Abominable KK Failures

Kevin Dorst
Massachussetts Institute of Technology
kmdorst@mit.edu

Abstract

KK is the thesis that if you can know p, you can know that you can know p. Though it's unpopular, a flurry of considerations have recently emerged in its favor. Here we add fuel to the fire: standard resources allow us to show that any failure of KK will lead to the knowability and assertability of abominable indicative conditionals of the form, 'If I don't know it, p.' Such conditionals are manifestly *not* assertable—a fact that KK defenders can easily explain. I survey a variety of KK-denying responses and find them wanting. Those who object to the *knowability* of such conditionals must either (i) deny the possibility of harmony between knowledge and belief, or (ii) deny well-supported connections between conditional and unconditional attitudes. Meanwhile, those who grant knowability owe us an explanation of such conditionals' *unassertability*—yet no successful explanations are on offer. Upshot: we have new evidence for KK.

KK is the thesis that if you're in a position to know p, you're in a position to know that you're in a position to know p. It's unpopular. In the years since Williamson (2000) showed it to conflict with plausible margin-for-error principles on knowledge, KK has mostly kept to the shadows. But no longer: a flurry of considerations in its favor have recently emerged—including theoretical pictures that explain why the margin-for-error principles may be intuitively plausible but false.[1] The spark of resistance has been kindled.

Here we add fuel to the fire. Standard resources allow us to show that any failure of KK will lead to the knowability and assertability of abominable indicative conditionals of the form, 'If I don't know it, p' (§1). Such conditionals are manifestly *not* assertable—nor, arguably, are they knowable. KK defenders have an easy explanation (§2). KK deniers owe us one.

[1] Cf. McHugh (2008); Cresto (2012); Cohen and Comesaña (2013); Greco (2014, 2015, 2017); Stalnaker (2015); Das and Salow (2016); Salow (2017).

I survey the possible responses and find them wanting. In order to avoid knowability, KK deniers must either (i) deny the possibility of harmony between knowledge and belief (§3.1), or (ii) deny well-supported connections between conditional and unconditional attitudes (§3.2). If they grant knowability, KK-deniers owe us an explanation of the *unassertability* of our conditionals. Yet none on offer is successful: it will not do to appeal to irrelevant antecedents (§4.1), or to Gricean norms (§4.2), or to a modification of the knowledge norm of assertion (§4.3), or to derivative norms of assertion (§4.4), or to other self-effacing speech acts (§4.5; cf. Williamson 2013 and Cohen and Comesaña 2013).

Upshot: KK deniers have some explaining to do.

1 Abomination

Give me your favorite KK failure, and I'll give you an unhappy consequence. Example: Kim is an unconfident examinee (Radford, 1966). Although she can know that Padua's in Italy (p), she can't know that she can know it:

(1) $Kp \wedge \neg KKp$

I'll make two assumptions—here I'll just state them; in §3 I'll defend them at length.

First assumption: if KK can fail, it can fail for a *diligent* agent—one who is sure of all and only the things she's in a position to know. Using sentential quantifiers, and 'C' to represent our agent's (probabilistic) credence function, my first assumption is:

> DILIGENCE: If it's possible that $(\exists p)(Kp \wedge \neg KKp)$, then it's possible that $(\exists p)(Kp \wedge \neg KKp)$ while $(\forall q)(Kq \leftrightarrow [C(q) = 1])$.

The idea here is simple. Being sure (or certain) of q captures what knowledge of q feels like 'from the inside'—if your certainty is justified and the world cooperates, it amounts to knowledge.[2] DILIGENCE captures the idea that if KK can fail, you needn't always be certain that it holds. Thus you can be sure of p without being sure that you know p. Moreover, these (un)certainties can be in lock-step with your knowledge.[3] I take DILIGENCE to be a default hypothesis—§3.1 defends it further.

Applying DILIGENCE to our case of the unconfident examinee, we can infer that there could be a *diligent* agent for whom KK fails. There seems to be no reason to think this agent couldn't be Kim herself—so for simplicity suppose it's so. Kim can

[2] If you think a notion of *belief* weaker than certainty corresponds to this internal component of knowing, the argument generalizes: see §3.1.

[3] Importantly, DILIGENCE does *not* say that KK can fail for an agent who *knows* that her certainties match what she can know—it says merely that KK can fail for an agent whose certainties *in fact* match what she can know.

know that Padua's in Italy but she cannot know that she can know it—and she's in fact certain of all and only the things she can know:

(2) $Kp \ \wedge \ \neg KKp \ \wedge \ (\forall q)(Kq \leftrightarrow [C(q) = 1])$

I claim that from (2) it follows that:

(3) Kim can know that if she can't know that Padua's in Italy, it's in Italy.

