
(which it turns out, unsurprisingly, that you have). I am telling you something true. Why is it that I count as
misleading you in this case rather than simply letting you be misled by your preconception that I would
only tell you that you had lost if I had insider information? It is, at least, not obvious how to draw the line
between misdirecting and allowing to be misled in such a way that the lottery case and Moorean paradoxes
come out on one side and the mistaken-for-a-secretary case comes out on the other. Thanks to an
anonymous reviewer for conversation on this point.

account) that you've lost the lottery. So, when someone says, "you lost the lottery," for instance, they violate the mandate of informativeness. We naturally resolve this violation by charitably assuming that the speaker be in an even better epistemic situation than the rest of us. That would explain why they flouted the conversational maxim of informativeness. But this response falls short. For it's still improper to flat-out assert "you lost the lottery," even when explaining lotteries to a lottery-novice who doesn't have any idea what will happen to their ticket.[35]

3.6 Knowledge-Oriented Varieties of Reasonability

Perhaps Lackey was too quick to concede that the relevant kind of reasonability did not secure the right judgment all by itself. Douven (2006, 2009) and Kvanvig (2009) propose broadly reasonability-based norms of assertion (rational credibility and justification respectively) that, they argue, secure the intuitive verdict in Moorean paradoxes and lottery propositions without any additional norms.

Thus Kvanvig says that assertion requires justification of "the kind that puts one in a position to know" (Kvanvig 2009, 149).[36] And Douven builds this principle into his account of rational belief: $RB_i(\varphi) \Rightarrow Cr_i(K_i(\varphi)) \geqslant Cr_i(\neg K_i(\varphi))$ (2009, 371). Roughly, if an agent rationally believes a proposition, then their credence that they know it is higher than their credence that they don't. On this account, one can't rationally believe that "*p*, but I don't know that *p*" because that would require believing that it's both more and less probable than not that I know that *p*.

---

[35] Note, too, that the assertion is not helped by adding more information: "Your ticket lost because there are so many tickets that it is exceedingly unlikely that you won," is just as bad if not worse than the less informative "your ticket lost."

[36] For Kvanvig, the justification need not *actually* put the speaker in a position to know since the content of the assertion could be false or Gettiered. But it can't be a defect in the speaker's justification that they are not in a position to know.

But Douven's principle is contentious. Plausibly, it *is* possible to rationally believe (or even know) a proposition while simultaneously having a higher credence that one does not know that proposition than that one knows it.

This is most compelling when an agent has misleading evidence that they do not know. Imagine, for instance, someone living in a community of Ungerian (1975) skeptics. Such skeptics maintain their ordinary beliefs, but because they think the demands of knowledge are very high, they think that (almost) none of their beliefs constitute knowledge. In such a community, agents might have very good testimonial evidence that they do not know what they believe—after all, everyone in that community says that Ungerian skepticism is true. But if they continued to form beliefs in the normal way (despite a commitment to an epistemological theory that these beliefs, however well-supported, never constituted knowledge) by perception, valid reasoning, etc., I'd be inclined to say that their ordinary, first-order beliefs were still rationally believed (and known!) even if their credence that those beliefs were knowledgeable was near 0.[37]

But there's a deeper dissatisfaction that affects both Douven's and Kvanvig's accounts. Why is it that justification requires being of the sort that puts one in a position to *know*? Or why does rational belief require thinking it more likely than not that one *knows* what one believes? Presumably because beliefs are aimed, in some sense, at knowledge. Otherwise, why care that one's beliefs fall short of knowledge rather than, e.g., truth or probable truth? For the varieties of reasonability that Douven and Kvanvig

---

[37] Admittedly, such a community would plausibly have a different norm of assertion than the one the community of contemporary, English speakers have developed. The restricted point here is that Douven's epistemic principle is contentious. Thanks to an anonymous reviewer for encouraging further clarity on this point.

care about (rational credibility and justification), one has only really been reasonable if one has done well enough with respect to knowledge.

This gives knowledge uncomfortable pride of place in apparent competitors to the thesis that knowledge is the epistemic property central to assertion. There's a principled reason Lackey appeals to an accuracy-based account of reasonability as she engages Williamson's arguments. For suppose that in order to be reasonable (or justified or rational) one aims not just to be accurate but to know, and that in order to properly assert one aims to be reasonable. Then in aiming to properly assert one would also aim to know, since it is by aiming at knowledge that one achieves the aim or reasonability. But once it's admitted that we aim to know in order to properly assert, it's hard to resist the pull of the knowledge norm. If knowledge is the ultimate aim of reasonability, then why should assertion aim at a merely derivative goal? Better to say that knowledge is the normatively guiding property for both belief and assertion.[38]

3.7 Putting the Audience in a Position to Know

Now that we've examined the importance of Moorean paradoxes, we're able to consider one more attempt to explain selfless assertions. Like Lackey, I have explained selfless assertions by dropping a belief requirement on assertion. But we both maintain that it is the epistemic position of the speaker that enables proper assertion. A more radical solution would be to deny that the speaker must be in any particular epistemic position to properly assert. García-Carpintero and Pelling have each proposed norms that focus not on the originating epistemic position of the speaker but on the resulting epistemic position of the audience:

---

[38] See §4 for more on the relationship between the norms of belief and assertion.

TKR: One must ((assert *p*) only if one's audience comes thereby to be in a position to know *p*) (García-Carpintero 2004, 134).

PKA: One's assertion that *p* is proper only if it is fit to give a hearer knowledge that *p* (Pelling 2013, 297).[39]

As I argue in more detail in the next chapter, such assertions cannot explain the felicity of certain second-person variants of Moorean paradoxes. For instance, it would be felicitous for Jean Valjean to tell Inspector Javert, "You'll never believe that I've changed even though I have!" This is something the speaker, Jean Valjean, is in a position to know. But Javert, the intended audience, couldn't possibly know that Javert has changed and that he'll never believe it: that would commit Javert (but not Valjean) to a Moorean paradox![40]

If it's impossible for one's audience to know a proposition, it is neither fit nor possible to put the audience in a position to know that proposition. The PtK-Norm rightly maintains that it is the epistemic position of the speaker rather than the audience that licenses proper assertion, while still allowing speakers the flexibility to make selfless assertions in the best interests of their audience.[41]

4. Assertion and Belief

There's another important argument against the PtK-Norm—especially favored by the knowledge camp—from the apparent similarities between assertion and belief. Williamson and Sosa both argue that the knowledge norm of assertion illuminates a deep parallel between assertion and belief. Williamson writes:

---

[39] In later work, Pelling modifies this view, claiming that that an interpersonal norm such as knowledge provision is better understood as a norm for the speech act of telling than the speech act of asserting, a distinction he makes in Pelling (2014). (He notes his change of view in 2013, 348, fn. 9.)

[40] Note that since the claim includes that Javert will *never* believe that Valjean has changed, Javert cannot come to know this proposition in the future as well as at the time of utterance.

[41] For more on what we can learn from second-person variants of Moorean Paradoxes, see Whitcomb (2013) and my discussion in chapter 4 of this dissertation.

> [O]ccurently believing p stands to asserting p as the inner stands to the outer. If so, the knowledge rule for assertion corresponds to the norm that one should believe p only if one knows p (Williamson 2000, 255-56).

Likewise, Sosa says:

> If knowledge is the norm of assertion, it is plausibly also the norm of affirmation, whether the affirming be private or public (Sosa 2011, 48).[42]

It is a negotiable feature of a theory of assertion that the norms of belief and assertion thus align. But suppose that Williamson and Sosa are right: the parallels between belief and assertion run deep. Insofar as knowledge is the norm of belief, austerity in our theorizing favors the knowledge norm of assertion.

The problem can be cast in a broader way. The norms governing belief *do* pass through the doxastic state of the asserter—the doxastic state of the asserter is what we are evaluating when we evaluate belief. We need an answer to the question of how best to evaluate an agent's beliefs. Perhaps an agent should believe *p* only if her belief constitutes knowledge; perhaps an agent should believe *p* only if her belief is rational or supported by good reasons. At any rate, once we've decided on an answer to that question, why not simply cross-apply *that* norm to assertion?

The first answer to this question is the principal retort to every objection from parsimony: the PtK-Norm gets more cases right, and this is what justifies a distinct norm for assertion. This is the conclusion for which the paper has argued so far. But I think we can do better than this. There are reasons to expect the norms of assertion and belief to separate and to separate in precisely the way the PtK-norm predicts.

---

[42] As Goldberg notes, these insights from Williamson and Sosa are anticipated by Dummett: "Judgment . . . is the interiorization of the external act of assertion" (Dummett 1973, 362; cited in Goldberg 2015b, 376).

Williamson and Sosa hold that belief and assertion share an important similarity (the same norm), but they also hold that there's an important difference: the difference between inner and outer, or private and public. It is this very difference that explains why the norms, too, diverge. There's a good reason that the norms for belief may require something (knowledge or rational belief) that passes through the doxastic state of the asserter—the inner, doxastic state *itself* is what is under evaluation. In contrast, the norm of assertion is not an evaluation of a doxastic state, but of an *outer* utterance.[43]

Closely related to this distinction between inner and outer is the distinction between other-affecting and non-other-affecting aims. Indeed, it is the other-affecting aim of assertions that requires that assertions be outer acts in a way that beliefs are not. Typically, when a speaker decides to assert, she thereby decides to try to bring about a certain attitude in her audience. In contrast, when an agent decides to believe, she thereby decides to bring about a certain attitude in herself. There may be exceptions, as when a cleverly framed defendant asserts their innocence against all hope of persuasion or of shifting the common ground. But even in these cases, the other-affecting quality of assertion is invoked to explain why such assertions feel helpless. To assert when no one will believe what you assert is typically to have one's act *frustrated*: to believe when no one will share your belief may be lonely, but it is not to have one's act frustrated in the same way.

Once we have it fixedly in mind that assertion is an outer and typically other-affecting act, it becomes puzzling why assertion should require the asserter to be in any particular doxastic state. And although we'll see some important exceptions to this

---

[43] Cf. Goldberg: "Assertion is a public act, whereas belief is neither an act nor a public matter, and there are reasons to think that both of these differences bear on the respective standards of each" (2015a, 167).

general rule in the next chapter, it's often the goal of an assertion to change the doxastic state of the *audience*. The doxastic state of the speaker is of secondary concern, if of any. This is why it makes sense to focus, as Lackey does, on assertions that are *selfless*. What Lackey identifies as the selfless quality of assertion is the consequence of its other-regarding character. Assertions can be selfless precisely because assertions are other regarding.

Consider: the hearer is at no disadvantage if she receives testimony from those who are merely in a position to know rather than those who actually know. The quality of information received is just as high. Propositions one is in a position to know have the quality Goldberg calls "aptness for communicating knowledge" (Goldberg 2015a, 6). And since the set of things a speaker knows is a subset of the things she is in a position to know, audiences do better with respect to the quantity of propositions from asserters who assert what they are in a position to know rather than merely what they know. Since the quantity of asserted propositions is larger and the quality of asserted propositions no worse, the PtK-Norm is an intuitively better policy for transmitting information.

It seems, therefore, that we do have good reason for having distinct norms for assertion and belief: assertion is immediately other-regarding in a way that belief is not. Indeed, we had to look little further than Williamson's and Sosa's own characterizations of the chief difference between belief and assertion—that of private and public—to find such a reason.

5. Are We in a Position to Know Too Much?

Someone might object: What we are in a position to know far outstrips what we know. This makes it too easy for someone to accidentally follow the norm. Suppose that

you are asked where Tom is and that you don't know the answer—you don't have any belief about where Tom is right now. But instead of confessing your ignorance, you pick a place at random: "Tom's at the bar." You brush off your inquisitor and give Tom's location not a second thought. As it turns out, Tom *is* at the bar. More than that, you remember that he told you yesterday that he would be at the bar at seven, and you know perfectly well that it's seven now. If you'd bothered to think about the question at all, if you'd bothered to put your evidence of Tom's testimony together with your evidence about the current time, you would have known that Tom is at the bar. So, you're in a position to know that Tom is at the bar. But the assertion seems awful! The PtK-Norm, it seems, cannot explain this.

