

THE EXPLANATORY POWER OF KK

Caleb Fawell

A thesis submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Philosophy in the College of Arts and Sciences.

Chapel Hill
2022

Approved by:

Ram Neta

James Pryor

Alex Worsnip

© 2022
Caleb Fawell
ALL RIGHTS RESERVED

ABSTRACT

Caleb Fawell: The Explanatory Power of KK
(Under the direction of Ram Neta)

The KK thesis says that if S knows that P, then S knows that S knows that P. Though controversial, it seems able to explain some otherwise puzzling data. For many, this is a strong consideration in its favor. In this essay, I will propose a different explanation of the data. My explanation is built around the following norm: one should assert that P only if one believes that one knows that P. Since this norm is more plausible than KK, the explanation I propose undermines the explanatory power of KK and so weakens the case for KK.

TABLE OF CONTENTS

SECTION 1: INTRODUCTION AND SETUP.....	1
SECTION 2: INTRODUCING A NORMATIVE REQUIREMENT.....	13
SECTION 3: THE POSITIVE PROPOSAL.....	19
SECTION 4: OBJECTIONS AND REPLIES.....	29
SECTION 5: CONCLUSION.....	31
REFERENCES.....	33

Section 1. Introduction and Setup

The KK thesis says that if S knows that P, then S knows that S knows that P. Though controversial, it seems able to explain some otherwise puzzling data. For many, this is a strong consideration in its favor.¹ In this essay, I will propose a different explanation of the data. My explanation is built around the following norm: one should assert that P only if one believes that one knows that P. Since this norm is more plausible than KK, the explanation I propose undermines the explanatory power of KK and so weakens the case for KK.

In this paper, I will focus on the argument for the KK thesis according to which it can (best) explain certain data. I do this because many hold that its ability to explain these data is a strong consideration in its favor, and so it would be noteworthy if there were an equally satisfying but less controversial explanation available.

Before examining the data which are commonly taken to support KK, it will be helpful to distinguish them from other data which, though similar, are not my current focus. Some of these involve sentences of the following forms:

- (1) It's raining, but I believe that it's not raining.
- (2) It's raining, but I don't believe that it's raining.

These are doxastic Moorean sentences. Following standard practice, we can call the first a “commissive” Moorean sentence and the second an “omissive” Moorean sentence. Moorean sentences are puzzling because they sound like contradictions, but are not.

¹ Notably Greco (2014), Greco (2015), Das and Salow (2018), and Cohen and Comesaña (2013). Even critics of KK seem to be somewhat moved by this line of thought (see Sosa (2009) and Holguín (2019)).

Moorean phenomena give rise to two puzzles. One is to explain why they sound infelicitous. Another is to show whether (and why) they are defective—that is, to go beyond explaining why they sound strange, and to explain what is in fact *wrong* with them. This might mean showing how they exhibit some kind of irrationality, arguing they manifest a failure to properly value knowledge, showing how they violate a certain norm which we have good reason to believe binds our practice of assertion, some combination of these, or something else entirely.

Understandably, these explanatory tasks are frequently run together: it is tempting to assume that explaining why an assertion is defective adequately explains why it sounds infelicitous. To be fair, these tasks are related, and it may be they are not *quite* as separate as I am suggesting. Still, this conception of the overall explanatory project gets at two important constraints on a full explanation of our reactions to Moore-style sentences: (a) the explanation should be phenomenologically and psychologically viable: it should explain the distinct *feeling* of infelicity which such sentences tend to cause, and why creatures like us could actually feel it; (b) the explanation should also explain what is in fact *wrong* with the sentences in question—for instance, why it is irrational (or otherwise defective) to assert them, or to believe the propositions they express (if nothing is wrong with the assertion or belief, it should give an error theory). As we will see, it is important to keep these questions distinct. For often, an explanation of what is defective about a certain assertion—for instance, that it violates a certain norm, or that it manifests a certain kind of irrationality—does not immediately explain why that assertion sounds infelicitous.

Moore's paradox is not my focus here, and I mention the sentence forms (1) and (2) only to set them aside. There is a third kind of Moorean sentence about which I have slightly more to say than (1) or (2), though. It is not my focus either, but paying some attention to it will help set

the stage for what comes later. This third kind of Moorean sentence is sometimes called an “epistemic” Moorean sentence:

(3) It’s raining, but I don’t know that it’s raining.

The two questions connected with doxastic Moorean phenomena—what is defective about such sentences, and why do they sound infelicitous?—are relevant for sentences like (3) as well.

I will not be focusing on sentences like (3) because there is reason to believe that we can explain both how they are defective and why they sound infelicitous without appealing to KK. For instance, one popular explanation involves noticing that whenever one asserts something, one represents oneself as knowing what one has asserted. But if I assert (3), then I represent myself as knowing that it is raining (*via* the first conjunct), while also asserting that I do *not* know this (*via* the second conjunct). What I represent contradicts what I assert. This is commonly taken to explain why (3) sounds infelicitous, and the violation of a knowledge norm like Williamson (2000)’s “one should assert P only if one knows P” is commonly taken to explain why such assertions are defective.^{2,3}

There is a fourth kind of troubling assertion which is importantly different from the three just advanced. Following Sosa (2009)’s terminology, we can call these “dubious assertions”.

They have the following form:

(4) It’s raining, but I don’t know if I know it’s raining.

² See DeRose (199: 597-598), Sosa (2009), and Williamson (2000: 253).

³ A different explanation of why (3) sounds infelicitous is that it simply violates Williamson’s knowledge norm. As it stands, though, such an explanation is incomplete. Why? Because all kinds of reasonable but false assertions sound felicitous, and yet violate Williamson’s norm. As it stands, this explanation is a good example of running together the distinct explanatory tasks outlined above. Of course, it is easy to modify this particular explanation and add that (3) not only violates Williamson’s norm, it does so a priori and straightforwardly. That, as an explanation, looks much better. The point is just that that modification was necessary. It is necessary not only to explain what is defective about a given infelicitous assertion, but to show how that defectiveness contributes to an explanation of why it sounds infelicitous (if it does at all).

This paper will focus on dubious assertions like (4).

The basic datum which many think KK can explain is that assertions like (4) both (a) sound infelicitous and (b) are, in some way, defective. My ultimate aim is to argue that an alternative explanation is available which relies on less contentious principles. Before I do that, though, I would like to focus on a few different ways one could try to use KK to explain (a) why assertions like (4) sound infelicitous.

The above explanation of why (3) sounds infelicitous does not immediately generalize to (4). Why not? Because although (4) is similar to (3) in that the speaker represents herself as knowing that P by asserting the first conjunct, (4) is unlike (3) in that, in the second conjunct, the speaker asserts that they do not know *whether* they know that P. Putting this formally: the speaker represents (*via* the first conjunct) that Kp , but asserts (*via* the second conjunct) that $\sim KKp$. While with (3) what the speaker asserts in the second conjunct contradicts what they represent as being the case in the first conjunct, this is not the case with (4).