I mean it as I say it—interpret (3) in English, as knowledge of an indicative conditional. It's a bad conditional—an abominable one. Let '$q \rightarrowtail p$' symbolize the indicative sentence 'if q, p' as uttered in the context under discussion (in this case, Kim's).[4] My claim is that (2) implies:

ABOMINATION: $K(\neg Kp \rightarrowtail p)$

Here's where we need my second assumption. First, an example: suppose that you are sure that it'll rain, and you leave open that it might be cold. Then—if you're coherent—you'll be sure that *if it's cold, it'll rain.* Generalizing, certainties are *stable*: if you're sure of p and you leave open that q, then you are sure that *if q, p*:

STABILITY: If $C(p) = 1$ and $C(q) > 0$, then $C(q \rightarrowtail p) = 1$

The case for STABILITY strong. It of course holds for the material conditional (since $C(q \supset p) \geq C(p)$). And—as discussed in §3.2—STABILITY follows from standard assumptions about the Stalnaker-Lewis 'nearest-worlds' conditional, from natural implementations of the Kratzer 'restrictor' conditional and the strict conditional, and from a special case of the widely-attested Ramseyan thesis your credence in an indicative conditional equals your corresponding conditional credence.[5]

DILIGENCE and STABILITY are all we need to derive our unhappy consequence. From (2) we know that Kim is diligent, so the fact that she can know that Padua's in Italy (Kp) means that she's also sure that it is:

[4] Strictly, I should be more careful about use/mention, and write context superscripts on all expressions to indicate where they are uttered. But—so long as we keep in mind that (e.g.) the semantic values of conditionals can vary across contexts—these technicalities won't matter for our purposes.

[5] Although STABILITY might seem reminiscent of the PRESERVATION condition that Bradley (2000) uses to prove a triviality result, there is a crucial difference. Bradley's condition says that $q \rightarrowtail p$ expresses a proposition such that *for any probability function* P, if $P(p) = 0$ and $P(q) > 0$, then $P(q \rightarrowtail p) = 0$. In contrast, STABILITY says that the proposition expressed by the conditional (if such there be) *in a given context* is coordinated with a *single* probability function—namely, the credence function of the agent of that context (see footnote 4). Thus STABILITY is analogous to the LOCAL PRESERVATION condition presented in Mandelkern and Khoo (2018); their proof that LOCAL PRESERVATION is nontrivial is analogous to my proofs in §3.2 that STABILITY follows from many (nontrivial) theories of conditionals.

(4) $C(p) = 1$

Since Kim *can't* know that she can know Padua's in Italy ($\neg KKp$), it likewise follows that she's not sure that she can know it: $\neg[C(Kp) = 1]$, i.e. $C(Kp) < 1$; thus she leaves open that she can't know it:

(5) $C(\neg Kp) > 0$

Applying Stability to (4) and (5), it follows that Kim is certain that if she can't know that Padua's in Italy, it's in Italy:

(6) $C(\neg Kp \rightarrowtail p) = 1$

Finally, since Kim is diligent, the things she's sure of are things she can know—so she can *know* that if she can't know that Padua's in Italy, it's in Italy:

Abomination: $K(\neg Kp \rightarrowtail p)$

Upshot: if KK fails—if Kim can know p without being able to know that she can—then a diligent counterpart of her can know the abominable conditional: 'If I can't know that Padua's in Italy, it's in Italy.'

For ease of exposition, I'll make one further assumption. Suppose that Kim (knows that she) has thought through the (relevant) questions carefully, and hence: she knows everything she's in a position to know.[6] Thus in our discussion we can replace the phrase 'Kim can know q' with 'Kim knows q'. Abomination can then be read as follows: 'Kim knows that if she doesn't know Padua's in Italy, it's in Italy.'

This is a bad result. First things first: it is simply an odd state to be in. Kim thinks to herself:

> 'Maybe I don't know whether Padua's in Italy. Now, if I don't know it's in Italy, then it is. And of course if I *do* know it's in Italy, then it is. So whether or not I know it, it's in Italy... But I'm still not sure whether I know it's in Italy.'

That's not a paradox. She doesn't *know* that the reasoning she went through was based on known premises (though it was), so she can't infer that she knows the conclusion (though she does). But it's an odd bit of reasoning indeed.