This problem of random assertions that happen to conform to the norm is not, however, unique to the PtK-Norm. It's a general feature of all norms, as Kant taught us, that norms give rise to at least two species of criticisms: (1) not acting in accordance with the norm, and (2) not acting out of respect for the norm. In the story above, Tom is criticizable for not acting out of respect for the norm even though he (by accident) acts in accordance with the norm.

Note that the primary competitors we have considered also face the same sort of objection. If you're in a position to know that Tom is at the bar, it's also reasonable for you to believe that Tom is at the bar, so the case above is equally a problem for the RtB-Norm. A slightly modified version of the case shows that it is also a problem for the K-Norm. Suppose that the speaker knew (perhaps non-occurrently) that Tom was at the bar but said so for reasons entirely unconnected to that knowledge. Irritated by the question about Tom's whereabouts, the speaker picks the location in town furthest from their

current location. It is only afterward—and with some regret—that the speaker later

consciously realizes that his answer helped rather than hindered the inquiry.[44]

Lackey proposes a strategy for handling such cases. She builds a second condition

into her norm: "if one asserted that *p*, one would assert that *p* at least in part because it is

reasonable for one to believe that *p*" (Lackey 2007, 608).[45] Analogous clauses could be

built for the PtK- and K-Norms, e.g., we could add the second condition that "if one

asserted that *p*, one would assert that *p* at least in part because one is in a position to

know that *p*." Adding such a clause seems to solve the problem. Asserting a proposition

at random is wrong since it is not asserted *because* of its epistemic standing, even if the

proper epistemic standing is (as it so happens) possessed by the asserter.

Despite the obvious appeal of Lackey's suggestion, I am hesitant to adopt it. Do

norms in general tend to have the form, "one ought to do the right thing" or "one ought to

do the right thing (at least in part) because it is the right thing?" Surely it is better to do

the right thing for the right reason, but is such a condition baked into the norm itself? At

least not obviously. For one thing, it makes sense to talk of someone doing the right thing

for the wrong reason—and even doing the right thing *only* for wrong reasons, so that their

right action is in no way caused or explained by its being the right thing to do. In Kantian

terms, one can act in accordance with duty without acting from duty. But building the

right causes, reasons, or expression[46] into the norm itself threatens to obfuscate this

distinction.[47]

---

[44] Cf. Turri (2011, 41) for a similar case.
[45] McKinnon follows Lackey's lead and builds a similar condition into her supportive reasons norm (McKinnon 2015, 52).
[46] As in Turri (2011).
[47] I'm grateful to Tom Kelly for conversation on this point.

Moreover, once we have this Kantian distinction in mind, it's no longer a mystery why we should say that asserting things at random is bad even if it should happen to satisfy the norm of assertion (whatever that norm may be). Agents can be criticized not only for violating the norm but for being insensitive to the guidance of the norm in asserting. It's a general feature of norms that there are these two levels of criticism, and all norms of assertion will need to invoke this distinction (or one like it) to handle the general problem of random assertions.

6. Conclusion

We began by sketching an account of an often employed but rarely fully characterized concept: the position to know. The position to know depsychologizes away from the agent so that it is possible for an agent to be in a position to know what they do not believe. But the position to know shares many important features with knowledge: importantly, being in a position to know that p entails that p is both true and knowable.

This unique combination suggests a promising norm of assertion: Agents should assert that p only if they are in a position to know it. This norm gets the right verdict in selfless assertions, lotteries, and Moorean paradoxes by balancing Williamson's (etc.) insight that knowledge is central to assertion with Lackey's insight that assertion is other regarding. The PtK-Norm thus foregrounds the public nature of assertion that allows asserters the flexibility to act in the best interests of their audience when internal psychological pressures would otherwise prevent them from communicating the knowable truth.[48]

P, BUT YOU DON'T KNOW THAT P

Abstract: It's sometimes assertorically acceptable to tell an audience, "*P*, but you don't know that *p*." This can seem puzzling: after all, one can never *inform* an audience of the asserted content by speaking thus. Nevertheless, such assertions can be conversationally useful, making it apparent where there are unbridgeable gaps in the perspectives of the speaker and audience. Such assertions also make trouble for the growing family of views about the norm of assertion that what licenses proper assertion is not the initiating epistemic position of the speaker but the (potential) resulting epistemic position of the audience. They also allow us to draw a distinction between the aims of assertion and its norm.

> *"There's something you must remember: You're braver than you believe,*
> *and stronger than you seem, and smarter than you think."*
>
> —*Christopher Robin, to Winnie the Pooh*

1. Introduction

Assertion is our most straightforward practice for informing others about the

world. To assert that *p* is (at least) to say a sentence that expresses *p* (in the given

context) in the right sort of way. As so often in philosophy, "the right sort of way" turns

out to be terribly difficult to characterize. But since Williamson (2000), some

philosophers have thought that part of the story involves assertion being subject to a

special, epistemic norm—at least a norm that applies to assertions that are straight-faced

and unhedged, spoken in ordinary contexts.[1] This paper, which belongs to that tradition,

---

[1] Qualifications of this sort appear as early as Williamson (2000) who excludes from his argument assertions with "a special jocular tone" (2000, 246). Goldberg (2015) argues that the default knowledge norm doesn't apply in conditions of diminished epistemic hope, and Fleisher (forthcoming) argues that assertion in the realm of certain research projects parallel norms for endorsement rather than belief. Benton and van Elswyk (2020) argue compellingly that we can weaken the obligation to follow a strong norm of assertion by hedging. It's consistent with all this that there is a default norm of assertion that applies unless there is some special explanation for why it does not in a given case or domain. This default interpretation of the norm of assertion will be the working hypothesis of this paper. This differs from some theorists, e.g., Kemp (2007) who suggests that the norms of assertion may vary across contexts according to a patchwork of conversational practices. I will make no assumptions about whether the norm of assertion should be understood as constitutive or merely regulative.

argues that second-person Moorean assertions make trouble for the growing family of views about the norm of assertion: that what licenses proper assertion is not the epistemic position of the speaker but of the audience.

But first, why think that the norms of assertion reference audiences in the first place? In direct challenge to the knowledge norm, and other norms requiring the speaker to believe what they assert, Jennifer Lackey proposes that in cases of selfless assertion, the asserter can help their *audience* come to believe what they themselves cannot. Consider Sebastian, the distraught doctor:

> DISTRAUGHT DOCTOR: Sebastian is [a] ...pediatrician [who] recognizes ...that all of the scientific evidence shows that there is absolutely no connection between vaccines and autism. However, shortly after his ...daughter received one of her vaccines, ...she was soon diagnosed with autism. …[T]he grief and exhaustion brought on by his daughter's recent diagnosis cause him to abandon his previously deeply-held beliefs regarding vaccines. …[W]hile performing a well-baby checkup..., the child's parents ask him about ...the rumors surrounding vaccines and autism. Recognizing both that the current doubt he has towards vaccines was probably brought about through [his] emotional trauma ...and that he has an obligation to his patients, ...Sebastian asserts, "There is no connection between vaccines and autism" …[although it's false that] Sebastian himself believes or knows this proposition. (Lackey 2007, 598–99)

Lackey concludes that "it is a mistake to require proper assertion to pass through the doxastic states of the asserter" (Lackey 2007, 600). Doffing the belief requirement on assertion allows speakers like Sebastian to help their audiences learn despite the epistemic limitations of the speaker. What Lackey recognizes is that assertion is, to some extent, *for* audiences—at least in the normal case.[2] Lackey thus proposes a norm of assertion that is *audience-accommodating*. It allows speakers to assert what

---

[2] This isn't to deny that some assertions are spoken out of a need to express oneself, regardless of whether any audience is present. Or that one's audience may be oneself, as, perhaps, when one writes in a diary. But paradigmatic cases of assertion involve an external audience. (Thanks to conversation with Matt McGrath here.)

they don't believe (in certain cases) if doing so promotes the epistemic well-being of

the audience:

> RTB-Norm: One should assert that *p* only if it is reasonable for one to believe that
> *p*. (Lackey 2007, 608)[3]

Since Lackey, several authors have crafted norms with the intent to capture her audience-

accommodating insight:

> SRN-Norm: One may assert that *p* only if one has supportive reasons for *p*.
> (McKinnon 2015, 4)[4]

> PtK-Norm: One should assert that *p* only if one is in a position to know that *p*.
> (Willard-Kyle 2020, 328)[5]

Unsurprisingly, disagreement persists: some philosophers think that our judgments about

Lackey's selfless assertions are misleading (e.g. Kvanvig 2011, 235) or that her cases can

be explained without giving up on belief-entailing norms (Sosa 2011, 47; Turri 2015).[6]

But in this paper, we'll be interested not in the objection that Lackey's audience-

accommodation goes too far but that it doesn't go far enough. Note that all three of the

audience-accommodating norms just considered still maintain that proper assertion

depends on the epistemic position of the speaker. One should assert only if it is

reasonable *for the speaker* to believe, if the *speaker* has supportive reasons, if the *speaker*

---

[3] Lackey's full formulation of the RTB-norm includes this second condition: "if one asserted that *p*, one would assert that *p* at least in part because it is reasonable for one to believe that *p*" (Lackey 2007, 608). She also complements this norm with a Not Misleading Norm of assertion. But what matters here is that the epistemic requirement only obliges speakers to say what it is reasonable *for* them to believe, not what they, in fact, reasonably believe.

[4] Again, McKinnon's full formulation includes two additional clauses:
ii. The relevant conventional and pragmatic elements of the context are present, and
iii. One asserts that p at least in part because the assertion that p satisfies (i) [that is, the supportive reasons condition] and (ii). (McKinnon 2015, 4)
But what matters here is that the epistemic requirement only obliges speakers to *have* supportive reasons, not to believe on that basis.

[5] I define the position to know this way: "S is in a position to know that *p* iff S could know that *p* given their actual epistemic position," (Willard-Kyle 2020, 330). Importantly, this conception of the position to know is factive (2020, 331). Cf. Williamson (2000, 95).

[6] I tackle these objections directly in Willard-Kyle (2020, 334–46).

is in a position to know. But once we've reoriented ourselves around the idea that

assertion is (typically) aimed at an audience, we may start to wonder whether it isn't the

epistemic position of the *audience* that really licenses proper assertion.[7,8]

This general strategy has been recently advanced by Manuel García-Carpintero,

Charlie Pelling, and Edward Hinchman. Their norms are not just audience-

accommodating but *audience-centric*:

> TKR: One must ((assert *p*) only if one's audience comes thereby to be in a
> position to know *p*) (García-Carpintero 2004, 134).

> PKA: One's assertion that *p* is proper only if it is fit to give a hearer knowledge
> that *p* (Pelling 2013, 297).[9]

> PTW: Assert [that *p*] only when you could provide testimonial warrant [that *p*] to
> a potential addressee (Hinchman 2020, 556).

Broadly speaking, these suggestive alternatives to a speaker-oriented literature echo the

Gricean theme that assertions aim to produce a doxastic attitude in the hearer.[10,11]

Moreover, they perform surprisingly well on test cases forged in a speaker-centered

---

[7] Thanks to D Black and Sandy Goldberg for helping me to appreciate the force of this question for audience-accommodating norms.

[8] Although speaker-centric and audience-centric norms are those most widely represented in the literature, they are not exhaustive. For instance, one could hold that it is the epistemic credentials of relevant communities (potentially including both the speaker and audience) that matters for assertion. Thanks to Jennifer Lackey for pointing out the breadth of the space of options.