This makes it easy to see why KK is such a tempting explanation of why (4) sounds infelicitous. Without KK, what the speaker represents by asserting the first conjunct (namely, Kp) is perfectly consistent with what the speaker asserts in the second conjunct (namely, $\sim KKp$). But if the KK thesis is true—that is, if Kp entails KKp —then Kp and $\sim KKp$ are inconsistent. This has the benefit of explaining (a) why (4) sounds infelicitous while keeping fairly close to the original explanation of why (3) sounds infelicitous. And if there is a knowledge norm on assertion, it also explains (b) why dubious assertions are defective.

As it stands, however, this explanation needs to be clarified. In particular, we need to clarify what kinds of “entailment” and “inconsistency” are at work in each of the explanations of why (3) and (4) sound infelicitous.

The original explanation of why (3) sounds infelicitous began by noticing that the speaker represents that Kp while also asserting that $\sim Kp$. These two contents are formally inconsistent. Plausibly, this is why (3) sounds infelicitous. Is there a way to use KK to show that what one represents in (4) is *formally* inconsistent with what one asserts? Probably not. In (4), one represents that Kp (*via* asserting the first conjunct), but one asserts that $\sim KKp$ (*via* the second conjunct). For these to be formally inconsistent, Kp would need to *formally* entail KKp . But even if some version of the KK thesis is true, this is surely false—for it is surely false that Kp entails KKp in the same way that $P \ \& \ Q$ entails P .

What if we depart somewhat from the explanation of why (3) sounds infelicitous, appeal to the KK thesis as a matter of a priori metaphysical necessity, and argue that in asserting something like (4), what one represents in asserting the first conjunct (namely, Kp) is a priori metaphysically inconsistent with what one asserts in the second conjunct (namely, $\sim KKp$)? Perhaps (4) sounds infelicitous because what one represents is a priori metaphysically inconsistent with what one asserts.

The problem with this explanation is that it proves too much. Consider an assertion of some complicated mathematical falsehood P . Such an assertion might sound perfectly felicitous. But then what the speaker represents in asserting that P (namely, Kp) is a priori metaphysically inconsistent with what she asserts (namely, P) because P is, itself, metaphysically inconsistent. If metaphysical inconsistency holding between what one asserts and what one represents is enough to generate infelicities, then the graduate student's assertion should sound infelicitous—but by hypothesis, it does not.⁴

⁴ Perhaps a more plausible theory would hold that if what one represents is a priori metaphysically inconsistent with what one asserts *but each are a priori metaphysically possible on their own*, then what one asserts will sound infelicitous. That does not seem totally right either, though. For suppose I am an engineer and yesterday, I calculated that a certain bridge is strong enough for my friend's truck to drive over it. Today, forgetting exactly which

So there is probably not a straightforward way to extend the explanation of why (3) sounds infelicitous to a KK-involving explanation why (4) sounds infelicitous. But this does not mean there is no way that KK could explain this. Indeed, Daniel Greco uses KK to offer an influential explanation of why dubious assertions like (4) sound infelicitous in his (2014). Unlike the above explanations which appeal to the idea that speakers represent that they know P whenever they assert that P, Greco's explanation builds on a knowledge norm of assertion according to which one may assert P iff one knows P. If this norm is correct and if KK holds, then one may assert that P...

...iff one knows that P (by Greco's knowledge norm), which is true...

...iff one knows that one knows that P (by the KK thesis), which is true...

...iff one may assert that one knows that P (by Greco's knowledge norm, again)

In other words, if KK is true and Greco's knowledge norm on assertion is correct, one may assert that P iff one may assert that one knows that P.

Thus if KK is true and Greco's knowledge norm is correct, one's permission to assert "P" is linked to one's permission to assert "I know that P." There is then, Greco claims, an obvious problem with assertions like (4). In such cases one asserts that P but, when prompted, refuses to assert that one knows that P: but then one must either have violated the knowledge norm of assertion in asserting that P (in the first conjunct), or one must be "*unnecessarily reticent* in refusing to go on to assert" that one knows that P by, in the second conjunct, confessing instead

calculations I performed yesterday, I come to believe that I relied on mathematical claim X in calculating the bridge's strength. In fact, X is false, and I did not rely on it at all when calculating the bridge's strength. Thinking I relied on X, the following would be perfectly coherent for me to assert: "The bridge is strong enough to support your truck's weight, but of course, if theorem X is false, I don't know that." What I represent with the first conjunct is Kp , and what I assert in the second conjunct is "if theorem X is false, then $\sim Kp$ ". Each of these are metaphysically possible. But since it is metaphysically necessary (a priori) that theorem X is false, together, they are metaphysically inconsistent. And yet the sentence sounds perfectly felicitous (at least to my ears), and does not seem to manifest any irrationality on the part of the asserter.

that one does not know whether or not one knows that P (Greco (2014: 668), emphasis mine).

Furthermore, if there is a knowledge norm on assertion, KK can explain (b) why assertions like (4) are defective. If KK is true, such assertions are unknowable.⁵ They will thus be defective because they violate whatever knowledge norm one has in mind.

Some take this point to also explain (a) why assertions like (4) sound infelicitous. Note, though, that this will give a complete explanation of why certain assertions sound infelicitous only if we also claim that it is sufficiently obvious, a priori, that this is what is going on (otherwise, any reasonable falsehood would sound infelicitous, since it is impossible to know a falsehood). Presumably, this depends on the claim that KK is itself a sufficiently obvious a priori truth. But as we will soon see (and as the fact that there is a wide literature on KK suggests), there is reason to believe it is not.

Many agree that with some plausible background assumptions, KK can explain why assertions like (4) sound infelicitous. Of course, different people work out the basic insight in different ways, but the important point is that Greco is not alone in thinking that KK can explain these cases: in general, both fans and foes of KK alike hold that it can (help) explain the infelicity of assertions like (4).⁶

There are a number of (infamous) cases in which KK seems to fail, however—cases where subjects seem to know that P without believing that they know that P. Assuming that

⁵ Why? Suppose KK is true, that an assertion is permissible only if the speaker knows it, and that someone asserts “p & ~KKp”. Their assertion will be permissible only if they are such that $K(p \ \& \ \sim KKp)$. This entails (by KK) that the speaker is such that $KK(p \ \& \ \sim KKp)$. Distributing each of the outside K’s yields $KKp \ \& \ KK(\sim KKp)$. Appealing twice to the factivity of K reduces the second conjunct to $\sim KKp$, and we are left with $KKp \ \& \ \sim KKp$, a contradiction. So if KK is true, it is impossible for an assertion of anything like (4), which has the form “p & ~KKp”, to satisfy a knowledge norm on assertion.

⁶ See, for instance, Cohen and Comesaña (2013: 24-25), Greco (2015), Das and Salow (2018), and Holguín (2019: 13). Sosa (2009), though he rejects both KNA and KK, seems open to the possibility that KK could at least *help* explain why assertions like (4) sound infelicitous.

knowledge requires belief, if these cases are coherent, then they are counterexamples to KK. The first case I have in mind is something like this:

The salesman. George is busy selling vacuums door-to-door. As such a salesman, he knows a great deal about the vacuums: how many colors they come in, for example. While he often wonders whether he will sell enough vacuums to meet his quota, he never wonders if he *knows* how many colors the vacuums come in. He is simply too busy selling vacuums to reflect on what he knows.