To draw it out, suppose we endorse a knowledge norm of action (cf. Hawthorne and Stanley, 2008)—roughly, you're warranted in acting on your knowledge. Suppose Jill offers Kim the following conditional bet: 'If you know Padua's in Italy, the bet's off;

[6] Since, conversely, she's in a position to know everything that she knows, we have a biconditional. Letting $\mathbb{K}q$ be the claim that she knows q and Kq be the claim that she's in a position to know q: ($\forall$ relevant q) $\mathbb{K}(\mathbb{K}q \leftrightarrow Kq)$.

if you don't know Padua's in Italy, you win \$0.01 if it is, and lose \$1.00 if it's not.' Knowing that if she doesn't know p, it's true, Kim should act on that knowledge and take the bet—aware, of course, that she'll make money iff she's being irrational. (If she doesn't know p, it's irrational to take such an imbalanced bet.)

If this reasoning and acting is odd enough—and, I think, it's pretty odd—then we have reason to think ABOMINATION is false: Kim *can't* know that if she doesn't know Padua's in Italy, it's in Italy.

But we can go further. Kim knows that if she doesn't know Padua's in Italy, it's in Italy. So given a knowledge norm of assertion (Williamson, 2000), she should be able to assert the abominable conditional (cf. DeRose, 1995; Greco, 2014):

(7) #If I don't know Padua's in Italy, it's in Italy.

Such a conditional is manifestly infelicitous. Of course, (7) might suggest a connection between Kim not-knowing p and its truth. But that's not the source of the problem. 'Even if...' conditionals disavow any such connection (example: 'Even if Bill doesn't study, he'll pass'). Yet the 'Even if...' variation on (7) is just as abominable:

(8) #Even if I don't know it, Padua's in Italy.

Further, 'Whether or not...' conditionals also disavow any such connection. Kim knows that if she doesn't know Padua's in Italy, it is. And Kim also knows that if she *does* know Padua's in Italy, it is—knowledge is factive, after all. Combined, she should be able to assert:

(9) #Whether or not I know it, Padua's in Italy.

But (9) is—yet again—abominable.

More can be said. It turns out that there are some contexts in which our conditionals are not abominable—witness:

(10) Suppose that Padua's in Italy. Then even if I don't know it, Padua's in Italy.

(10) sounds fine because the first sentence guarantees that the ensuing context *presupposes* that Padua's in Italy—all possibilities that are live for the purposes of (that moment of) the conversation are ones in which Padua's in Italy (Stalnaker, 1970a, 1998).

My claim is that abominable conditionals are infelicitous in contexts that *don't* presuppose the truth of their consequents.[7] They are analogous to Moorean sentences of

[7]Two notes. First, sometimes we use the word 'know' in a way that is not operative in our context. In response to a skeptic who questioned all knowledge I might say, 'Even if I don't *know* that I have hands, I have hands.' But I might equally say, 'Even if I don't *know* that I have hands, I know I have hands'. Thus '*know*' must be getting a different interpretation than 'know'.

Second, some constructions trigger a presupposition that p even when embedded—e.g. 'Ada doesn't

the form *p, but I don't know it*. To see this, notice that although Moorean sentences are uniformly infelicitous when stated outright, they are fine in contexts that presuppose the truth of their first conjuncts:

(11) Suppose that it'll rain in two weeks. Then it'll rain, but I don't know that it will.

Moreover, Moorean sentences are also felicitous when ascribed third-personally or embedded in non-factive constructions:

(12) Padua's in Italy, but Jim doesn't know that it is.

(13) Sue said that Padua's in Italy but that I don't know that it is.

Likewise for abominable conditionals:

(14) Even if Jim doesn't know that Padua's in Italy, it's in Italy.

(15) Sue said that even if I don't know Padua's in Italy, it's in Italy.

The felicity of (10), (14), and (15) shows that—like Moorean sentences—there is nothing *semantically* defective about our conditionals. Rather, their infelicity must stem from the pragmatics of asserting them. The natural hypothesis, then, is that our explanation of the infelicity of abominable conditionals should be the same as our explanation of the infelicity of Moorean sentences.[8]

Why are Moorean sentences infelicitous (when they are)? The first part of the story is fairly standard. Moorean sentences are *blindspots*: consistent claims that a speaker cannot know (Sorensen, 1988; Cresto, 2017).[9] By the knowledge norm of assertion, you have epistemic warrant to assert p only if you know p (Williamson, 2000). Therefore you never can have epistemic warrant to assert a Moorean sentence.