[9] In later work, Pelling modifies this view, claiming that an interpersonal norm such as knowledge provision is better understood as a norm for the speech act of telling than the speech act of asserting, a distinction he makes in Pelling (2014). (He notes his change of view in 2013, 348, fn. 9.) The distinction between telling and asserting will be addressed in §3.

[10] Grice describes proper assertion as essentially aiming at producing a mental state in the audience: "A meant$_{NN}$ something by x" is roughly equivalent to "A uttered x with the intention of inducing a belief by means of the recognition of this intention" (Grice 1957, 384).
There are differences, to be sure, between Grice's view and the norm-oriented approach favored by Hinchman, Pelling, and García-Carpintero, and I don't mean to claim that any of these three must be interpreted as *Griceans*. Grice focuses on belief not knowledge; none of the norms under consideration is committed to the view that meaningful utterances are accompanied by intentions-recognizing intentions; however, all three share a commitment to the idea that assertion (meaningful utterance) is aimed, in some sense, at producing a doxastic state in the audience.

[11] See Benton (2016), however, for a defense of the claim that, properly read, Grice is an ally of the knowledge norm.

context. Note, for instance, that one cannot put someone in a position to know that they've lost the lottery on merely probabilistic grounds.[12] And they get the right verdict in Lackey's selfless assertion cases, since none of the norms requires any particular doxastic attitude in the speaker.

It's a little trickier to determine whether or not audience-centric norms can explain the infelicity of Moorean paradoxes—utterances of the form, "p, but I don't know that p."[13] But there's at least a case to be made on their behalf. Suppose that the knowledge one would gain from testimony is defeated if the audience learns that the speaker doesn't know what they've asserted.[14] Then asserting, "p, but I don't know that p" would never generate knowledge that p in one's audience, since one would always be giving them a defeater for p along with the testimony that p.[15,16] Arguably, then, these audience-centric norms get the right verdicts in the cases that have most vexed the speaker-centric literature. And so, there's reason to think that switching the focus of assertions from speakers to audiences is exactly what the literature needs.

Nevertheless, making this switch gets us into trouble. In particular, it falsely predicts (I will argue) that second-person variants of Moorean paradoxes are always

---

[12] See Williamson (2000).

[13] Here, I simply assume the majority report that first-person Moorean propositions are not properly assertable. One notable way to push back against the majority view is to focus on assertions in contexts where guessing seems appropriate: "The Colts are going to win the Super Bowl. Of course, I don't know that for sure, but that's my take." For papers on how a knowledge-friendly norm of assertion might handle such cases, see especially Benton (2012) on prediction and Benton and van Elswyk (2020) on hedges.

[14] This is consistent with the possibility that a hearer can acquire knowledge from a speaker who in fact does not know what they assert (so long as the hearer doesn't *know* that the speaker is ignorant). That's good news since the possibility of knowledge from ignorant testifiers is crucial for Lackey's cases of selfless assertions. Cf. Lackey (1999, 2007).

[15] I owe thanks to conversation with Sandy Goldberg on this point.

[16] This move does, however, raise some hard questions for audience-centric norms: *Why* does learning that the speaker doesn't know, believe, etc. what they assert act as a defeater? Plausibly, because there's a presumption that the speaker has high epistemic standing for what they assert. But whence this presumption if the norm of assertion is silent on the epistemic standing of the asserter? For the most direct treatment of this challenge, see Hinchman (2013).

infelicitous. But although such expressions sometimes have an air of self-defeating futility, some second-person variants of Moorean Paradoxes are perfectly acceptable. These cases make trouble for audience-centric norms and also show the limitations of a Gricean view whereby the primary intention guiding assertion is to bring about in one's audience the belief that the asserted proposition is true (Grice 1957).

This paper, then, has two closely related goals. The first is to argue that even if—with Lackey, McKinnon, and myself—one rejects the view that proper assertion requires speaker-belief, one should hold onto the traditional view within the literature on the norm of assertion that it is epistemic position of the *speaker* that licenses assertion rather than of the *audience*. Assertoric license comes speaker-side, not audience-side. The second goal is to explore second-person variants of Moorean paradoxes—too long upstaged by their first-person counterparts—and to suggest that they often contribute to a conversation despite necessarily failing to inform their audiences by making perspectival gaps between addresser and addressee salient. Thinking about uninformative assertions will also give us the chance to reflect on the relationship between *aims* and *norms*: good norms allow us to systematically achieve our aims, but that doesn't mean that the content of our aims and our norms will be particularly similar.

In §2, I present felicitous assertions of second-person variants of Moorean paradoxes and discuss how they can be conversationally useful. In §3, I argue that these (and related) assertions pose a difficult challenge to audience-centric norms of assertion before turning to objections in §4. In §5, I reflect on lessons learned about the relationship between the aims of assertion and its norm(s).

2. What Can't Be Taught

This section has two goals. First, I argue that some second-person Moorean assertions are properly assertable. This will set up §3 in which I argue that the existence of felicitous second-person Moorean assertions is at odds with extant audience-centric norms of assertion: second-person Moorean assertions never put their audiences in anything like a position to know the asserted content. Second, I argue that second-person Moorean assertions can be conversationally useful in virtue of making salient the presence of two distinct perspectives, the addresser's and the addressee's. Reflections on the usefulness of Moorean paradoxes will inform the discussion of assertion's aims in §5.

2.1 Moorean Sentences

I can't know that "*p*, but I don't know that *p*." If I did, then I'd know *p*. But if I knew that *p*, I couldn't know that I *don't* know that *p*. So, I can't know both conjuncts together.[17] Similarly, on the assumption that knowledge entails belief, I can't know that "*p*, but I don't believe that *p*." If I did, then I'd know *and so believe* that *p*. But if I knew that *p* then I'd believe it, and so I couldn't know that I *don't* believe that *p*. Nevertheless, you, reader, can know the proposition, "*p*, but *you*"—that is, I the author—"don't know (or believe) that *p*." After all, there are many true things that I don't know. And there is no contradiction in *your* knowing that *p* and *your* knowing that I don't know it. In short, people cannot know of themselves that "*p*, but I don't know that *p*," but they can know of others, "*p*, but you don't know that *p*." The same goes for states that are entailed by knowledge, such as "*p*, but you don't believe that *p*," or "*p*, but you're not justified in

---

[17] See Moore (1942) and Williamson (2000).

believing *p*," and so on. We're necessarily better at detecting present false belief in others than we are in ourselves.[18]

Moorean paradoxes and their ilk show that some propositions are *partial*: they are knowable by some but not by others. Partial propositions make trouble for information-sharing. As concerns knowledge, second-person Moorean sentences are determined to fall on those without ears to hear: what is non-paradoxical in the mouth of the speaker is transformed into a paradox in the ear of the addressee. For while *I* can know <*p* and you don't know that *p*>, *you* can't know <*p* and you don't know that *p*>. You can't know both conjuncts together.

Nevertheless, it is (I will argue) sometimes perfectly acceptable for speakers to utter second-person Moorean propositions. This is somewhat surprising. What could speakers hope to accomplish by telling their audience something they could never learn? We'll discover that such assertions often manifest the presence of two *perspectives*: the asserter's and the audience's. They make clear that the asserter and speaker have different *sets of belief* and also that different *sets of potential beliefs* are rationally or psychologically available to them. Making these perspectives salient is a valuable tool for shaping the future direction of a conversation.

2.2 Felicitous Second-Person Moorean Assertions

Let's now consider acceptable second-person Moorean assertions. First, imagine a conversation with a climate change denier. Normally, when conversationalists disagree, it's appropriate for disputants to share their reasons to try to resolve the dispute before

---

[18] The relevance of second-person Moorean sentences has been strangely muted in a literature on norms of assertion that is mesmerized by their first-person counterparts. One exception is Dennis Whitcomb (2013) who considers second-person variants in an illuminating way to reflect on the nature of evidence.

moving on: we tend to seek as wide a common ground as possible. But let's suppose that in this case you can tell that a conversation about climate change will be unproductive. Your conversation partner is showing bad epistemic faith by blatantly ignoring evidence—they are actively, not just passively, ignorant.[19] It becomes clear that a belief in the truth of global warming is not psychologically accessible for them. Although you yourself are quite confident that climate change is a real phenomenon, your belief is based largely on the testimony of the scientific community: you don't have access to much first-order evidence that someone who distrusts the scientific community might find persuasive. And, frankly, you're tired and aren't likely to be charitable enough to have a pleasant or productive conversation. After listening to a couple of minutes of your friend's arguments, you finally assert:

1. Global warming is really happening, even though you don't believe it.

There's no way your friend can come to know that they themselves don't believe in global warming *and* that it's real. But (1) is fine to assert, nonetheless.

If you haven't informed—or even attempted to inform!—your friend, what, if anything, have you accomplished? You've expressed your public perspective. By asserting, you have declared a *stance*: you're *committed* to the world's being a certain way, a way that involves your interlocutor's being ignorant of global warming. When we assert, we do (at least) two things. We put forward the asserted content with assertoric force. But we also represent ourselves as being assertorically committed to it. We put ourselves or our assertoric authority behind the assertion. This is why we can be properly blamed when we assert falsely or baselessly without excuse.

---

[19] Cf. Mills's distinction between passive and active ignorance (as applied to white ignorance): "Imagine ...an ignorance that is active, dynamic, that refuses to go quietly" (2007, 13).

Often, we assertively express our stances in order to get our interlocutors to agree with us about the asserted content. That cannot be the goal in (1), but that does not mean the assertion is futile. What, then, have you accomplished? That answer may depend on the specific relationship between you and your friend, but one thing you may have done is to change the direction of the conversation, or perhaps even to end it. You've expressed a public stance that is *unshareable* with your interlocutor. By doing so blatantly, you make it clear that you have renounced the goal of changing your friend's mind on climate change or of trying to see things from their perspective. This may be especially helpful if you, the speaker, judge that airing the hearer's objections would be counterproductive. Your perspectives are too far apart to beneficially engage in conversation on the subject, at least for the time being.

Before considering other cases, we should make an important clarification that bears on the argument in §3. I said that there's no way your friend can come to know the asserted proposition. But that's not quite right. Your friend can't come to know the asserted proposition *right now*.[20] Suppose that five years from now, your friend has ceased to be a climate change denier—they could then come to believe, perhaps partly on the basis of your past testimony, "Global warming is really happening, even though I *didn't* believe it (five years ago)." So, perhaps second-person Moorean sentences *are* capable of putting their audiences in a position to know. What's unique about them is that they don't put their audiences in a position to know them at the time of utterance.

---

[20] If propositions are tensed, then the initial example already shows that some propositions are assertable despite being unknowable for the audience, since even if the audience later learned, "*p*, but I *didn't* know that *p*," they wouldn't be learning the very same proposition. But let us grant the objector that propositions are untensed.

Indeed, some assertions of second-person Moorean propositions might be best understood this way.[21] But not all. Note that in the climate change case, even if it's possible that your friend ultimately comes to believe what you've asserted, getting your friend to know what you say just isn't the point of your assertion. You had given up hope of gaining a convert. Indeed, by expressing your perspective in this way and potentially ending the conversation, you are probably making it less likely not more that they will come to learn anything about global warming from you. That should make us suspicious that the assertion is ultimately licensed by its downstream ability to put the audience in a position to know.

But second, we can change the proposition so that your friend can never come to know it. Suppose you know your friend is set in their climate-change-denying ways, and that they'll never be persuaded. Hoping to avoid a recurrence of frustrating conversations, you assert:

1*. Global warming is really happening, even though you'll *never* believe it.