George knows how many colors the vacuums come in. But he has never reflected on whether or not he knows how many colors the vacuums come in, and so it is possible that he simply has no beliefs about whether or not he knows how many colors the vacuums come in. Note: this is not to say that failing to reflect on P entails that one has no belief as to P; it is only to say that failing to reflect on P makes it *possible* that one lacks beliefs as to P. If (as many think) belief is necessary for knowledge, then George may not know that he knows how many colors the vacuums come in. George may thus be in a KK failure.

The obvious way to resist this case would be to argue that George does, in fact, have the required belief (perhaps implicitly). But it is not obvious that he *must* have such a belief. Moreover, if KK is true, George must not only believe that he knows, he must believe that he knows that he knows that he knows ... that he knows that P (for every finite number of iterations). Again, this very well might be possible, but claiming that it is *necessarily* the case whenever one knows that P is a non-trivial explanatory burden for advocates of KK.

Another difficult case for advocates of KK to explain is the following:

The therapist. A philosopher—in particular, one who believes that belief is necessary for knowledge—is seeing a therapist. During their first session, the therapist (somewhat incautiously) tells the philosopher that given his behavior, he does not believe his parents love him. The philosopher believes the therapist, even though the therapist seems to be making a mistake—when he thinks of his relationship with his parents, it seems to him that they really do love him. As it turns out, the therapist *is* making a mistake: the philosopher believes (and knows!) that his parents love him. At the same time, the

philosopher believes that he does not believe his parents love him. He reflects: “belief is necessary for knowledge, so I must not know my parents love me either.”

In this case, the philosopher believes he does not know that his parents love him, even though he knows his parents love him. Believing one does not know that P is *not* the same thing as *not believing* one knows that P, of course, but the two often go together. In particular, it certainly seems *possible* that the philosopher would *also* fail to believe that he knows that his parents love him. If belief is necessary for knowledge, then in this case, the philosopher is in a KK failure.

Finally, there are especially infamous “unconfident examinee” cases. These cases typically go something like this:

The unconfident examinee: Mary, a high schooler, has spent weeks studying for the AP US history exam. She has memorized countless facts, and has written practice essays making use of those facts. Unfortunately, Mary has severe testing anxiety. When Mary comes to the multiple-choice section on the exam, she starts to doubt her memory. Nevertheless, she puts her best guess down for each question. As it turns out, she gets every question correct.

The usual moral is that for some historical propositions P, Mary knows that P, but she does not believe that she knows it.

This case is especially infamous because it is tempting to argue either that Mary’s doubts defeat her first-order justification for believing that P and thus destroy her first-order knowledge that P, or that Mary does not, in fact, believe that P, or both. On the other hand, for some (especially for those of a more externalist bent), it is still quite intuitive that Mary nevertheless knows that P.⁷

⁷ There are also cases where subjects do not have the concept *knowledge* and so, the thought goes, cannot believe that they know anything at all. I mention this only briefly because, while it does pose a *prima facie* challenge for KK, discussion of it would take us too far afield.

Together, these cases put a significant amount of pressure on the KK theorist. I do not mean to claim that they decisively refute the KK thesis—only that they have the potential to make the KK thesis a costly position to maintain.

Now, even if some of these cases are counterexamples to KK, one might not think they are very interesting counterexamples. For one might take it that the real point of KK is that if one knows something, one has already done everything that one needs to do, epistemically if not psychologically, in order to know that one knows it. In line with this thought, many advocates of KK hold a weaker version of KK.⁸ According to this weaker version, if you know that P, then you are automatically *in a position to know* that you know that P. Thus, there are at least two versions of the KK thesis:

Strong KK: Necessarily, for all S and P, if S knows that P, then S knows that S knows that P.

Weak KK: Necessarily, for all S and P, if S knows that P, then S is in a position to know that S knows that P.

The notion of being in a position to know threatens to be somewhat obscure, but the gist is that if one is in a position to know that P then one has done everything one needs to do, epistemically if not psychologically, in order to know that P; for instance, one may not have formed the belief that P, but one has good evidence for P, one is not in a Gettier case, etc..⁹ Those who find KK plausible but are worried by the (purported) counterexamples outlined above often retreat to Weak KK, for cases like *the salesman* and *the therapist* are almost certainly not counterexamples to Weak KK. And while cases like *the unconfident examinee* may still be counterexamples to

⁸ For instance McHugh (2010), Das and Salow (2018), and Dorst (2019). Dorst (2019) works with a slightly different principle: if one is in a position to know that P, then one is in a position to know that one is in a position to know that P. Still, Dorst's version of KK is "weak" in the sense that simply failing to have the relevant higher-order belief furnishes no counterexample.

⁹ This is how Smithies (2012: 268) describes being in a position to know. See Williamson (2000: 95) for a related explication.

Weak KK, since it is contentious whether subjects in such situations actually preserve their first-order knowledge, moving to Weak KK still manages to avoid the most troubling cases for Strong KK without surrendering its basic insight.¹⁰ There is thus much to be gained by moving from Strong KK to Weak KK.

For those who sympathize with the KK thesis's fundamental insight, giving up Strong KK and retreating to Weak KK is a tidy way to avoid having to deal with the most troubling (alleged) counterexamples countenanced above. It is, however, more difficult to see how Weak KK could explain the infelicity of assertions like (4). Recall:

(4) It's raining, but I don't know if I know it's raining.

Suppose we wanted to offer a Weak KK-based explanation of (4) that parallels Greco (2014)'s (which appeals to Strong KK). Such an explanation would look something like this: when one asserts (4), if one is following the knowledge norm of assertion, one must know that it's raining. If Weak KK holds, then one is thereby in a position to know that one knows it's raining.

In Greco's original explanation, from the fact that one knows it's raining follows the fact that one knows *that* one knows it's raining. Thus, given Greco's knowledge norm, one will be permitted to assert that one knows that it's raining. He concludes that either one's reticence to assert this in the second conjunct is inappropriate, or that one's assertion of the first conjunct must have been impermissible. Thus, either way, the asserter was behaving inappropriately.

But if we can only appeal to Weak KK, from the fact that one knows it's raining, it only follows that one is in a position to know that one knows that it's raining. It does not seem very

¹⁰ On the other hand, the subject of *the unconfident examinee* might, after all, be in a position to know that they know the relevant historical facts. If this is true, *the unconfident examinee* is not a counterexample to Weak KK.

likely that merely being in a position to know makes assertion permissible.¹¹ So in refusing to assert that one knows that it's raining, one is no longer being “unnecessarily reticent” to assert something that one is permitted to assert. Thus, it is unlikely that there is a straightforward way to adapt Greco's original explanation of the fact that (4) sounds infelicitous into an explanation which appeals only to Weak KK.

Now, of course, there may be some other explanation of the fact that (4) sounds infelicitous which makes central use of Weak KK. This does not seem especially likely either, though. For if Weak KK but not Strong KK holds, one may very well know that P, be in a position to know that one knows that P, and yet not actually know that one knows that P—perhaps, for instance, one has only recently begun to reflect on whether or not one knows that P. If this were one's situation, then an assertion like (4) would seem to simply give voice to that state. In other words, the state that an assertion like (4) describes—the state that must obtain for an assertion like (4) to conform to a factive norm on assertion—could very well be one in which Weak KK holds.¹² This suggests that ultimately, we should not expect Weak KK to be able to explain what is wrong with assertions like (4). Of course, the situation very well may be more complicated, and it might turn out that ultimately, Weak KK *is* at the root of what's wrong with assertions like (4). In my view, however, these considerations make this rather unlikely.