Or so goes the standard story. It works fine for contexts with no special presuppositions, but it must be refined to handle the fact that Moorean sentences are felicitous in contexts that presuppose the truth of their first conjuncts. The crucial point is that—properly formulated—the knowledge norm should be understood *incrementally*, as requiring knowledge of the content you are *adding* to the common ground in making your assertion (Ichikawa 2017; cf. Moss 2012). There are a variety of ways to make this thought precise, but the basic idea is straightforward: if our conversation presupposes some claim q (which I may not know to be true), I have epistemic warrant to assert another claim p so long as I know that *if* q is true, then p is too—even if I don't know p

know that John is coming to the party' seems to presuppose that John *is* coming (Kiparsky and Kiparksy, 1970). Thus sometimes the context in which our conditional is asserted can be made to presuppose its consequent without surrounding material.

[8]Thanks to a referee for emphasizing this analogy.

[9]*Reductio:* suppose you know *p but I don't know p*. Then you know p, and you know that you don't know p. But if you know p, you *can't* know that you don't know p (knowledge is factive). Contradiction.

unconditionally. For example, if you say, 'Let's assume it'll rain tomorrow', I can follow up with, 'Okay. The picnic will have to be canceled'. I may not know that the picnic will have to be canceled, but my assertion is felicitous because I know that *given our presuppositions* (namely, that it will rain tomorrow) the picnic will have to be canceled. More generally, the idea behind an incremental knowledge norm of assertion is this. Given a context set $\mathbb{C}$—the set of possibilities consistent with our presuppositions—you have epistemic warrant to assert p when *conditional on* $\mathbb{C}$, you know p. This refinement of the knowledge norm permits an assertion of the Moorean sentence *p but I don't know it* in a context that presupposes p: all you must know is that *if* p is true, then 'p but I don't know that p' is true—as you will if you know that you don't know p.

In short, the patterns of (in)felicity of Moorean sentences are explained by the proper, *incremental* formulation of the knowledge norm of assertion, plus the fact that Moorean sentences are blindspots. Since the patterns of (in)felicity of our abominable conditionals are precisely parallel to those of Moorean sentences, the natural hypothesis is that abominable conditionals are *also* blindspots. In other words, Abomination is false: Kim *can't* know that if she doesn't know it, p.

If we endorse this natural hypothesis, we need KK. For from a failure of KK we derived Abomination. Contraposing: if Abomination is false, KK holds. The only way to resist this conclusion is to deny Diligence or Stability—§3 argues that this strategy is untenable.

Another option is to deny our natural hypothesis: perhaps, unlike Moorean sentences, abominable conditionals *can* be known—but some *other* feature explains their patterns of (in)felicity. I'll consider a variety of versions of this strategy—§4 argues that none of them work. In order to explain the infelicity of our abominable conditionals, we *must* deny Abomination.

That means we must endorse KK. Strictly speaking, this is all I need for my argument. But I'll offer you more: I'll argue that KK is also *sufficient* to explain the unknowability of abominable conditionals. The next section explains why.

2 Solution

In this section—and only this section—I will assume that what you can know is closed under known consequence. Combining this with KK, it follows that Kim knows an abominable conditional only if she knows that its antecedent is false: $K(\neg Kp \rightarrowtail p)$ only if KKp. Why? Suppose Kim knows that if she doesn't know p, then p. Knowledge of this indicative conditional implies knowledge of the material conditional: $K(\neg Kp \supset p)$. Moreover, since Kim knows that knowledge is factive, she knows (materially) that if she *does* know p, then p: $K(Kp \supset p)$. Conjoining these pieces of knowledge, it follows that

she can know p: Kp. Applying KK, it follows that she knows that she knows p: KKp.

So—given KK and closure—Kim knows an abominable conditional only if she rules out its antecedent. This will be crucial in explaining the unknowability of such conditionals. In fact—since indicative conditionals are standardly taken to presuppose the possibility of their antecedents (e.g. Willer, 2017)—we might think that this fact *on its own* suffices to explain why Kim can't know our abominable conditionals.

But it doesn't. Sometimes we *can* know an indicative conditional while ruling out its antecedent. Witness:

(16) I know that Oswald shot Kennedy. But I also know that if he didn't, then someone else did.

(17) I know that plenty of people exist. But I also know that if no one exists, then I don't exist.