This new assertion is properly assertable even though your friend can *never* come to know (1*), even in the future: (1*) is unknowable for audiences at any time. I'll continue to focus on present-tense versions of Moorean paradoxes because doing so allows us to see a wider range of their conversational uses. But in each case, I will include an "eternal counterpart" via footnote. We will return to discussion about whether audience-centric norms can explain such assertions in §3.

---

[21] E.g. assertions of the form "You'll [only] understand that *p* when you're older" (assuming that understanding is factive).

Let's consider another example of a felicitous second-person Moorean assertion that has a different conversational purpose. Imagine this conversation between a conservative who watches Fox News and a liberal who believes Fox News is unreliable.

2. Liberal: Even Fox News says that Trump's justification for his immigration policy greatly exaggerates the risk of terrorism.

    Conservative: Wait a minute, I thought you didn't trust Fox News.

    Liberal: Of course, <u>Fox News is unreliable, but you don't know that</u>.[22] I'm trying to make an argument that persuades *you*.

Liberal is trying to give evidence that meets Conservative where they are. Instead of offering the argument that they themselves would find most convincing, they are giving an argument that Conservative would find persuasive. By taking up Conservative's perspective in this way, Liberal ends up citing some evidence that they believe is weak (the testimony from Fox News) because they think Conservative may take it to be strong evidence. When Conservative astutely notices that Liberal is citing an argument that Liberal wouldn't accept as persuasive (though of course, Liberal thinks the premises are *true*), Liberal responds with a second-person variant of a Moorean paradox. In doing so, they mark a sharp distinction between their two perspectives. This, in turn, allows the explanation that the argument is targeted solely at Conservative's perspective.

In our story, Liberal and Conservative are trying to arrive at common ground despite having very different political starting points. Liberal's second-person Moorean assertion allows them to signal where they want to find common ground (on immigration) and where they are, for the moment at least, willing to agree to disagree (on the reliability of Fox News).

---

[22] Eternal Counterpart: "Fox News is unreliable, but I know you'll never know that." Or "Fox news is unreliable, but I know you'll never be persuaded of that."

Recall that assertions typically invite challenge from one's audience. If Liberal had simply said, "Fox News is unreliable," then they would have invited dispute over whether or not Fox News is reliable. But by embedding that claim within a second-person Moorean sentence, Liberal effectively manages to *set the record straight* on Fox News while simultaneously *avoiding debate* about Fox News.[23]

But this isn't like the global warming case, in which one speaker asserts something that cannot be common ground in hopes that the conversation will end. Rather, by expressing an unshareable perspective about the reliability of Fox News (Conservative can't know that they don't know Fox News is unreliable *and* that it's unreliable), Liberal hopes to *continue* the conversation by *redirecting* it back to the topic of immigration. Once again, the perspective expressed by the assertion helps direct the future path of the conversation even though it (necessarily) fails to knowledgeably transmit the asserted content.

Not every second-person, Moorean-style assertion is quite so blatant. Recall Christopher Robin's conversation with Pooh:

3. You're braver than you believe (Geurs 1997).[24]

Less lyrically: "You're braver than degree X, but you don't believe that you're braver than degree X." This is something that Pooh can't know given Pooh's first-person perspective. In some ways, that's the point. Christopher Robin is *inviting* Pooh to see himself from a different perspective than his own: Through Christopher Robin's eyes, Pooh sees a braver bear. Of course, if Pooh does eventually manage to change their perspective in the right way, similar assertions will no longer be true: if Pooh believes he

---

[23] I'm grateful to conversation with Matt McGrath on this point.
[24] Eternal Counterpart: "You're braver than you could ever believe that you are."

is brave enough, it will no longer express something true to tell Pooh that he is "braver than he thinks." But the point of the assertion was to change Pooh's view of himself, not to get Pooh to believe the asserted proposition. What these cases show us is that sometimes assertion isn't (just) about communicating information: sometimes it's about influencing perspectives.

I've said that expressing a second-person Moorean sentence emphasizes the presence of two distinct perspectives, the speaker's and the audience's. Second-person Moorean sentences (or at least their eternal counterparts) are not rationally accessible to the people they are told to: they can only be received as paradoxes.

But this is only the first step. Once this distinction between perspectives is drawn, it can be put to various uses. In the global warming case, the speaker draws the distinction to forswear the goal of engaging with the audience's perspective. But in the Fox News case, the speaker draws the same distinction not to avoid engagement but to make it possible—to advance an argument that will appeal to the audience given their vantage point though it holds little appeal to the speaker from theirs. In both cases, the distinction is drawn so that the speaker can choose the terms of (dis)engagement with an audience whose perspective differs from their own.

Other times, the distinction is drawn not to tell the audience how the speaker is choosing to engage the audience's perspective, but to change the audience's perspective. Plausibly, this is what happens to Pooh. Christopher Robin's assertion is best understood as an invitation for Pooh to change his perspective on himself, to see himself through Christopher Robin's eyes. That is, after all, a more accurate perspective regarding Pooh's

bravery. In each of the cases, the primary goal is not to communicate information, but to direct, clarify, entrench, or change the perspectives operative in the conversation.

3. Against Audience-Centric Norms of Assertion

Following Williamson (2000), first-person Moorean paradoxes have been a central data point to explain (or explain away) in the literature on the norm of assertion. Second-person variants have largely been ignored. As we've seen, however, second-person variants of Moorean paradoxes are theoretically interesting. They have the intriguing twin properties of (i) necessarily failing to inform the intended audience of the asserted content and yet (ii) potentially being both permissible and conversationally useful assertions.

The felicity of some second-person Moorean sentences is, I will now argue, a mark against extant audience-centric norms of assertion. Recall the three audience-centric norms with which we began this paper:

> TKR: One must ((assert p) only if one's audience comes thereby to be in a position to know p) (García-Carpintero 2004, 134).

> PKA: One's assertion that p is proper only if it is fit to give a hearer knowledge that p (Pelling 2013, 297).

> PTW: Assert [that p] only when you could provide testimonial warrant [that p] to a potential addressee (Hinchman 2020, 556).

I address each in turn.

3.1 TKR: Transmission of Knowledge Rule

Suppose that one should assert that p only if one's audience comes thereby to be in a position to know that p (TKR). Should we then expect to be able to tell someone second-person Moorean sentences such as, "p, but you don't believe that p"? No, for it's impossible for one's intended audience to thereby learn that proposition. And if it's

impossible for an agent to know a proposition, then they are not in a position to know it.[25]

TKR wrongly predicts that second-person variants of Moorean propositions should be

unassertable.

3.2 PKA: Provision of Knowledge Account

Suppose that one should assert that *p* only if it is *fit* to give a hearer knowledge

that *p* (PKA). Should we then expect to be able to tell someone, "*p*, but you don't believe

that *p*"? It seems not. For plausibly if it's impossible for an agent to know a proposition,

then it is not fit to give them knowledge that *p*. Especially if the reason for the

impossibility is that the proposition is just not the sort of proposition that can be known

by the addressee, and not, for instance, that the agent just happens to lack the cognitive

resources or strength of evidence to know it.

But perhaps this depends on what exactly is meant by "fitness." Does what

Pelling means by "fitness" rule out the felicity of second-person Moorean assertions?

Here, things get tricky. Pelling defines fitness to give knowledge in terms of faultiness:

an assertion is fit iff it is not faulty (Pelling 2013, 303). But Pelling only gives us a

definition of *evidential* faultiness: an assertion is evidentially faulty iff something about

the "evidential basis explains why it doesn't give knowledge in the normal way" (Pelling

2013, 302). And the way in which second-person Moorean assertions are unfit to give

knowledge has nothing to do with evidence. Notice, for instance, that the evidence I (the

speaker) have for the proposition <*p*, but you don't believe that *p*> could be arbitrarily

---

[25] I defend this principle in Willard-Kyle (2020, 337–38).

strong: I (the speaker) could know it with rational certainty. But no matter how strong my evidence, you will never be able to know it.[26]

So, if second-person Moorean propositions are faulty with respect to giving an addressee knowledge, they are faulty in some non-evidential way. Pelling only discusses ways that assertions can be evidentially faulty or unfit to give knowledge. But since second-person Moorean sentences clearly can't give knowledge to their addressees, they seem to be faulty or unfit in *some* way: perhaps they are *propositionally* faulty in virtue of there being something about the proposition (or proposition-addressee pairing) that explains why it does not give knowledge in the normal way.

Although this explanation of the way that second-person Moorean sentences are unfit to give their addressees knowledge goes beyond Pelling's official account, it agrees with Pelling's overall gloss on the concept of fitness. Why introduce a notion of "fitness to give knowledge" in the first place? Because (Pelling rightly notes) an agent might assert that *p* very quietly, or in a language the addressee doesn't understand, or in an epistemic context in which there is too much misleading evidence for the addressee to take the speaker's word, and so on (Pelling 2013, 296). In *these* sorts of situations, it is *fit* to give the addressee knowledge. They *would* be in a position to know it in better auditory, conversational, or epistemic contexts even though, contingently, the addressee won't *actually* end up in a position to know the asserted content.

Intuitively, second-person Moorean propositions are radically unlike the sorts of cases that motivate Pelling's introduction of "fitness." It's a necessary feature of second-

---

[26] Indeed, as Whitcomb (2013) argues, the *addressee's* evidence for <p, but you [the addressee] don't know that p> could be arbitrarily strong (short of giving knowledge) and yet fail to put them in a position to know it.

person Moorean assertions—not a contingent fact about the local context—that addressees cannot come to know them. It seems, then, that such assertions should count as unfit to give knowledge.

And, indeed, elsewhere, Pelling argues that a certain proposition is fit to give knowledge because *if* the addressee forms a true belief in the normal way on the basis of the speaker's testimony that the proposition is true *then* the addressee would also thereby acquire knowledge (Pelling 2013, 310–311). But second-person Moorean propositions necessarily fail this test. Pooh could come to truly believe, "I am braver than I believe," not recognizing the incoherence this commits him to—he is, after all, a *silly* old bear— but Pooh would not thereby come to know the proposition, which is unknowable for him. That second-person Moorean sentences necessarily fail this positive test for fitness is further evidence that they are unfit to give an addressee knowledge.

Where does this leave us? We've been trying to evaluate whether or not it is fit to give knowledge of a second-person Moorean proposition to an addressee. I've suggested that the intuitive answer is *no*. After all, it's conceptually impossible for the addressee to come to know it. And to borrow a test that Pelling sometimes employs, if the addressee were to truly believe it on the basis of testimony they would not thereby count as knowing it. Admittedly, the reason knowledge isn't given in such cases doesn't have to do with evidential deficiencies, but that (I suggest) is because evidential deficiencies aren't the only way fitness to give knowledge can be revoked. Although second-person Moorean sentences are unfit for giving knowledge, they can be proper to assert. The PKA thus delivers the wrong verdict.

3.3 PTW: Provision of Testimonial Warrant

Let us now turn to Hinchman's norm: one should assert that *p* only when one can provide testimonial warrant that *p* to a potential addressee (PTW). There are two notable ways that Hinchman's norm differs from previous audience-centric norms: (1) the shift from *knowledge* to *testimonial warrant*, and (2) the shift from *actual* to *potential* addressee. Will either of these adjustments rescue audience-centric norms?

Hinchman shifts from knowledge to testimonial warrant, but there are a number of other epistemic properties weaker than knowledge that an audience-centric account could shift to in order to avoid the challenge from second-person Moorean assertions. I will address this class of views in detail in §3.4. But first, I'll say something brief about why we should be suspicious that the shift to warrant in particular will solve the problem.