¹¹ Though see Willard-Kyle (2020) for a recent dissenting view. And notice: for Greco's explanation to work, it is not enough for the assertion to be permissible. It must also be something like ‘necessary if given the opportunity’. To me, it seems even less likely that being in a position to know could play this role.

¹² It may be that while, given Weak KK, each conjunct could be separately known, still, assuming the *conjunction* is known violates Weak KK, and thus Weak KK can after all explain why asserting the conjunction is defective. One might then claim that this is sufficiently obvious a priori to explain why the assertion sounds infelicitous. But this path is not so smooth. For one thing, it will only explain what is wrong with *believing* “p & ~KKp” by positing a knowledge norm for belief, which is contentious. Second, it is not clear that it could also explain what is wrong with asserting “p”, and then (slightly later) asserting “~KKp”, without positing a norm according to which one should know the conjunctions of one's assertions. Even a weak version of such a norm could prove quite troublesome.

Those sympathetic to KK are thus caught in something of a dilemma: on the one hand, they may opt for Strong KK. There is some promise that the data could help motivate this position, but at the same time, anyone who defends Strong KK will have to address a broad range of apparently plausible counterexamples. On the other hand, one could opt for Weak KK in the hopes of preserving the core of the view while avoiding some of its stronger commitments. But since it is less likely that Weak KK can explain the data and these data are often taken to be a strong consideration in favor of KK, moving to Weak KK could be more expensive than it initially seemed.

Section 2. Introducing a Normative Requirement

I will now start to put forward my positive view. The first step is to introduce and motivate two norms. This will occupy the entirety of the present section. In the next section, I will use these norms to explain the data often taken to support KK. Since it is more plausible that these norms are correct than that Strong KK is true, this explanation lessens Strong KK's explanatory power and so weakens the overall case for KK.¹³

The two norms I will introduced are both “BK”—believe that you know—norms. One is on assertion, and the other is on belief.

BKNA: One should assert that P only if one believes that one knows that P.

BKNB: One should believe that P only if one believes that one knows that P.

Plausibly, both “assertion” and “belief” are polysemous. It will thus be helpful to narrow our attention to particular uses of each. I have in mind the use of “assertion” on which it picks out acts where subjects mean to, among other things, commit themselves to the truth of the asserted

¹³ Moreover, even if Weak KK can furnish a satisfying explanation, it has baggage of its own (see Liu (2020) and Williamson (2000)).

content. For example: when I want to let my friend know the bus comes by every half hour, I assert “the bus comes by every half hour.” An example of something which is *not* an assertion (in the relevant sense) may be my friend’s utterance at the game, later that night, of “they’re going down!”. By “belief,” I mean to pick out the propositional attitude paradigmatically expressed by assertions.¹⁴ Assertions and beliefs are thus closely linked; consequentially, there is a sense in which BKNA and BKNB are two sides of the same coin.

Both BKNA and BKNB are meant to express conditional obligations. If you like, feel free to understand them as “wide scope” principles. The point is that they forbid *both* asserting (or believing) that P *and* failing to believe that one knows that P. I am not the first to suggest such norms.¹⁵ Importantly, I do not mean these norms to be “the” constitutive norms for either belief or assertion. Since they only forbid certain combinations of beliefs (and assertions), they are each compatible with plenty of other norms.¹⁶

To me, at least, these norms are fairly intuitive: it makes sense that one should assert that P only if one believes that one knows it, and since assertion is the paradigmatic expression of belief, BKNB makes just as much sense. If you do not find these norms immediately plausible, though, there are two lines of thought which can each give both norms some principled support.

¹⁴ Following Hawthorne, Rothschild, and Spectre (2016), some hold that belief is “weak”: that is, that one may have sufficient evidence to justifiably believe that P without having sufficient evidence to assert that P, and that belief that P may even be compatible with credence less than .5 in P. I am not inclined to think that the state Hawthorne, Rothschild, and Spectre are identifying is really belief. Even if Hawthorne, Rothschild, and Spectre are right, though, there is surely a stronger propositional attitude available—that which is paradigmatically expressed by assertions. It is in *that* attitude that I am here interested.

¹⁵ Huemer (2007) and (2011) and Smithies (2012) propose similar norms.

¹⁶ In particular, BKNA is meant to be compatible with knowledge norms on assertion which merely forbid both asserting P and failing to know that P. It will not be compatible with knowledge norms on assertion which say that knowing that P permits one to assert that P, for—as *the salesman* and *the therapist* from §2 (allegedly) show—one may know that P without believing one knows it.

The first line of thought draws out some connections between belief and assertion, and then between assertion and reasons for action. It then appeals to the idea that one should use P as a reason for action only if one knows that P. Overall, its goal is to highlight some important connections between belief (and assertion) and knowledge.

To begin, belief and assertion are intimately related. Assertions express beliefs, and there is a sense in which one is ready to assert that P iff one believes that P. The best way to get a grip on the relevant sense of “ready” is to look at some examples. If one is about to engage in some Frankfurtian bullshitting or is about to lie, there is certainly some sense in which one is ready to assert P, but it is not the relevant sense. Conversely, even if there are various practical considerations which disfavor asserting P (of relevance or etiquette, for instance), one may still believe P and so be ready, in the relevant sense, to assert P. One may also be ready to assert P and yet not be actively or consciously thinking that P. There is thus an important and intuitive sense in which assertion and belief are intimately related: assertions express beliefs, and there is a sense in which one believes that P just in case one is ready to assert that P.

Assertion is importantly connected to the audience’s reasons for action as well. In paradigmatic cases of asserting that P, one invites one’s audience to use P as a premise in at least some instances of practical reasoning. For instance: if I assert that it’s raining, I am inviting you to reason from “it’s raining” to conclusions such as “I should grab my umbrella”. Of course, asserting that P does not mean that the speaker is inviting the audience to go and use P as a premise in every possible kind of practical environment. For instance, if P is the proposition that the left sandwich is made with almond butter and the right sandwich is made with peanut butter, I may assert P and yet not mean to invite you to repeat it to your friend who is allergic to peanut butter. Still, I mean to invite you to act on the basis of P in at least some more ordinary

circumstances—for instance, if I know that you prefer peanut butter to almond butter. So in asserting that P, one invites one’s audience to use the claim that P as a premise in practical reasoning in at least some environments, even if not in all.

Furthermore, it is plausible that one should invite (or be ready to invite) one’s audience to practically reason from P (in at least some circumstances) only if one believes that they would be justified in so reasoning.¹⁷ For instance, suppose Scott Pilgrim is ready to assert that it’s snowing to Ramona Flowers. If he did not believe that Ramona would be justified in grabbing an extra coat on the basis of his assertion, intuitively, something would be wrong—intuitively, there is something defective about this combination of states. Again, Scott need not believe that Ramona would be justified in (say) betting \$10,000 on the claim that it’s snowing. But for a wide variety of ordinary actions, Scott should be ready assert that it’s snowing only if he believes that Ramona would be justified in performing (or refraining from) that action on the basis of his assertion.