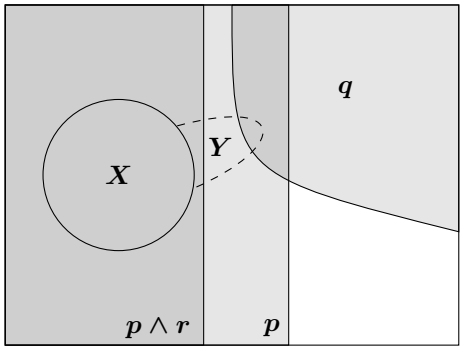
On the other hand, sometimes we can't. Suppose you tell me that Steph stole some cookies from the jar. Then:

(18) I know Steph stole some cookies from the jar. But I *don't* know that if she didn't, then someone else did.

(After all: for all I know, if Steph didn't steal cookies, then no one did.)

What's going on with (16)–(18)? Proposal: in (16) and (17) I know the conditional because my knowledge of it *outstrips* my knowledge that the antecedent is false; not so in (18). If you remove my knowledge that Oswald shot Kennedy, then I still know that *someone* did—and so still know that if Oswald didn't, someone else did. If you remove my knowledge that anyone exists, then I still know logic—so still know that if no one exists, then I don't exist. In contrast: if you remove my knowledge that Steph stole some cookies, then for all I know *no one* stole cookies (that was my only evidence that someone did)—so for all I know, if Steph didn't steal cookies, then no one did.

Upshot: knowledge of an indicative conditional must be *robust* with respect to the antecedent—if we minimally weaken your knowledge so that it doesn't rule out the antecedent, then you must still know the conditional. Precisely, let K_q be an operator corresponding to the minimal weakening of your knowledge that leaves open q, i.e. the strongest portion of your knowledge such that $\neg K_q \neg q$. Think of K_q like this. Take the set of worlds X that are consistent with your knowledge. You know p (Kp) iff every X-world is a p-world. Now take the set of (epistemically) closest worlds Y that contains some q-worlds, and add it to X. You q-know p ($K_q p$) iff every $X \cup Y$-world is a p-world. For example in the following picture, you both q-know and know p (since $X \cup Y \subseteq p$), but you only know (you don't q-know) $p \wedge r$ (since $X \cup Y \not\subseteq (p \wedge r) \supseteq X$):



My proposal is that you can know a conditional 'If q, p' iff you q-*know* 'If q, p', i.e. iff upon minimally weakening your knowledge to leave open q, you still know 'If q, p':

ROBUSTNESS: $K(q \rightarrowtail p)$ iff $K_q(q \rightarrowtail p)$

Some observations:

First, ROBUSTNESS is a constraint—not an analysis—so there is no problem with its circularity.

Second, since it doesn't matter precisely how we fill in the formal details of the K_q operator, I'll simply work with an intuitive understanding of it.[10]

Third, ROBUSTNESS is trivial in cases where your knowledge leaves open q—for if $\neg K \neg q$, then the state that results from weakening your knowledge as little as possible to make it leave open q is simply your original knowledge state: $K_q = K$. Moreover, if we assume that knowledge of 'if q, p' is knowledge of a proposition, then the right-to-left direction of ROBUSTNESS is also trivial: K_q is defined to be weaker than K, so $K_q(r)$ implies $K(r)$—and in particular, $K_q(q \rightarrowtail r)$ implies $K(q \rightarrowtail r)$.

Fourth, ROBUSTNESS follows the natural hypothesis that you know 'If q, p' iff you have *conditional knowledge* of p given q. The standard Levi Identity from the belief revision literature states that you know p conditional on q iff: if we first contract your knowledge so that it leaves open q, and then add q to this contracted knowledge state, the resulting knowledge-state implies p (Levi, 1977; Gärdenfors, 1981; Stalnaker, 1984, 2006). Under the plausible assumption that when you leave open q, you know p conditional on q iff you know $q \supset p$, we can equivalently state the Levi Identity as follows. You know p conditional on q iff: upon weakening your knowledge to leave open q, you know the material conditional $q \supset p$. This proposal shares ROBUSTNESS's key structural features of weakening your knowledge-state to leave open q. If we further make

[10]The details clearly *can* be filled in, e.g. by understanding it as the result of an AGM contraction operator that removes $\neg q$ from your knowledge (Alchourrón et al., 1985). Although this theory is standardly developed syntactically, it can be given a possible-worlds semantics like the one I use in the text (Stalnaker, 2006).

the (standard) assumption that when your knowledge leaves open q you know $q \rightarrowtail p$ iff you know $q \supset p$, then ROBUSTNESS follows.