Even though warrant is weaker than knowledge, it is plausible that it is conceptually linked to knowledge. Hinchman writes that "the core thought" behind his proposal is that "the normative aim of assertion is not simply to express knowledge but to give your addressee knowledge" (Hinchman 2020, 579, note 12). Warrant of the sort that can't even *potentially* put an addressee in a position to know doesn't do enough to properly promote the aim of giving the addressee knowledge. But if having testimonial warrant *does* entail at least the possibility of being in a position to know, then no addressee can have testimonial warrant for a second-person Moorean assertion. So, it seems that Hinchman's conception of warrant will either (a) be too weak to secure the core knowledge-transmitting thought behind his proposal or else (b) face the same challenge from second-person Moorean assertions as audience-centric accounts that feature knowledge.

We'll reconsider conceptions of warrant and related epistemic states more carefully in §3.4. For now, we turn to Hinchman's second refinement from *actual* addressee to *potential* addressee. In second-person Moorean assertions, although one's intended audience cannot come to know (or be warranted in believing) the asserted content, third parties can. For instance, although Pooh cannot learn (when addressed to him) that "You are braver than you believe," Piglet can learn the same proposition in a different guise: "Pooh is braver than *he* believes." So, perhaps Christopher Robin is licensed to tell Pooh "You are braver than you believe," not because *Pooh* (the actual addressee) could receive testimonial warrant but because *Piglet* (a potential addressee) could.

Relatedly, Hinchman makes a nuanced distinction between mere assertions and tellings. Mere assertions need not have any particular audience (as when one speaks aloud to oneself) although (according to Hinchman) they are still guided by our ability to inform a possible interlocutor. But tellings take a particular audience into account: "to count as telling A that *p* in full sincerity S must intend that A would be fully warranted in believing that *p* on S's say-so" (Hinchman 2020, 570).[27]

This maneuver successfully deflects the objection as (so far) developed for audience-centric accounts of assertion. But the problem quickly resurfaces in two ways. First, the cases of second-person Moorean assertions considered in this paper are not mere assertions but also tellings. When Christopher Robin says, "You are braver than you believe," he is *directly addressing* Pooh; he is not merely saying the sentence (as it so happens) in Pooh's presence. So, even if this audience-centric norm of (mere) assertion

---

[27] As noted earlier, Pelling makes a similar distinction and ultimately classifies his norm as a norm on telling rather than asserting (Pelling 2013, 348, fn. 9).

can allow for second-person Moorean sentences, the accompanying view about telling falters.

Second, we can imagine cases in which permissible second-person Moorean assertions rule out *any* potential addressee coming to learn the asserted content. Suppose, for instance, that Christopher Robin and Pooh were the only two people who ever existed or would exist. Pooh would be Christopher Robin's only potential addressee; nevertheless, it would remain assertorically proper for Christopher Robin to tell Pooh "You're braver than you believe."

Or consider Cassandra, to whom Apollo gave both the gift of prophecy and the curse of never being believed when prophesying. When Cassandra foretells the destruction of Troy, she asserts properly even though, given her curse, she cannot be believed. Or take the boy who cried wolf. If he has cried wolf loudly and widely enough, then when he finally does meet the wolf, he might properly (and tragically) assert:

4. There really is a wolf, even though no one (else) could ever (reasonably) believe me (that there is a wolf)!

This, too, is a second-person Moorean sentence, and so unknowable (and unable to produce testimonial warrant). But neither Cassandra nor the boy (who, in our versions of the stories, are wise to the predicaments they have landed in) can imagine any potential audience who would come to believe them by their say-so.

Hinchman nuances audience-centric norms in two interesting ways: shifting from knowledge to warrant and from actual addressee to potential addressee. The shift to a potential addressee is not sufficient to block the challenge from felicitous second-person Moorean assertions. Let's now turn to audience-centric views that focus on weaker epistemic requirements than knowledge.

3.4 Providing Less Than Knowledge

Let's take a step back. All extant proposals of audience-centric norms involve knowledge, at least indirectly. And that is what makes *unknowable* (for the addressee) second-person Moorean sentences so tricky to handle on these accounts. This is no accident: there's good reason to think that knowledge and assertion are, in some way, intimately related.[28] But there's conceptual space for audience-centric proposals that focus on some epistemic accomplishment less demanding than knowledge.[29] For instance, one could go in for this norm: Assert that *p* to some addressee A only if you could thereby make it *justified* for A to believe that *p*.

In this section, I'm going to suggest that we should be skeptical that weakening audience-centric norms of assertion in this way will be very helpful. There are two basic reasons for this. The first is that some weaker epistemic states that one might try to induce in one's audience are still tethered to knowledge in such a way that they won't be able to handle second-person Moorean assertions either. Consider again the concept of "warrant." I suggest that the following principle is true:

> **No Known Unknowability**: If S is warranted in believing that *p*, then it's *not* the case that: S knows <S is not in a position to know that *p*>.

Someone who knows that they aren't in a position to know that *p* yet believes *p* anyway believes *p* recklessly.[30] They have formed their belief despite knowing that it isn't good enough for knowledge. They are committed to a belief-version of a (first-person)

---

[28] Apart from the audience-centric proposals already considered, that knowledge is (in some way) intimately connected to assertion has been defended in Unger (1975), Williamson (2000), DeRose (2002), Benton (2011, 2016), Sosa (2011), Turri (2011), Goldberg (2015), Benton and van Elswyk (2020) and Willard-Kyle (2020).

[29] Thanks especially to Elise Woodard for raising this point.

[30] I mean that they have behaved *epistemically* recklessly. There may be pragmatic or moral reasons to believe what we know we cannot know.

Moorean paradox: they are doxastically committed to "*p*, and I don't know that *p*." At least they are committed to each conjunct.

As established, addressees are not in a position to know second-person Moorean assertions. And the addressees may well *know* that they are not in a position to know such assertions. In fact, we often *assumed* that addressees were aware that they couldn't know what the speaker was saying in our explanations for why such assertions could be conversationally useful. Since second-person Moorean sentences can be asserted even when their audiences know that they are unknowable for them, then if **No Known Unknowability** is a constraint on warrant, it's not true that all proper assertions can provide testimonial warrant to their addressees.

The point generalizes. For it's hard to ascribe positive epistemic status to agents who believe what they know they can't know. Consider the following analogues:

> **J-No Known Unknowability**: If S is *justified* in believing that *p*, then it's *not* the case that: S knows <S is not in a position to know that *p*>.[31]

> **Ra-No Known Unknowability**: If S is *rational* in believing that *p*, then it's *not* the case that: S knows <S is not in a position to know that *p*>.[32]

> **Re-No Known Unknowability**: If it is *reasonable for S to believe* that *p*, then it's *not* the case that: S knows <S is not in a position to know that *p*>.[33]

---

[31] Although Kvanvig defends a justified belief norm of assertion for *speakers*, he writes that the justification relevant to assertion is of the sort "that puts one in a position to know" (Kvanvig 2009, 149).

[32] Igor Douven defends a stronger variant of this principle: $RB_i(\varphi) \Rightarrow Cr_i(K_i(\varphi)) \geqslant Cr_i(\neg K_i(\varphi))$ (2009, 371). Roughly, rationally believing a proposition entails having a higher credence that one knows it than that one doesn't.

[33] Although explicitly working within idealized model, Williamson (2013) defends a principle that suggests several versions of **No Known Unknowability**: $Bp \vDash \neg K \neg Kp$ (if one believes that *p*, then one doesn't know that one doesn't know that *p*). Assuming <S knows that S is not in a position to know that *p*> entails <S knows that S does not know that *p*>, then $Bp \vDash \neg K \neg PtKp$. *A fortiori*, whenever one *justifiedly* or *rationally* or *reasonably* or *warrantedly* believes that *p*, one doesn't know that one is not in a position to know that *p*.

166

These are not entirely uncontroversial, and I don't have space to argue for each of these

here.[34] Nevertheless, the plausibility of each principle shows that it's not so easy for

audience-centric norms to avoid the challenge of permissible second-person Moorean

assertions just by swapping out knowledge for another epistemic category.

The second reason to be suspicious that switching epistemic targets would save

audience-centric norms is the prevalence of permissible second-person Moorean-*like*

sentences. Consider this general schema:

Never Will You Ever: *p*, even though never will you ever φ that *p*.

What I want to suggest now is that (as we've already seen in part), there are permissible

assertions of this general form that substitute something other than "know" for φ.

Consider these assertions:

5. Global warming is really happening, even though you'll never *believe/think*
   that's true.

6. You're been suckered into a conspiracy theory, and I fear there's no getting
   out. Global warming is really happening, even though you'll never *reasonably
   believe* it.

7. I see why, given the echo chamber that you are stuck in and the authorities you
   have trusted, you are not rationally entitled to believe that global warming is
   happening. Still, global warming is really happening, even though you'll never
   be *rationally entitled to believe* it.

Are such sentences a problem for all audience-centric norms of assertion that are weaker

than knowledge? Not necessarily in the same way that second-person Moorean assertions

are. The proposition "*p*, but you don't *know* that *p*" (and, for that matter, "p, but you

---

[34] Lackey (2007) seems committed to the falsity of **Re-No Known Unknowability**, for instance. Those
who think that belief is a very weak kind of commitment in the mode of Hawthorne, Rothschild, and
Spectre (2016) may think that it is correspondingly easy to have positively evaluable beliefs in things
known to be unknown.

don't *believe* that *p*") is *unknowable* for the addressee; but "*p*, but you don't *believe* that *p*" is not, perhaps, strictly *unbelievable* for the addressee.[35]

Still, I think such assertions are awkward for audience-centric accounts to explain. Suppose that some instantiation of the following norm is true:

Norm: Assert that *p* only if by so doing you will potentially enable your addressee to φ that *p*.

And suppose also that S could properly assert to an audience A a proposition with this form:

Never Will You Ever φ: *P*, even though never will you ever φ that *p*.

If S's assertion is true, then A will never ever φ that *p*. So, asserting that *p* to A would not potentially enable A to φ that *p*. So, according to **Norm**, asserting *p* to A would violate the norm of assertion.

But as we've seen, there are many acceptable ways to assert propositions that instantiate the **Never Will You Ever** schema. They don't crash. And the certainly don't crash as decisively as first-person Moorean assertions.

Does every possible audience-centric account fall prey to a version of this challenge? It's hard to say with certainty. There are a lot of values we could plug in for φ. Nor should any single piece of data singlehandedly rule a norm in or out of contention: there's a lot of linguistic data to accommodate. Nevertheless, it's striking that second-

---

[35] It seemed plausible earlier in the paper to suppose that Pooh stumbled into believing "I am braver than I believe," not recognizing the incoherence this committed him to. See also Sorensen (1985, 496) on nonobvious Moorean sentences.

But some (perhaps idealized) models *do* rule out the believability of *blatant* Moorean assertions (like "*p*, but you don't know that *p*"). Robert Stalnaker defends a principle for strong belief whereby: if S believes that *p*, then S believes that S knows that *p* (2006, 179). When combined with the principle that beliefs are consistent (2006, 179), Stalnaker's picture entails that "*p*, but I don't know that *p*" is not just unknowable but unbelievable. For if an agent did believe it, they would believe both that they know that *p* and that they don't know that *p*, and this is inconsistent.

person Moorean assertions (and variations thereof) tend to sound much better than their first-person counterparts. This is a serious mark against audience-centric views. Maintaining that the epistemic position of the speaker, not of the audience, licenses proper assertion allows us to preserve the insight that not all assertion is intended to communicate information. Sometimes, assertion can be productive by expressing the speaker's assertoric stance even when convincing the addressee is impossible.

4. Objections

In this section, I consider three objections: (1) that non-audience centric norms cannot explain why some second-person Moorean sentences *do* sound paradoxical, (2) that second-person Moorean "assertions" may belong to some other speech-act type, and (3) that second-person Moorean assertions are not strictly felicitous but rather (often) blameless violations of an audience-centric norm of assertion.