Now, a common thought is that one should use P as a premise in one’s practical reasoning only if one knows that P. This is, for instance, a popular diagnosis of what is problematic in the following practical reasoning:

My lottery ticket is going to lose.

Therefore, I should sell it for a penny.

This is a paradigmatically bad piece of practical reasoning (Hawthorne 2004: 29). A plausible explanation is that anyone who engages in it (without inside information) does not know the first premise, and so cannot rely on it in drawing practical conclusions. The verdict commonly drawn is that one may practically reason from P only if one knows that P.

¹⁷ Or at least, in cases where one’s audience does not take one to be lying, does not take one to be a complete ignoramus on the subject at hand, etc.. For our present purposes, we can set these cases aside.

Insofar as this is plausible and straightforward it is plausible that, in some external or ideal sense of “should”, one should believe that in general, subjects permissibly reason from P only if they know it. If one should, in fact, believe this, and if (as just argued) one should assert P (or be ready to assert it, that is, believe P) only if one believes that one’s audience could justifiably reason from it, it follows that: one should assert P (or be ready to assert it, that is, believe P) only if one believes that one’s audience knows that P.

Now, one should believe that S knows that P only if one believes that one knows P oneself. So at the end of the day, stringing these (in some sense external or ideal) “oughts” together, it follows that: one should assert that P (or be ready to assert it, that is, believe P) only if one believes that one knows that P.

Now, this is hardly an airtight argument for BKNA or BKNB. For one thing, it may not be that reasoning from P requires knowing P. For another, a number of “oughts” are getting strung together above, and one may doubt that—given the complexity of the reasoning—the concluding “ought” really binds us non-ideal agents. Even so, the argument gains plausibility if we keep in mind that the “oughts” should be understood in an external sense—that is, in a sense similar to that in which one *ought* to do the right thing (even if one does not believe that it is the right thing), or to that in which I *should* have bought extra food for the party tonight (even though I had no way of knowing fifteen extra guests would arrive unannounced). Totally discounting the validity of such imperatives seems unwise. Furthermore, this argument helps show how BKNA and BKNB connect with some intuitive expectations we have for each other: if someone asserts something, we generally take them to be offering it to their audience to use as a premise in (at least some instances of) practical reasoning. Perhaps because of this, we are inclined to be upset on learning that an asserter did not take themselves to know what they

asserted. So BKNA and BKNB, at the very least, seem to cohere well with expectations we have for each other as asserters.

Are there not circumstances where violations of BKNA and BKNB seem acceptable? Certainly. But the mere fact that we relax our expectations in certain environments does not mean that the norms codifying those expectations are not in effect. Excusable violations of a norm should not be confused with counterexamples to that norm, and since BKNB and BKNA do not involve an “all things considered” obligation, there will certainly be cases where other, more important norms take over—for instance, in some cases, one really should lie to protect one’s friends.¹⁸

The second line of argument applies primarily to BKNA. In virtue of the connections between assertion and belief outlined above, it supports BKNB as well. The central idea is that BKNA can coopt the linguistic evidence for Timothy Williamson’s knowledge norm of assertion (Williamson (2000)). Insofar as one doubts the knowledge norm and still wants to explain Williamson’s data, then, his data provide excellent evidence for BKNA (and so, given the intimate relation between belief and assertion, BKNB as well).

Williamson (2000) considers two kinds of linguistic data.¹⁹ The first is that, when one asserts that P, it is often appropriate for one’s interlocutors to respond with “how do you know that?”. As Williamson notes, “how is it that P?” presupposes that P, and so the felicity of this response suggests that we expect speakers to know what they assert. But this question can also be seen as presupposing that someone who asserts that P takes themselves to know that P, so at the very least, this linguistic pattern is consistent with BKNA. The other datum is the fact that

¹⁸ One might also worry that BKNB requires us to have infinitely many beliefs. I discuss this worry in §5.

¹⁹ In chapter 11. There are some other data, but the point can be made with just these.

epistemic Moorean sentences like “it’s raining, but I don’t know that it’s raining” ((3), above) sound infelicitous. This is also consonant with BKNA—for we expect someone who asserts that it’s raining to believe that they know that it’s raining, and yet in the epistemic Moorean case, they also seem to believe that they do not know it.²⁰

On the other hand, even if a norm like Williamson’s is correct, the fact that BKNA coheres well with his linguistic data is still a mark in its favor. And to the extent that a norm like Williamson’s merely forbids asserting P while failing to know that P, such a norm is compatible with BKNA.²¹ Moreover, such a norm could even help furnish a principled basis for BKNA: that one should assert P only if one believes that one knows it, may be partly justified by the fact that one should assert P only if one knows it. Thus, those who favor such a knowledge norm on assertion have a straightforward reason to accept BKNA.

One last consideration in favor of BKNB and BKNA is that, as I will shortly argue, they can explain why believing (or asserting) propositions like (4) is defective (and sounds infelicitous). This is hardly an independent reason to favor them. Nevertheless, it is worth keeping in mind.

Section 3. The Positive Proposal

This section argues this paper’s primary contention: that BKNA can explain both why certain assertions (a) are defective and (b) sound infelicitous without appealing to *any* version of

²⁰ Williamson explicitly considers something like BKNA and acknowledges that it can explain much of the linguistic data as well. He worries, though, that it would make a poor constitutive norm for assertion (260). But this is no problem for my account, since I am not claiming that BKNA is a constitutive norm on assertion, nor am I claiming that it is the only normative constraint involving assertion.

²¹ This may not be the right interpretation of Williamson’s norm, since Williamson seems to take it that knowledge makes assertion “correct” (241) and thus “warranted” (243). Still, it is an open question how Williamson’s sense of “correct” relates to the kind of *permission* which is the dual of the “ought” employed in BKNA and BKNB. In any case, this is why the text says “a norm *like* Williamson’s”.

KK. It will argue this piecemeal, first by using BKNA to explain what is defective about assertions like (4), and then by integrating this into a fuller explanation of why such assertions sound infelicitous. This overall explanation of the data thus keeps the twin explanatory questions outlined in §2—why is the assertion defective, and why does it sound infelicitous?—distinct.²²

I will first explain why assertions of (4) are defective. In short, they are defective because those who assert them will either violate BKNA, have a defective belief, or violate one of two very plausible rational principles which I am about to introduce. The rational principles I have in mind are these:

Principle 1: S ought to be such that they believe they know P & Q only if they believe they know P and they know Q. Schematically:

$$O(\sim B(K(p \ \& \ q)) \vee B(Kp \ \& \ Kq))$$

Principle 2: S ought to be such that if they know that P entails Q, then they believe P only if they also believe Q. Schematically (and employing the material conditional):

$$O(K(p \ \text{entails} \ q) \supset (\sim Bp \vee Bq))$$

Just like BKNA and BKNB, these principles are merely meant to merely prohibit certain combinations of states. The first principle forbids both believing one knows P & Q while failing to believe both that one knows P and that one knows Q; the second forbids knowing that P entails Q, believing P, and yet failing to believe Q. I take these to be fairly minimal aspects of a theory of rationality and so I will not argue for them here.