Finally, ROBUSTNESS gets our cases right. Why do I know 'If Oswald didn't shoot Kennedy, then someone else did'? Because if you weaken my knowledge so that I leave open that he didn't, then I still know that *someone* did, so this weakened state knows the conditional. Why do I know 'If no one exists, then I don't exist'? Because if you weaken my knowledge so that I leave open that no one exists, I'll still know the logical truth that *no one exists* implies *I don't exist*. Why *don't* I know 'If Steph didn't steal the cookies, then someone else did'? Because if you weaken my knowledge so that I leave open that she didn't, then I leave open that no one stole any cookies—so I don't know that that if Steph didn't steal any cookies, someone else did.

Upshot: ROBUSTNESS is a well-motivated constraint. And when combined with KK, it rules out knowledge of abominable conditionals. For if KK holds, then it's a structural constraint on knowledge—meaning that the minimal weakening of your knowledge that leaves open q (K_q) will obey KK: $K_q p \rightarrow K_q K_q p$. Roughly, what this means is that when we weaken your knowledge to leave open that you don't know p, we thereby weaken it to leave open $\neg p$. Since you can know an abominable conditional only if you know the consequent, it follows that your weakened knowledge state cannot know the conditional—hence, by ROBUSTNESS, *you* cannot know an abominable conditional.

Now precisely. By definition, weakening your knowledge to leave open that you don't know p leaves you unable to know that you know p, i.e. you cannot $\neg Kp$-know that you know p: $\neg K_{\neg Kp} \neg (\neg Kp)$, i.e.

(19) $\neg K_{\neg Kp} Kp$

Since (by definition) you $\neg Kp$-know p only if you know p (i.e. $K_{\neg Kp}(p) \Rightarrow Kp$), it follows that you don't $\neg Kp$-know that you $\neg Kp$-know p

(20) $\neg K_{\neg Kp} K_{\neg Kp}(p)$

By KK, you $\neg Kp$-know p only if you $\neg Kp$-know that you $\neg Kp$-know p (i.e. $K_{\neg Kp}(p) \Rightarrow K_{\neg Kp} K_{\neg Kp}(p)$). Contraposing, and combined with (20), you don't $\neg Kp$-know p:

(21) $\neg K_{\neg Kp}(p)$

As established above, you can know an abominable conditional $\neg Kp \rightarrowtail p$ only if you know p (i.e. $K(\neg Kp \rightarrowtail p) \Rightarrow K(p)$). By parallel reasoning, you can $\neg Kp$-know an abominable conditional $\neg Kp \rightarrowtail p$ only if you $\neg Kp$-know p (i.e. $K_{\neg Kp}(\neg Kp \rightarrowtail p) \Rightarrow K_{\neg Kp}(p)$). Contraposing, from (21) it follows that you can't $\neg Kp$-know the conditional:

(22) $\neg K_{\neg Kp}(\neg Kp \rightarrowtail p)$

Finally, since you can't $\neg Kp$-know the conditional, and $\neg Kp$ is the conditional's antecedent, by Robustness you can't know it:

$\neg$Abomination: $\neg K(\neg Kp \rightarrowtail p)$

Upshot: KK rules out the possibility of knowing our abominable conditionals. So KK defenders can endorse the natural hypothesis that abominable conditionals are infelicitous for the same reason that Moorean sentences are: they are unknowable.[11]

What about KK-deniers? They can either object to my claim that KK failures lead to the knowability of abominable conditionals, or else allow that such conditionals are knowable but offer an alternative explanation for why they are infelicitous. §3 replies to the first objection by defending my premises; §4 replies to the second by showing that the explanations on offer do not succeed.

3 Opposition: Knowability

I used two premises in my derivation of Abomination from a failure of KK. Diligence said that if KK can fail, then it can fail for an agent whose certainties match her knowledge; Stability said that if you are certain of p and leave open q, you are certain of $q \rightarrowtail p$. §3.1 defends the first; §3.2 defends the second.

3.1 Defending Diligence

Using sentential quantifiers and C to represent our agent's probabilistic credences, Diligence says that if KK fails then it can fail for a diligent agent who is certain of all and only the claims that she can know:

> Diligence: If it's possible that $(\exists p)(Kp \wedge \neg KKp)$, then it's possible that $(\exists p)(Kp \wedge \neg KKp)$ while $(\forall q)(Kq \leftrightarrow [C(q) = 1])$.