Objection 1: But many second-person Moorean sentences *do* sound paradoxical, or at least very odd. Imagine asserting, out of the blue,

8. ?It's raining, but you don't know that it's raining.

It's hard to make sense of this assertion. Audience-centric norms have the advantage when explaining why utterances like (8) sound odd.

Answer: One thing to note is that although (8) sounds decidedly odd in most contexts, it doesn't obviously sound odd in the same way that first-person Moorean assertions do. As Sorensen notes, "In the first person examples [of Moorean assertions],

the speaker seems to be contradicting himself. In the second person examples, the speaker seems to be saying something which is self-defeating" (Sorensen 1985, 491).[36]

I think that these two different phenomenologies—the self-contradictory phenomenology and the self-defeating phenomenology—point to two different explanations of infelicity. Moreover, while I think it is self-defeating in many contexts to utter a second-person Moorean assertion, there is an explanation for why second-person Moorean assertions are *not* self-defeating when speakers do note expect (or aim) to be believed. Assertions of second-person Moorean propositions that are *not* self-defeated in the relevant way can be felicitous and, thus, still speak against audience-centric norms of assertion.

For second-person Moorean assertions, much hangs on whether the audience is prepared to take the speaker at their word. Suppose that the speaker does expect to be taken at their word. Then asserting, "*p*, but you don't believe that *p*" is what Hintikka (1962, 90–91) calls an "antiperformatory" assertion: asserting the proposition makes it false. For assuming that the audience believes the speaker, the audience will come to believe that *p*. And then it won't be true that "you [the audience] don't believe that *p*." When the speaker is trusted, second-person Moorean assertions are self-defeating in the sense that asserting them has the immediate effect of making them false.

Similar observations apply to second-person Moorean assertions of the form "*p*, but you don't *know* that *p*." For, assuming that the audience trusts the speaker and believes that *p* on the speaker's say so, then if the speaker had any right to assert *p* and *p*

---

[36] Sorensen explicitly says this regarding propositions of the form <*p*, but you don't believe that *p*>. When later discussing propositions of the form <*p*, but you don't *know* that *p*>, Sorensen does say that there is a temptation to "dismiss [them] as contradictions," though less so than for first-person versions (1985, 503).

is true, the audience will typically come to know that p. So, if p is true, then the second-person Moorean assertion will be false in virtue of being uttered. And if p is false, then it will already have been false. Either way, the truth of the assertion is defeated.

But what about cases in which audience is not prepared to take the speaker's word on the subject? When audiences don't trust speakers, the self-defeating tendency of second-person Moorean assertions is defused. This is exactly where felicitous, second-person Moorean assertions can be found.

When second-person Moorean assertions are self-defeating, it is because when the speaker says p this has the effect of getting the audience to believe (and in the good case *know*) that p, and so the second conjunct of the Moorean assertion becomes false through utterance. When audiences aren't prepared to take the speaker's word for it, this effect is removed, and so second-person Moorean assertions do not become false through utterance.

Here is an example. In *The Lord of the Rings*, Frodo meets a mysterious character named Strider who claims to be a good guide to Rivendell—a place Frodo desperately needs to go. But cautious of accepting an unknown guide, Frodo demands of Strider an account of himself, to which Strider wryly responds, "[W]hy should you believe my story, if you do not trust me already? Still here it is" (Tolkien 1984, 178). Strider might just as well have said,

9. You won't be able to know what I'm about to tell you about myself (namely p), but p.

This is a second-person Moorean assertion. In context, (9) is a perfectly sensible thing to assert. And it's sensible to assert at least in part because Strider does not expect Frodo to believe the story about himself or to be able to know that his story is true just by his say-

so. And indeed, in each of our original cases, there is at least some doubt as to whether the speaker will be believed. The climate change denier won't stop being a climate change denier at their friend's say-so, for instance.

When audiences clearly expect to be believed, second-person Moorean sentences tend be awkward: they are self-defeating. But one thing we've learned from §2 is this: sometimes, we can properly assert even when we don't expect our assertion to be believed. If audience-centric norms of assertion were true, then *all* second-person Moorean assertions would be infelicitous. But at least when speakers expect not to be believed, they can felicitously assert second-person Moorean propositions.

Objection 2: The apparent assertions of this paper aren't assertions at all: they only look that way because they are in the indicative.[37] When I say, e.g., "Global warming is really happening, even though you don't believe it," I may be *stopping* the conversation or *expressing* my frustration that we can't communicate about a topic I care about, but I'm not asserting. Any *real* assertion (the objector insists) would involve aiming to get the audience to be positioned to know what I say thereby!

Answer: But even if the utterances are *stopping* or *expressing* something, they are accomplishing that task *by* asserting something. The reason the conversation (sometimes) stops or frustration (or another attitude) is expressed is because the audience takes the speaker to have really taken a stance on the asserted content—a stance that the addressee cannot share.

Moreover, there's good reason to think that second-person Moorean sentences can be genuine assertions. First, we typically take first and third person versions of Moorean

---

[37] I'm grateful to Laura Callahan and Sandy Goldberg for raising the possibility of this interpretation.

sentences to be assertions, so it would be odd if second-person variants were exceptional. If we *didn't* take first-person Moorean sentences to be assertions, their infelicity wouldn't shed any light on the nature of assertion. But the infelicity of first-person Moorean sentences is frequently interpreted as evidence for some norms of assertion and against others.

Second, uttering second-person Moorean sentences seems to enable knowledge by testimony to bystanders who overhear an assertion but are not among the audience. Suppose that Christopher Robin tells Pooh, "You're braver than you believe" in the presence of Piglet. As noted before, Pooh cannot learn the content of the sentence, but Piglet can. That is, Piglet can learn that "Pooh is braver than Pooh believes." Moreover, Piglet seems to learn this in the normal way that one acquires knowledge through (overhearing) testimony. And this is hard to explain if Christopher Robin has not genuinely asserted something with the content "Pooh is braver than Pooh believes." If Christopher Robin were merely *expressing*, "Hurray for Pooh's bravery," or if Christopher Robin were merely *recommending* that Pooh increase his estimation of his own bravery (as one could do for pragmatic reasons as well as veritistic ones), then it would be an awful coincidence that Piglet is in a position to learn, apparently by testimony, the very proposition Christopher Robin would otherwise be asserting.

Finally, Pooh can *object* to Christopher Robin's utterance as one would object to an assertion. In response to "You are smarter than you think," Pooh can reply, "No, that is not true: I really am a Bear of Very Little Brain." But "that's not true" is not a reasonable response to a (mere) expressive or recommendation. Note, too, that Pooh's objection lacks the sauciness of deliberate misinterpretation that typically accompanies

responses that target the grammatical mood of an utterance rather than its intended conversational force. By objecting, Pooh is *not* acting like the smart-aleck who passively responds to the order "Can you pass the salt?" by merely answering as though to a (genuine) question, "Indeed, I can," without bothering to reach for the shaker.

Objection 3: Second-person Moorean paradoxes are improper, as audience-centric norms indicate. But they are *excused*. That's why the examples in this paper sound ok even though they are, strictly speaking, improper.[38]

Answer: Many factive norms of assertion utilize some kind of distinction between primary and secondary impropriety (cf. DeRose 2002) or blameworthy and blameless violations of the norm (cf. Williamson 2000). And our ears are not always sufficiently attentive to the differences between utterances that are blameless violations and justified non-violations. So, the audience-centric theorist is well within their rights to argue that the examples of this paper are *excused* or *blameless* violations, but violations, nonetheless.

But in order for that response to be satisfying, we'd need an explanation why second-person variants of Moorean paradoxes are often excused. And it's not obvious what that explanation would be. Two prominent kinds of excuse are (a) ignorance that one is violating the norm or (b) overriding pragmatic or moral reasons to violate the norm. But (a) the protagonists in our example are not necessarily ignorant that they are telling someone something that they can't know; indeed, some of them are counting on it. So, ignorance is not an applicable excuse. And although moral and pragmatic reasons can rarely be fully excised from any example, (b) the cases of this paper do not appear to rely

---

[38] I'm grateful to Dennis Whitcomb for suggesting this objection.

on cases with trumped up pragmatic or moral stakes. Thus, the examples of this paper are best understood as fully proper assertions and not excused violations.

5. Aims and Norms

Second-person Moorean assertions make trouble for audience-centric norms of assertion. But those norms are trying to latch onto something important: the Gricean thought that assertion is *for* audiences.

In this section, I'll argue that respecting the difference between aims and norms allows us to preserve much of value from audience-centric proposals: audiences feature prominently in the *typical aims* of assertion, and these aims play a role in guiding the speaker-side norm. It's not the case, however, that it is always the aim of an asserter to *inform* an audience, nor (I will suggest) is there an *essential* link between assertion and some audience-involving aim or other.

Aims are importantly different from norms. Aims define success whereas norms include a notion of permissibility (Marsili 2018, 642).[39] Permissible failures and impermissible successes are both conceptually possible. Consider the norms that govern good driving: stop at stop signs, drive on the right side of the road, don't drive too fast, etc. None of these is an *aim* of driving. People do not drive *in order* to drive on the right side of the road. The aim is (generally) to get where you're going.

We can distinguish between the aims of a practice as a whole and the aim of a particular move within that practice (cf. Kemp 2007). My goal in swerving left at a particular moment may be to avoid crashing into another car. But my reason for making

---

[39] This is not to deny that overlap is possible. Success can be normatively required as when "failure is not an option." But overlap is not conceptually required.

that particular move is not my reason for driving. We can distinguish, too, between aims that are more or less intimately connected to a practice. Ishani Maitra explains:

> Any baseball player has a variety of purposes while playing the game. These may include: whiling away some time, hitting a home run, entertaining spectators, and winning the game, to mention just a few. However, one of these purposes—namely, winning...—seems more intimately connected to the game than the others, in the sense... that it is a purpose that baseball players have *qua* players of the game (Maitra 2011, 284–285).

We've distinguished norms from various kinds of aims. The state tells you *how* to drive not *where* or *why*. But that is not to say that the norms and aims are unrelated. The norms on driving do (generally) enable people to fulfill their (typical) driving aims in a fair and orderly way. Not having a norm to drive on one side of the road would lead to more accidents and stop people from achieving their aim of safely getting where they are going. The norm "Always drive in circles" wouldn't be a good norm precisely because that norm would undermine drivers' typical aim of getting folks where they're going.

The relationship between norms and aims is not always straightforward. Some norms for driving help the aims of pedestrians rather than drivers, for instance, by requiring drivers to stop at crosswalks occupied by pedestrians. And even the norms of driving that ultimately promote our driving aims in the long term often *frustrate* our driving aims in the short term. This is why people (too) often violate the driving norms in order to achieve their driving aims. At a surface level, at least, a speed limit is directly at odds with my aim to get somewhere quickly. The same goes for the norm to stop at red lights, even when there doesn't seem to be anyone else around. But by all abiding by the speed limit and stop-at-red-lights norms, we generally and collectively do better at safely fulfilling our driving aims than if we did not.

Just as the rules of the road help us (collectively and generally) achieve our typical driving aims—even though the aims are not referenced in the rules—so a speaker-side norm of assertion can help us achieve our typical audience-involving aims. Suppose, for instance, that the aim of a speaker in making a particular assertion is to inform an audience of the asserted content. Succeeding in this aim requires convincing the audience that the asserted content is worthy of their belief. [40] If the practice of assertion includes a norm that prohibits asserting content that is not belief-worthy for the speaker, then addressees are more likely to feel it is safe to believe the asserted content. At least if the norm is enforced well-enough and if the speaker is likely enough to abide by the norm. [41] After all, there is (defeasible) reason to think that if something is belief-worthy for you (the asserter), then it's belief-worthy for me (the addressee).[42]

I don't claim that this is the *only* story that one could tell linking the norms of assertion with (one of) its typical aim(s), namely, to inform audiences. In fact, I think a similar story is available to the audience-centric norms that I have otherwise been criticizing. The modest point here is that it's possible for speaker-side norms of assertion to nevertheless recognize that informing an audience is an aim, or even the principal aim, of asserters. As we've seen, the content of the norms of a practice don't need to be particularly similar to the content of its typical aims. And it turns out that speaker-side norms of assertion can be at least partly explained in terms of facilitating a practice that makes it easier for us to collectively achieve our audience-side aims. And so, if those

---

[40] I leave open for our purposes whether truth, reasonability, or knowledge, etc., qualifies as belief worthy.