These principles put us in a position to see what is wrong with an assertion of (4). To see this, I will start by assuming someone has asserted (4) and then, through a sequence of stages, see what must have been the case for them to have satisfied BKNA, principle 1, and principle 2.

²² In many ways, this project is on friendly terms with the suggestions offered by Benton (2013). It differs from Benton's, though, on whether there are some situations where assertions (and beliefs) of the form "p & ~KKp" would be rational (and sound felicitous). Benton thinks there are; I do not.

It will turn out that this is only possible if they have a certain belief which I will argue is (perhaps rationally) defective. Before going through those stages, though, it is crucial to notice the following: since both principles 1 and 2, as well as BKNA, could be satisfied in different ways, at each stage, different “options” will be available. For instance, at Stage 3, in order to satisfy principle 2, the asserter must either be such that $B(Kp \wedge K \sim KKp)$, or must be such that $\sim BK(p \wedge \sim KKp)$. At this stage, there are thus three options: $B(Kp \wedge K \sim KKp)$, $\sim BK(p \wedge \sim KKp)$, or violate principle 2. Since our goal is to argue that something *must* be wrong, we cannot simply assume that the third option obtains and the speaker violated principle 2. Likewise (as the sequence of stages will make clear), if the second option were to obtain, then the speaker would violate either principle 1 or BKNA (given how previous stages played out). Thus, we will have to assume that the first option obtains, and the speaker is such that $B(Kp \wedge K \sim KKp)$.

Analogous pieces of reasoning will be relevant for each of the four stages. As it turns out, at each stage, the constellation of norms permits only a single way forward (this will all be clearer when we look at the stages themselves).

Of course, actual subjects might “take different paths”, so to speak, in which cases they will violate at least one of BKNA, principle 1, or principle 2. What the four-stage argument below shows is that having a certain defective belief is the only way for someone who asserts (4) to satisfy BKNA, principle 1, and principle 2. Thus, something is *guaranteed* to go wrong with assertions of (4)—either one of BKNA, principle 1, and principle 2 are violated, or the speaker has a defective belief. That is why assertions of (4) are defective.

Stage 1. Suppose a subject S sincerely asserts something like (4). The form of what is asserted is:

$$p \wedge \sim KKp$$

Stage 2. S's assertion will satisfy BKNA only if S believes they know what they have asserted. Schematically:

$$BK(p \wedge \sim KKp)$$

Stage 3. Now, S can satisfy principle 1 in one of two ways. The first is to lack the belief from step 2, that is, to be such that $\sim BK(p \wedge \sim KKp)$. In this case, their assertion would violate BKNA. The other way is to believe they know each conjunct, that is:

$$B(Kp \wedge K\sim KKp)$$

Since we are trying to show that something must be wrong, we will assume they satisfy principle 1 in the second way.

Stage 4. Since knowledge is factive, I assume S knows that $Kp \wedge K\sim KKp$ entails $Kp \wedge \sim KKp$. On the plausible assumption that S ought not give up this piece of knowledge, there are two ways left to satisfy principle 2. The first way is to fail to have the belief from step 3, that is, to be such that $\sim B(Kp \wedge K\sim KKp)$. That would force S to violate either principle 1 or BKNA, though. Thus, we will assume that S satisfies principle 2 by believing $Kp \wedge \sim KKp$, that is, by being such that:

$$B(Kp \wedge \sim KKp)$$

Thus, the only way for an assertion of (4) to satisfy BKNA, principle 1, and principle 2 is if the speaker believes they know P and that they do not know whether they know P—that is, that they themselves are presently in a Strong KK failure.

Intuitively, this belief is defective. I will argue that in a moment. But before I do, the present point is not to be missed: the above argument shows that the defectiveness of asserting something like (4) *reduces*, at least in part, to the defectiveness of believing that one oneself is

presently in a Strong KK failure. This is progress: if we can show what is wrong with that belief, we can show what is wrong with asserting something like (4).²³

If one believes that one oneself is presently in a Strong KK failure, one is in a (perhaps rationally) defective doxastic state. There are a few reasons why this is so. The first is that one can know, *a priori* and without too much difficulty, that one’s belief cannot amount to knowledge. For supposing one’s belief is knowledge, one quickly arrives at a contradiction:

- (i) suppose one knows one is in a Strong KK failure: $K(Kp \ \& \ \sim KKp)$
- (ii) distribution of K: $KKp \ \& \ K\sim KKp$
- (iii) factivity of K: $KKp \ \& \ \sim KKp$ (contradiction)

Thus, one’s belief cannot be knowledge; moreover, one can know that this is the case *a priori* and fairly straightforwardly. Importantly, this is so regardless of whether or not Strong KK failures are, in fact, possible. Insofar as belief aspires to knowledge—and plausibly there is *some* sense in which it does, if only that codified by BKNB—this belief is defective.²⁴

Furthermore, this belief directly violates BKNB. For if we represent “I know that P” by Q, the believer both believes that Q but disbelieves that they know Q. Moreover, it tends to violate BKNB in a second way as well. On even brief reflection, given the simplicity of the above proof, it will be nearly impossible to believe that one knows “I know P, but I do not know whether I know P”, and so it will be nearly impossible to satisfy BKNB as long as one believes “I know P, but I don’t know whether I know P.”

²³ There is a parallel argument which employs BKNB (in stage 2) to show the exact same thing is what is wrong with the belief that “ $p \ \& \ \sim KKp$ ”. This is important: to only explain what is wrong with *asserting* something like (4) is to give, at best, an incomplete explanation (compare: explanations of Moorean assertions which appeal only to Gricean implicatures).

²⁴ To say belief aspires to knowledge is not necessarily to say that knowledge is a (fundamental) norm on belief. For example, the fundamental norm of belief may be truth, but some kind of aspiration to knowledge may fall out of this fundamental norm (for a similar line of thought, see Sylvan (2020)).

On the other hand, if one satisfies BKNB by believing “I know that: I know P, but I don’t know whether I know P”, one believes something which entails, a priori and almost immediately, an explicit contradiction. As a merely conditional obligation, BKNB says one must either believe this, or give up the belief that one is in a Strong KK failure. Insofar as BKNB is true and we should not believe things which a priori entail explicit contradictions, then, one should not believe that one is in a Strong KK failure.²⁵

Finally, consider the assertoric expression of this belief: “I know P, but I do not know whether I know P”. This assertion is clearly defective. But intuitively, the problem is not just with *asserting* it—the problem is with believing it.

So the belief that “I know that P, but I don’t know whether I know that P” is a defective belief. Is it a *rationally* defective belief? It may be—but this depends on whether you take BKNB to express a rational constraint on belief, whether beliefs ought, as a matter of rationality, to aim at knowledge, and whether the fundamental problem with asserting “I know that P, but I don’t know whether I know that P” is a matter of irrationally believing it. I think these are plausible views (especially the last), but I will not argue for them here. After all, there may be another, more fundamental sense in which the belief is defective—say, that it manifests a failure to respect knowledge. All I need to show is that assertions of (4) can satisfy BKNA (and principles 1 and 2) *only if* the speaker believes they themselves are presently in a Strong KK failure, and that such a belief is defective. It follows that such assertions are defective.