Three clarifications:

First, Diligence does not say that KK can fail for an agent who *knows* (or believes) that her certainties match her knowledge. If agents can always tell whether or not they

[11] *Question:* how can *adding* a principle (like KK) to our logic of knowledge *remove* a consequence (like $\neg Kp \rightarrowtail p$)? This can sound paradoxical—logical consequence is monotonic, after all. *Answer:* distinguish two questions. First question: (i) Given a fixed set of propositions (someone's knowledge), does $\neg Kp \rightarrowtail p$ follow from it? If the answer to (i) is 'yes' for a given set of propositions when we don't assume KK, then adding KK to our logic of knowledge will not make the answer 'no'. But (i) is not our question—instead we ask: (ii) Is there a set of propositions that could both be someone's body of knowledge and also entail $\neg Kp \rightarrowtail p$? If the answer to (ii) is 'yes' when we don't assume KK (as I've argued it is), then the answer may become 'no' once we add KK (as I've argued it does)—for KK restricts the sets of propositions that can constitute someone's body of knowledge.

are certain of a given claim, then that latter claim is straightforwardly false. For if you know that your certainties match your knowledge, that you are certain of p, and that you are not certain that you know p, then you can come to know the Moorean conjunction 'I know p but I don't know that I know it'—which, of course, you can't. DILIGENCE does not have this consequence, for an agent can be diligent without knowing that she is diligent. (The biconditional in the universally quantified statement in DILIGENCE is a material biconditional.)

Second, DILIGENCE does not assert a *general* link between the possibility of having a given knowledge-state and the possibility of having that knowledge-state while being diligent. That claim is straightforwardly false: it's possible to know that you're not diligent, but it's not possible for a diligent agent to know that she's not diligent. DILIGENCE does not have this consequence, for all it asserts is that there is no systematic reason why KK failures *as a kind* cannot happen to a diligent agent.

Third, we don't even need the full strength of DILIGENCE. Our derivation only required an agent who knows p, doesn't know that she knows p, and obeys the following three instances of DILIGENCE: $Kp \leftrightarrow [C(p) = 1]$; $KKp \leftrightarrow [C(Kp) = 1]$; and $K(\neg Kp \rightarrowtail p) \leftrightarrow [C(\neg Kp \rightarrowtail p) = 1]$.

Clarifications in hand, DILIGENCE is surely the default hypothesis. Can KK deniers find principled reasons to reject it? I see three concerns: (i) perhaps certainty is *too strong* a state to be the internal component of knowledge; (ii) perhaps—since certainty is closed under consequence—DILIGENCE is in tension with fallibilism; and (iii) perhaps there's something illegitimate about applying DILIGENCE to an indicative conditional. I'll take them in turn.

Objection: We know plenty of things we're not certain of. In fact, it is precisely when we know p without knowing that we know p that we can know p without being certain of it.

Reply: We can modify the argument. Regardless of what you think about certainty, there must be *some* notion that corresponds to the internal component of knowledge. Let's call that notion *belief*.[12] Then everyone should agree that if KK can fail, it can fail for an agent whose *beliefs* match her knowledge:

> DILIGENCE$_B$: If it's possible that $(\exists p)(Kp \wedge \neg KKp)$, then it's possible that $(\exists p)(Kp \wedge \neg KKp)$ while $(\forall q)(Kq \leftrightarrow Bq)$.

Moreover, for this strong notion of belief that corresponds to the internal component of knowledge, an analogue of STABILITY is very plausible: if you believe p, and q is consistent with your beliefs, then (if you're coherent) you believe 'If q, p':

[12] Bearing in mind that this is a term of art and that the natural-language word 'belief' arguably picks out a much weaker state—note the felicity of sentences like 'I don't know if it'll rain, but I believe it will' (Hawthorne et al., 2016; Dorst, 2017).

STABILITY$_B$: If Bp and $\neg B\neg q$, then $B(q \rightarrowtail p)$.

Given the natural assumption that you believe 'If q, p' iff you have a conditional belief in p given q, STABILITY$_B$ corresponds to a standard axiom of belief-revision.[13]

Using reasoning precisely parallel to that used in §1, DILIGENCE$_B$ and STABILITY$_B$ imply that if KK can fail, it's possible for an agent to know an abominable conditional. In short: objecting that certainty is too strong is no way out.

Objection: The argument is inconsistent with a certain brand of fallibilism. Consider a fallibilist who thinks that you can know that you'll be at work tomorrow even though you can't rule out the possibility that you'll drop dead tonight. If we assume STABILITY and that you are diligent, it would follow that you can know the abominable conditional, 'If I drop dead tonight, I'll be at work tomorrow.' But you can know no such thing.