[41] Whether the audience needs to know that I am likely to abide by the norm to be properly informed depends at least in part on debates between reductionists and anti-reductionists about testimony.

[42] This general story is not new. But one important feature for our purposes is that the story is told in terms of belief-worthiness rather than belief. This allows us to accommodate Lackey's cases of selfless assertion, in which speakers assert something that is not *believed* by the speaker (e.g. that vaccines do not cause autism) but is *belief-worthy* for the speaker.

speaker-side norms do a better job of explaining the linguistic data, then the mere fact

that audience-side aims go unmentioned in speaker-side norms is not a serious mark

against them.

Perhaps informing an audience (or putting them in a position to know) is the aim

most intimately connected with the practice of assertion.[43] Still, I think we should be

cautious of thinking that informing an audience is in any way *essential* to the practice of

assertion. Once a set of norms is in place, they can enable the pursuit of aims other than

the dominant aim (if there is one). Consider again the case of driving. If there is a

dominant aim of driving,[44] it is plausibly to get people where they want to go. Certainly,

we would not have developed sprawling highway systems and intricate vehicular

regulations if driving were not a very useful mode of transportation. But once in place,

one can pursue less typical aims while abiding by the norms of driving. One can joyride.

Or take a passenger where they *don't* want to go (like a child to the dentist). One can

even drive aimlessly, as a pedestrian might wander. Even if the norms of assertion are

principally meant to *facilitate* audience-centric aims, they can *host* a much wider variety

of speaker-aims.

As joyriding does for driving, second-person Moorean assertions showcase

appropriate, if perhaps atypical, aims that someone might have for making an assertion.

As argued in §2, they can help speakers to highlight a perspectival gap, to end a

---

[43] This is contentious. Marsili (2018) makes a plausible case that the aim of assertion is truth, a thesis also considered in Dummett (1973). But dialectically, we may grant it to those who find audience-centric norms of assertion appealing.

[44] And there may not be a dominant aim of driving. Is there an aim that all drivers have *qua* drivers in the way that all competitors have the aim to win *qua* competitors? Not obviously. As Maitra (2011) critically notes, the literature on the norm of assertion has been dominated by an analogy to the norms in games. It is hoped that thinking about the norms of a different kind of activity, such as driving, suggests different parallels.

conversation, to agree to disagree, to advance an argument that meets the audience where they are, or to change an addressee's perspective. When Christopher Robin tells Pooh, "You're braver than you (will ever) believe," he succeeds in the aim of *encouraging* Pooh even though he necessarily fails to *inform* him of the asserted content.[45]

The typical or dominant aim of that speakers have in assertion is indeed to communicate information to one's audience—perhaps even to put them in a position to know. But there can be a wide range of individual aims in making an assertion, not all of which need conform to the typical aim. One could hold onto the idea that although an individual may properly aim for any number of things in making an assertion, the aim of *the practice* is to inform audiences. This may be right. But if so, we should be careful not to enforce too much conformity between the aim of the practice and the (even cooperative, audience-driven) aims that speakers might have for employing that practice. Assertion may have developed as a practice for sharing information,[46] but the practice is elastic, allowing speakers to permissibly pursue a wide range of creative purposes.

6. Conclusion

Audience-centric norms like Pelling's, García-Carpintero's, and Hinchman's are a helpful corrective to a literature that has too often excluded concern for the hearer, but they do so at the expense of the speaker. Assertion is (typically) an activity for pairs: the speaker and the hearer. The *aims* of assertion are often hearer-directed, but assertoric *license* comes speaker-side: the speaker's epistemic position is the grounds that enables

---

[45] Of course, sometimes we (e.g.) encourage someone *by* informing them of what we say. But this cannot be the case here since Pooh cannot learn the asserted content.

[46] In comparing the practice of assertion to the practice of driving, I've suggested that there's a degree of contingency about what norms apply to the practice of assertion such as we have it. But I hope to remain neutral on whether those norms are constitutive of our practice or not. If they *are* constitutive of our practice of assertion, then I take it that it is also a matter of contingent fact that we developed a practice of *assertion* rather than another practice like it but different in essence.

proper assertion to the audience. Audience-accommodating norms—like Lackey's, McKinnon's, and mine—enable flexible transmission of information by not requiring that speakers attain the same epistemic success they hope to impart to their hearers; moreover, they do this while avoiding the false predictions that audience-centric norms encounter in second-person Moorean sentences by maintaining that it is the epistemic position of the speaker that licenses proper assertion.

Examining second-person Moorean sentences more carefully reveals that such assertions can help both speaker and hearer by making salient their distinct perspectives. Whether it is by stopping a conversation or starting one, inviting someone to change their perspective or daring them to do so, these assertions enable the speaker and hearer to negotiate the direction of their conversation across perspectival gaps. This in turn highlights the diversity of proper assertoric aims that are compatible with a simple epistemic norm.[47]

BIBLIOGRAPHY

Alston, W.P. 1971. "Varieties of Privileged Access." *American Philosophical Quarterly* 8 (3): 223–241.

Alston, W.P. 1980. "Level-Confusions in Epistemology." *Midwest Studies in Philosophy* 5 (1): 135–150.

Anderson, C. A., M. R. Lepper, and L. Ross. 1980. "Perseverance of Social Theories: The Role of Explanation in the Persistence of Discredited Information." *Journal of Personality and Social Psychology*, 39 (6): 1037–1049.

Anscombe, G.E.M. 1957. *Intention*. Basil Blackwell.

Antill, G. 2020. "Epistemic Freedom Revisited." *Synthese* 197: 793–815.

Archer, A. 2018. "Wondering About What You Know." *Analysis* 78 (4): 596–604.

Audi, R. 1994. "Dispositional Beliefs and Dispositions to Believe." *Noûs* 28 (4): 419–434.

Benton, M.A. 2011. "Two More for the Knowledge Account of Assertion." *Analysis* 71 (4): 684–687.

Benton, M.A. 2012. "Assertion, Knowledge and Predictions." *Analysis* 72 (1): 102–105.

Benton, M.A. 2013. "Dubious Objections from Iterated Conjunctions." *Philosophical Studies* 162: 355–358.

Benton, M.A. 2016. "Gricean Quality." *Noûs* 50 (4):689–703.

Benton, M.A. and P. van Elswyk. 2020. "Hedged Assertions." In *The Oxford Handbook of Assertion*, edited by S. Goldberg, 245–263. Oxford: Oxford University Press.

Berker, S. 2013a. "Epistemic Teleology and the Separateness of Propositions." *Philosophical Review* 122: 337–394.

Berker, S. 2013b. "The Rejection of Epistemic Consequentialism." *Philosophical Issues* 23: 363–287.

Black, C. 1971. "Knowledge Without Belief." *Analysis* 31 (5): 152–158.

Brown, J. 2013. "Infallibilism, evidence and pragmatics." *Analysis* 73 (4): 626–635.

Carrol, L. 1899. *Through the Looking-Glass and What Alice Found There*. Macmillan and Co., Limited.

Chisholm, R.M. 1982. *The Foundations of Knowing*. Minneapolis: University of Minnesota Press.

Clifford, W.K. (1877) 1999. "The Ethics of Belief." In *The Ethics of Belief and Other Essays*, edited by T. Madigan, 70–96. Prometheus.

Coates, A. 2012. "Rational Epistemic Akrasia." *American Philosophical Quarterly* 49 (2): 113–124.

Dahlback, M. Forthcoming. "Infinitely Permissive." *Erkenntnis*.

DeRose, K. 2002. "Assertion, Knowledge, and Context." *The Philosophical Review* 111 (2): 167–203.

Descartes, R. 1984. *The Philosophical Writings of Descartes, Vol. II.* Translated by J. Cottingham, R. Stoothoff, and D. Murdoch. Cambridge University Press.

Douven, I. 2006. "Assertion, Knowledge, and Rational Credibility." *The Philosophical Review* 115 (4): 449–485.

Douven, I. 2009. "Assertion, Moore, and Bayes." *Philosophical Studies* 144: 361–375.

Drake, J. 2017. "Doxastic Permissiveness and the Promise of Truth." *Synthese* 194: 4897–4912.

Dummett, M. 1973. *Frege: Philosophy of Language*. Harvard University Press.

Egan, A. 2008. "Seeing and Believing: Perception, Belief Formation and the Divided Mind." *Philosophical Studies* 140 (1): 47–63.

Fantl, J. and M. McGrath. 2002. "Evidence, Pragmatics, and Justification." *Philosophical Review* 111 (1): 67–94.

Fantl, J. & M. McGrath 2009. *Knowledge in an Uncertain World*, Oxford University Press.

Feldman, R. 2002. "Epistemological Duties." In *The Oxford Handbook of Epistemology*, edited by P.K. Moser. Oxford University Press.

Fleisher, W. 2018. "Rational Endorsement." *Philosophical Studies* 175: 2649–2675.

Fleisher, W. Forthcoming. "Endorsement and assertion." *Noûs*.

Firth, R. 1981. "Epistemic Merit, Intrinsic and Instrumental." *Proceedings and Addresses of the American Philosophical Association* 55 (1): 5–23.

Flores, C. and E. Woodard. Manuscript. "Epistemic Vigilance: In Defense of a Norm on Evidence-Gathering."

Foley, R. 1991. "Evidence and Reasons for Belief." *Analysis* 51 (2): 98–102.

Friedman, J. 2017. "Why Suspend Judging?" *Noûs* 51 (2): 302–326.

Friedman, J. 2018. "Junk Beliefs and Interest-Driven Epistemology." *Philosophy and Phenomenological Research* 97 (3): 568–583.

Friedman, J. 2019a. "Checking Again." *Philosophical Issues* 29: 84–96.

Friedman, J. 2019b. "Inquiry and Belief." *Noûs* 53 (2): 296–315.

García-Carpintero, M. 2004. "Assertion and the Semantics of Force-Makers." In *The Semantics/Pragmatics Distinction*, edited by C. Bianchi, 133–166. CSLI Publications.

Geurs, K., dir. 1997. *Pooh's Grand Adventure: The Search for Christopher Robin*. Burbank, CA: Walt Disney Home Entertainment. VHS.

Gillies, A.S. 2001. "A New Solution to Moore's Paradox." *Philosophical Studies* 105 (3): 237–250.

Goldberg, S.C. 2015a. *Assertion: On the Philosophical Significance of Assertoric Speech*. Oxford: Oxford University Press.

Goldberg, S.C. 2015b. "Recent Work on Assertion." *American Philosophical Quarterly* 52 (4): 365–380.

Goldman, A. 1999. *Knowledge in a Social World*, Oxford University Press.

Goldman, A. 2001. "The Unity of the Epistemic Virtues." In *Virtue Epistemology: Essays on Epistemic Virtue and Responsibility*, Abrol Fairweather and Linda Zagzebski (eds.), Oxford University Press: 30–48.

Goldman, H.S. 1978. "Doing the Best One Can." In *Values and Morals*, A.I. Goldman and J. Kim (eds.), D. Reidel Publishing Company, 185–214.

Grice, H.P. 1957. "Meaning." *Philosophical Review* 66 (3): 377–388.

Grice, H.P. 1969. "Utterer's Meaning and Intention." *Philosophical Review* 78 (2): 147–177.