The task now is to explain why assertions like (4) sound infelicitous.

²⁵ And insofar as having this impermissible belief that $Kp \ \& \ \sim KKp$ is the only way for an assertion to satisfy BKNA/BKNB, principle 1, and principle 2, it follows that one should not have asserted/believed (4) in the first place.

It is important to notice at the outset that there is a difference between hearing (4) as infelicitous, and hearing (4) and simply being confused. If one is confused, one struggles to grasp the speaker's point; if one hears an assertion as infelicitous, one grasps the speaker's point, but has the feeling that that point (or the speaker's performance) is somehow defective. Dubious assertions like (4) are likely to be confusing because they contain an iteration of "knows that". If one is confused by this, one may be tempted to interpret the second conjunct as retracting or retroactively hedging the first. I would not be surprised if, in the real world, this was the most common reaction. It is not, however, the reaction that some claim KK can explain, and so it is not the reaction presently at issue. What we are concerned to explain is the infelicity. Thus we should set aside cases where the audience is merely confused and instead turn our focus to cases where the audience has a particular feeling which we have been calling "infelicity"—plausibly, the feeling that something about the speaker's performance must have been, somehow, defective.²⁶

When one hears (4) as infelicitous, one gets the distinctive feeling that the speaker has done something—one does not know exactly *what*—wrong. The phenomenology is similar to hearing a quickly presented, somewhat dubious argument. In each case, one is certain some mistake must have been made, but has trouble pointing out exactly where the mistake is.²⁷

When one has this distinctive feeling after hearing a dubious argument, it seems plausible that at least sometimes, it is because one has some faint intellectual grasp of what is actually wrong with the argument. The present suggestion is that something similar is happening with the

²⁶ Of course, there may be some sense in which confusing assertions are infelicitous—perhaps they violate some Gricean maxim or other. But in this case, to say that (4) sounds infelicitous is to draw attention to a more specific piece of phenomenology.

²⁷ I felt this way when I first heard St. Anselm's ontological argument. I still feel this way.

infelicity which accompanies dubious assertions like (4): when we hear someone assert “it’s raining, but I don’t know if I know it’s raining” and we feel as if the speaker must have done *something* wrong, it is because we have some intuitive grasp of what they have, in fact, done wrong.

Now, admittedly, this thought is somewhat speculative. But it is not meant to be radically revisionary. For instance, a similar idea seems implicit in Greco (2014)’s explanation of why assertions like (4) sound infelicitous. Given that this is a plausible explanation of the phenomenology of hearing a dubious argument and that the infelicity accompanying assertions of (4) shares that phenomenology, this principle should be fairly plausible.

The task now is to spell out exactly what we have a grip *on* when we hear assertions like (4) as infelicitous. My suggestion is that we have a grip on the explanation I gave above. That is: when we hear (4) as infelicitous, we intuitively (though partially and weakly) understand that the speaker could satisfy BKNA *only* by having the defective belief that they both know P yet do not know if they know that P (that is, the defective belief that $Kp \ \& \ \sim KKp$), and that our (partially and weakly) grasping this causes us to feel that *something* must be wrong with the assertion. There are thus two aspects of my explanation of why assertions like (4) sound infelicitous: first, the relevant feeling of infelicity comes from intuiting (perhaps imperfectly) what is defective about them; second, that I have successfully explained (above) what is defective about them, and it is this explanation that we intuitively (though weakly) grasp.

Why should we think my above explanation of why assertions of (4) are defective is the very explanation we intuitively (though weakly) grasp when we hear such assertions as infelicitous? First, we should note that the explanation we weakly intuit, whatever it is, must have the right amount of complexity to explain the phenomenology: it must neither be too

complex, nor too straightforward. For it cannot be too complex, or else we would be utterly unable to grasp it, and so—assuming that weakly intuiting what is wrong with the assertion explains why it sounds infelicitous—there would be no accompanying sense of infelicity; at the same time, it cannot be too straightforward, or else the statement would sound just as infelicitous as (for instance) omissive Moorean assertions like “it’s raining but I don’t believe it”, which it does not.

My explanation of why dubious assertions like (4) are defective seems to strike the right balance. For first, while it is complicated, it is not so complicated as to be utterly incomprehensible. Consider a more colloquial response to an assertion of (4):

You just said that it’s raining, but that you don’t know if you know that. Well, I can’t understand why you would assert that it’s raining unless you thought you knew it was raining. But you also thought you didn’t know if you knew it was raining... alright... but then, how can you think you know one thing, yet at the same time think that you don’t know whether or not you know it?!

This is a (very) streamlined version of the above explanation. It nevertheless captures that explanation’s core moves. To me, it certainly seems like a thought someone who heard (4) could intuitively grasp, at least to some degree. It is thus plausible to say that those who hear (4) as infelicitous have some grasp of the above explanation.²⁸

Moreover, we commonly hold each other to BKNA and principles 1 and 2. Consider first our disposition to feel upset on learning that an asserter did not believe they knew what they said. This seems like good evidence that in ordinary life, we hold each other to BKNA. Principle 1 certainly seems common enough as well—we expect people to believe they know P & Q only if they believe they know P and they know Q. Finally, principle 2 is something we rely on all the

²⁸ Remember that we have confined our attention to subjects who hear (4) as infelicitous, not those who hear (4) and are simply confused by the iteration of “knows”. The claim is not that the explanation is clear enough to be grasped by *everyone*—only that it is clear enough to be grasped by everyone who genuinely hears (4) as infelicitous.

time in criticizing each other: I expect my friend, who knows that the busses do not run during US holidays and believes that tomorrow is a US holiday, to also believe that the bus will not be running tomorrow (and to plan accordingly!). This lends credibility to the claim that we have some intuitive grasp of the above explanation of what is wrong with an assertion of something like (4): if we were not comfortable using these norms, there would be no reason to think that we had any grasp whatsoever of the above explanation. Since these norms come naturally, though, some (perhaps streamlined) version of the explanation above should be accessible to everyone.

At the same time, the above explanation is still somewhat complicated. The crucial point is that given the phenomenology, this fact works in my favor. For we should expect the true explanation of what is wrong with assertions of (4) to be somewhat complicated. If it were much simpler—assuming that intuitively grasping it explains our feeling that something must be wrong—a different phenomenology would accompany assertions of (4). In particular, if the true explanation of why assertions of (4) are defective were much simpler than the explanation I gave above, assertions of (4) would both sound much more infelicitous than they do, and we would have a more immediate sense of what exactly is wrong with them.

Thus, the complexity of the above explanation is commensurate with the phenomenology of the infelicity accompanying assertions of (4). It is not so complex as to rule out being grasped to any extent, which coheres with the fact that assertions of (4) sound infelicitous to some degree or other. At the same time, it is not so simple as to be perfectly and immediately grasped, and this coheres with the facts that assertions of (4) sound less infelicitous than (say) omissive Moorean assertions like “it’s raining but I don’t believe it”, and that we have a weaker grasp of what exactly is defective about them.

Again, this proposal relies on the idea that the infelicity accompanying assertions like (2) and (4) is generated by the audience's grasping (though faintly) why such assertions are defective and that, moreover, the degree of infelicity and its particular phenomenology are linked to the complexity of that explanation. As I said earlier, this principle is somewhat speculative. At the same time, it seems to be in line with how many understand explanations of linguistic infelicity, and—given the phenomenology—seems intuitively plausible. Most importantly, the overall explanation in which it is embedded seems, on the whole, much more plausible than any explanations which depend on Strong KK. This explanation thus undercuts the explanatory power of Strong KK. And insofar as Strong KK's ability to explain dubious assertions like (4) is a strong point in its favor, the present explanation of both what is wrong with dubious assertions like (4) and why they sound infelicitous significantly weakens the case for KK.²⁹

Section 4. Objections and Replies

In this section, I will consider two objections. The first is that BKNA is too strong a normative requirement—simply put, there seem to be some circumstances C and propositions P where assertions of P in C sound perfectly acceptable, and yet it does not seem reasonable to expect the asserter to believe that they know that P. For instance, (5):

(5) Northwestern's going to lose!

An overzealous fan could assert this at the beginning of the game, and yet we would not expect them to believe that they know it. This suggests that BKNA is too strong.

²⁹ If I am right and KK does not have the explanatory power many take it to have, moving from Strong KK to Weak KK may not be as expensive as it originally seemed. In a certain frame of mind, then, this paper could be seen as advancing a reason to prefer Weak KK to Strong KK.

In reply, the assertions with which BKNA is concerned are only to include assertions which are “serious” in some relevant sense, and (5) is not one of those. To put this point another way: “assertion” is polysemous, and (5) is not an assertion in the sense that I am using the term. One way to see this is by paying attention to the fact that someone who asserts (5) in a certain tone of voice and context is not inviting their audience to do anything on the basis of its content.³⁰ For instance, if a Northwestern fan heard someone shout (5) and decided to immediately head home, we would hope that the speaker would stop them and clarify that they did not really mean to claim that Northwestern was going to lose, and tell them that, in fact, Northwestern might still win, so it *is* worth staying to watch the game.

A related worry is this: if such assertions are not governed by BKNA, why do sentences like (6) still sound infelicitous?

(6) Northwestern’s going to lose, but I don’t know if I know they are.
My guess is that if we heard someone assert something like (6), we would expect the speaker to think that the question of whether or not they knew Northwestern was going to lose was relevant. But plausibly, this is not the case with the original assertion. This change of context may bring with it the familiar expectation that the speaker believes that they know what they have asserted, and so the explanation offered in §4 straightforwardly applies.

Another worry is that BKNB requires one to have infinitely many beliefs, and since this is not a realistic expectation, BKNB is not a legitimate norm on belief. In reply, the first thing to notice is that BKNA does not have this problem, and even without BKNB, BKNA still gives a decent explanation of why dubious assertions like (4) are defective and sound infelicitous. More importantly, though, even though BKNB requires us to have infinitely many beliefs, that need

³⁰ Though I do not mean to claim that inviting one’s audience to act as if P in asserting P *suffices* to make one’s assertion *serious* in the relevant sense.

not be a mark against it—that may just place it in the realm of ideal obligations, just like the requirement to believe every tautology. Such requirements are still requirements. Just as excuses come easily with complicated tautologies, so we may have an excuse for violating BKNB past three or four iterations—and conversely, in simple cases with only one or two iterations, we take the BK norms to actually bind us, just like we take the requirements to believe simple tautologies to actually bind us. Importantly, though, when I appeal to BKNA and BKNB above, I never go past one or two applications of each norm. So this worry should not affect the argument from §4.

Section 5. Conclusion

In this paper, I started by outlining the explanatory work to be done, namely: why are certain assertions defective, and why do they sound infelicitous? I then sketched a commonly accepted answer which appeals to the KK thesis. Many take this explanation (or one like it) to be successful, and thus attribute a significant degree of explanatory power to KK. While such an explanation may be successful, though, the relevant version of KK—what I called “Strong KK”—is seriously challenged by the apparent possibility of subjects who believe that P, but simply do not believe that they know that P. These cases have moved many to adopt a weaker version of the KK thesis which claims that one knows that P only if one is also in a position to know that one knows that P. The problem with this weaker thesis, I argued, is that it is less clear that it can explain the data after all.

I then introduced two BK norms, norms that one should assert (or believe) that P only if one also believes that one knows that P. I argued that these norms can explain the data originally taken to motivate Strong KK—that is, they can explain both why dubious assertions like (4) are defective, and why they sound infelicitous.

Thus while Strong KK, considered in isolation, may give a somewhat satisfying explanation of the data, it is open to some disconcerting counterexamples; moreover, while Weak KK avoids the counterexamples, it is less suited to explain the data. Since there is an alternative explanation of the data which not only avoids appealing to any version of KK but makes use of less contentious principles, overall, KK does not have the explanatory power it is often taken to have.³¹

³¹ One surprising upshot is that by making the move from Strong KK to Weak KK much less expensive, it strengthens Weak KK's position relative to Strong KK.

REFERENCES

- Cohen, Stewart and Comesaña, Juan. “Williamson on Gettier Cases and Epistemic Logic”. *Inquiry* 2013.
- Das, Nilanjan and Salow, Bernhard. “Transparency and the KK Principle”. *Noûs* 2018.
- DeRose, Keith. “Epistemic Possibilities”. *The Philosophical Review* 1991.
- Dorst, Kevin. “Abominable KK Failures”. *Mind* 2019.
- Greco, Daniel. “Iteration and Fragmentation”. *Philosophy and Phenomenological Research* 2014.
- Greco, Daniel. “Iteration Principles in Epistemology I: Arguments For”. *Philosophy Compass* 2015.
- Hawthorne, John, Rothschild, Daniel, and Spectre, Levi. “Belief is weak”. *Philosophical Studies* 2016.
- Holguín, Ben. “Indicative Conditionals Without Iterative Epistemology”. *Noûs* 2019.
- Huemer, Michael. “Moore’s Paradox and the Norm of Belief”, in *Themes from G.E. Moore: New Essays in Epistemology and Ethics*, edited by Susana Nuccetelli and Gary Seay, Oxford University Press 2007.
- Huemer, Michael. “The Puzzle of Metacoherence”. *Philosophy and Phenomenological Research* 2011.
- Liu, Sebastian. “(Un)knowability and knowledge iteration”. *Analysis* 2020.
- McHugh, Conor. “Self-Knowledge and the KK Principle”. *Synthese* 2010.
- Smithies, Declan. “The Normative Role of Knowledge”. *Noûs* 2012.
- Sosa, David. “Dubious Assertions”. *Philosophical Studies* 2009.
- Sylvan, Kurt. “An Epistemic Non-Consequentialism”. *The Philosophical Review* 2020.
- Williamson, Timothy. *Knowledge and Its Limits*. Oxford University Press 2000.
- Willard-Kyle, Christopher. “Being in a Position to Know is the Norm of Assertion”. *Pacific Philosophical Quarterly* 2020.