Hall, R.J. and C.R. Johnson. 1998. "The Epistemic Duty to Seek More Evidence." *American Philosophical Quarterly* 35 (2): 129–139.

Harman, G. 1986. *Change in View*. Cambridge, Mass.: MIT Press.

Hawthorne, J. 2004. *Knowledge and Lotteries*. Oxford University Press.

Hawthorne, J., D. Rothschild, and L. Spectre. 2016. "Belief is weak." *Philosophical Studies* 173: 1393–1404.

Hernandez, I. and J.L. Preston. 2013. "Disfluency Disrupts the Confirmation Bias." *Journal of Experimental Social Psychology*, 49: 178–182.

Hinchman, E.S. 2013. "Assertion, Sincerity, and Knowledge." *Noûs* 47 (4): 613–646.

Hinchman, E.S. 2020. "Assertion and Testimony." In *The Oxford Handbook of Assertion*, edited by S. Goldberg, 555–579. Oxford University Press.

Hintikka, J. 1962. *Knowledge and Belief*. Ithaca: Cornell University Press.

Hintikka, J. 1976. *The Semantics of Questions and the Questions of Semantics*. Cornell University Press.

Horowitz, S. 2014. "Epistemic Akrasia." *Noûs* 48 (4): 718–744.

Humberstone, I.L. 1992. "Direction of Fit." *Mind* 101 (401): 59–83.

Jackson, F. 1973. "Is There a Good Argument Against the Incorrigibility Thesis?" *Australasian Journal of Philosophy* 51 (1): 51–62.

Jager, R. 1970. "Truth and Assertion." *Mind* 79 (314): 161–69.

James, W. 1907. *The Will to Believe and Other Essays in Popular Philosophy*. Longmans Green and Co.

Kahneman, D. 2001. *Thinking, Fast and Slow*. Farrar, Straus, and Giroux.

Kavka, G. 1983. "The Toxin Puzzle." *Analysis* 43 (1): 33–36.

Kelly, T. 2003. "Epistemic Rationality as Instrumental Rationality: A Critique." *Philosophy and Phenomenological Research* 66 (3): 612–40.

Kelly, T. 2014. "Evidence Can Be Permissive." In *Contemporary Debates in Epistemology*, edited by Steup, Turri, and Sosa, 298–313. Blackwell.

Kelp, C. 2014. "Two for the Knowledge Goal of Inquiry." *American Philosophical Quarterly*, 51 (3): 227–32.

Kelp, C. 2015. "Sosa on Knowledge, Assertion and Value." *Erkenntnis* 80: 229–237.

Kemp, G. 2007. "Assertion as a Practice." In *Truth and Speech-Acts*, edited by D. Greimann and G. Siegwart, 106–129. Routledge.

Klayman, J. 1995. "Varieties of Confirmation Bias." *The Psychology of Learning and Motivation* 32: 385–418.

Klayman, J. and Y. Ha. 1987. "Confirmation, Disconfirmation, and Information in Hypothesis Testing." *Psychological Review* 94 (2): 211–28.

Koehler, J.J. 1993. "The Influence of Prior Beliefs on Scientific Judgments of Evidence Quality." *Organizational Behavior and Human Decision Processes* 56 (1): 28–55.

Kopec, M. 2015. "A Counterexample to the Uniqueness Thesis." *Philosophia* 43: 403–409.

Kvanvig, J.L. 2003. *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press.

Kvanvig, J.L. 2009. "Assertion, Knowledge, and Lotteries." In *Williamson on Knowledge*, edited by D. Pritchard and P. Greenough, 140–160. Oxford University Press.

Kvanvig, J.L. 2011. "Norms of Assertion." In *Assertion: New Philosophical Essays* edited by J. Brown and H. Cappelen, 233–250. Oxford University Press.

Lackey, J. 1999. "Testimonial Knowledge and Transmission." *The Philosophical Quarterly* 49 (197): 471–490.

Lackey, J. 2007. "Norms of Assertion." *Noûs* 41 (4): 594–626.

Lasonen-Aarnio, M. 2010. "Unreasonable Knowledge." *Philosophical Perspectives* 24: 1–21.

Lebens, S. 2015. "God and his Imaginary Friends: A Hassidic Metaphysics." *Religious Studies* 51 (2): 183–204.

Levi, I. 1991. *Fixation of Belief and Its Undoing: Changing Beliefs Through Inquiry*. Cambridge University Press.

Lewis, D. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy* 74 (4): 549–567.

Lord, C.G., L. Ross, and M.R. Lepper.1979. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence." *Journal of Personality and Social Psychology*, 37 (11): 2098–2109.

Maitra, I. 2011. "Assertion, Norms, and Games." In *Assertion: New Philosophical Essays,* edited by J. Brown and H. Cappelen, 277–296. Oxford University Press.

Marsili, N. 2018. "Truth and Assertion: Rules Versus Aims." *Analysis* 78 (4): 638–648.

McGrath, M. Forthcoming. "Being Neutral: Suspension of Judgment, Agnosticism, and Inquiry." *Noûs*.

McGrath, M. Manuscript. "Epistemic Norms for Waiting to Make up your Mind (or Suspending Judgment)."

McKinnon, R. 2013. "The Supportive Reasons Norm of Assertion." *American Philosophical Quarterly* 50 (2): 121–135.

McKinnon, R. 2015. *The Norms of Assertion: Truth, Lies, and Warrant*. New York: Palgrave Macmillan.

Mills, C.W. 2007. "White Ignorance." In *Race and Epistemologies of Ignorance*, edited by S. Sullivan and N. Tuana, 13–38. State University of New York Press.

Moon, A. 2018. "The Nature of Doubt and a New Puzzle About Belief, Doubt, and Confidence." *Synthese* 195 (4): 1827–48.

Moore, G.E. 1942. "A Reply to my Critics." In *The Philosophy of G.E. Moore,* edited by P. Schlipp, 543–667. Evanston.

Moyer, M.C. and K. Syrett. 2019. "The Semantics of Questions." *WIREs Cognitive Science* 10 (1513): 1–10.

Moyer, M.C. and K. Syrett. Forthcoming. "The Study of Questions." *WIREs Cognitive Science*.

Nelson, M.T. 2010. "We Have No Positive Epistemic Duties." *Mind*, 119 (473): 83–102.

Olsson, E.J. and D. Westlund. 2006. "On the Role of the Research Agenda in Epistemic Change." *Erkenntnis* 65(2): 165–183.

Pelling, C. 2013. "Assertion and the Provision of Knowledge." *The Philosophical Quarterly* 63 (251): 293–312.

Pelling, C. 2014. "Assertion, Telling, and Epistemic Norms." *Australasian Journal of Philosophy* 92 (2): 335–348.

Plato. 2002. *Five Dialogues: Eutyphro, Apology, Crito, Meno, Phaedo*, translated by G.M.A. Grube and revised by J.M. Cooper. Hackett.

Platts, M. 1979. *Ways of Meaning*. Routledge and Kegan Paul.

Pollock, J. 1983. "How do you Maximize Expectation Value?" *Nous* 17 (3): 409–421.

Raleigh, T. 2017. "Another Argument Against Uniqueness." *Philosophical Quarterly* 67 (267): 327–346.

Raz, J. 1975. *Practical Reasons and Norms*. Hutchinson.

Reisner, A. 2009. "The Possibility of Pragmatic Reasons for Belief and the Wrong Kind of Reasons Problem." *Philosophical Studies* 145: 257–272.

Reisner, A. 2013. "Leaps of Knowledge." In *The Aim of Belief*, edited by Timothy Chan. Oxford University Press.

Reisner, A. 2014. "A Short Refutation of Strict Normative Evidentialism." *Inquiry* (5): 1–9.

Roberts, C. 2012. "Information Structure in Discourse: Toward an Integrated Formal Theory of Pragmatics." *Semantics and Pragmatics* 5 (6):1–69.

Roush, S. 2003. *Tracking Truth*. Oxford University Press.

Sapir, Y. and P. van Elswyk. Manuscript. "The Ignorance Norm."

Schaffer, J. 2007. "Closure, Contrast, and Answer." *Philosophical Studies* 133 (2): 233–255.

Schroeder, M. 2012. "The Ubiquity of State-Given Reasons." *Ethics* 122 (3): 457–488.

Sextus. 2000. *Outlines of Scepticism*. Edited by Annas, J. and J. Barnes. Cambridge University Press.

Shah, N. 2009. "The Normativity of Belief and Self-Fulfilling Normative Beliefs." *Canadian Journal of Philosophy* 39 (SI): 189–212.

Shoemaker, S. 1963. *Self-Knowledge and Self-Identity.* Cornell University Press.

Sorensen, R.A. 1985. "Pure Moorean Propositions." *Canadian Journal of Philosophy* 15 (3): 489–505.

Sosa, E. 2007. *A Virtue Epistemology: Apt Belief and Reflective Knowledge*. Oxford University Press.

Sosa, E. 2011. *Knowing Full Well*. Princeton University Press.

Sosa, E. 2015. *Judgment and Agency*. Oxford University Press.

Sosa, E. Forthcoming. *Epistemic Explanations: A Theory of Telic Normativity, and What it Explains*. Oxford University Press.

Stalnaker, R. 1978. "Assertion." In *Pragmatics*, edited by P. Cole, 315–32. New York Academic Press.

Stalnaker, R. 2006. "On Logics of Belief and Knowledge." *Philosophical Studies* 128 (1): 169–199.

Stanley, J. 2008. "Knowledge and Certainty." *Philosophical Issues* 18: 35–57.

Sylvan, K. 2018. "Veritism Unswamped." *Mind* 127 (506): 381–435.

Tolkien, J.R.R. 1984. *The Fellowship of the Ring*. Easton Press.

Turri, J. 2011. "The Express Knowledge Account of Assertion." *Australasian Journal of Philosophy* 89 (1): 37–45.

Turri, J. 2015. "Selfless assertions: some empirical evidence." *Synthese* 192 (4): 1221–1233.

Unger, P. 1975. *Ignorance: A Case for Scepticism*. Oxford: Oxford University Press.

Velleman, J.D. 1989a. "Epistemic Freedom." *Pacific Philosophical Quarterly* 70: 73–97.

Velleman, J.D. 1989b. *Practical Reflection*. Princeton University Press.

Velleman, J.D. 1992. "The Guise of the Good." *Noûs* 26 (1): 3–26.

Velleman, J.D. 2000. "The Possibility of Practical Reason." Oxford University Press.

Weatherson, B. Manuscript. "Do Judgments Screen Evidence?"

Weiner, M. 2005. "Must we know what we say?" *The Philosophical Review* 114 (2): 227–251.

Whitcomb, D. 2013. "One Wage of Unknowability." *Synthese* 190 (3): 339–352.

Whitcomb, D. 2017. "One Kind of Asking." *The Philosophical Quarterly* 67 (266): 148–168.

Willard-Kyle, C. 2017. "Do Great Minds Really Think Alike?" *Synthese* 194 (3): 989–1026.

Willard-Kyle, C. 2020. "Being in a Position to Know is the Norm of Assertion." *Pacific Philosophical Quarterly* 101 (2): 328–352.

Williams, B.A.O. 1966. "Consistency and Realism." *Proceedings of the Aristotelian Society, Supplementary Volumes*, 40: 1–22.

Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

Williamson, T. 2013. "Response to Cohen, Comesaña, Goodman, Nagel, and Weatherson on Gettier Cases in Epistemic Logic." *Inquiry* 56 (1): 77–96.

Williamson, T. 2014. "Very Improbable Knowing." *Erkenntnis* 79: 971–999.

Worsnip, A. 2018. "The Conflict of Evidence and Coherence." *Philosophy and Phenomenological Research* 96 (1): 3–44.