Reply: Two replies. First, such fallibilism is *not* inconsistent with my argument. For DILIGENCE does *not* assert a general connection between the possibility of a given knowledge-state and the possibility of having a diligent agent with that knowledge-state. Perhaps cases where you know p even though you can't rule out all counterpossibilities are cases where you have no diligent counterparts. All DILIGENCE assumes is that if KK fails, it is not a *necessary* truth that all KK failures are of this type.

Second, I think we should reject this brand of fallibilism. The problem is that it allows blatant failures of closure—allows that you might know you'll be at work tomorrow, know (materially) that if you're at work tomorrow then you won't drop dead tonight, but be unable to know that you won't drop dead tonight. It's well known that such a fallibilism permits abominable conjunctions (DeRose, 1995), e.g.

(23) #For all I know I'll drop dead tonight—but I know I'll be at work tomorrow.

My argument for KK is of a piece with this abominable-conjunction argument against fallibilism. Of course, plenty of fallibilists are not convinced by the argument from (23)—so they may similarly be unconvinced by my argument from abominable conditionals. This I grant—if you are okay with the closure failure in (23), perhaps you should be okay with abominable conditionals. But the most popular KK-denying pictures are *not* okay with (23) (e.g. Sorensen 1988 and Williamson 2000), for they make much use and significance of the (potential) closure of knowledge. So despite its tension with certain brands of fallibilism, my argument has plenty of teeth.

Objection: Precisely because of the interaction between probabilities and conditionals, we should not think that conditionals express propositions (Adams, 1975; Edgington, 1995; Bennett, 2003). Although DILIGENCE is kosher when applied to sentences that express propositions, it is not so for conditionals.

[13]The axiom, called 'Vacuity', holds that if Bp and $\neg B\neg q$, then $B(p|q)$ (Hansson, 2017).

Reply: We can modify the argument to avoid applying DILIGENCE to conditional sentences. Here's how. Since we regularly ascribe knowledge of indicative conditionals ('Pat knows that if Don doesn't come to the party, I'll be upset'), such a non-propositionalist view of conditionals must have a story about what we are up to when we do so. The view holds that to have a certain credence in a conditional $C(q \rightarrowtail p)$ really just is to have a corresponding conditional credence $C(p|q)$—rather than an unconditional attitude toward a conditional proposition, we have a conditional attitude toward an unconditional proposition. Likewise, then, such a view should say that *knowledge* of an indicative conditional $K(q \rightarrowtail p)$ is really a form of *conditional knowledge* $K(p|q)$. On this picture, to know that 'If Padua's in Italy, then it's not near Berlin' is to be such that if your knowledge were updated with the claim that Padua's in Italy, the resulting state would know that Padua's not near Berlin. Precisely:

> CONDITIONAL KNOWLEDGE: $K(q \rightarrowtail p)$ iff $K(p|q)$

This view has no quarrel applying DILIGENCE to non-conditional sentences, so it should accept a version that connects conditional certainties (in unconditional sentences) to conditional knowledge:

> CONDITIONAL DILIGENCE: Restricting quantifiers to non-conditional sentences, if it's possible that $(\exists p)(Kp \wedge \neg KKp)$, then it's possible that $(\exists p)(Kp \wedge \neg KKp)$ while $(\forall q, r)(K(q|r) \leftrightarrow [C(q|r) = 1])$.

We don't need STABILITY for this version of the argument—we simply need the assumption that your unconditional attitudes are those you have conditional on a tautology. (Precisely: $Kp \leftrightarrow K(p|q \vee \neg q)$ and $[C(p) = 1] \leftrightarrow [C(p|q \vee \neg q) = 1]$.) Given this, CONDITIONAL KNOWLEDGE, and CONDITIONAL DILIGENCE, it follows that the possibility of a KK failure implies the knowability of abominable conditionals.[14]

[14] *Proof.* If it's possible for KK to fail, by CONDITIONAL DILIGENCE, it's possible for KK to fail for an agent whose conditional certainties (in unconditional propositions) matches her conditional knowledge:

(24) $Kp \wedge \neg KKp \wedge (\forall q, r)(K(q|r) \leftrightarrow [C(q|r) = 1])$

From Kp we have $K(p|q \vee \neg q)$, so by (24) we have $[C(p|q \vee \neg q) = 1]$, and hence:

(25) $C(p) = 1$

Meanwhile from $\neg KKp$ we have $\neg K(Kp|q \vee \neg q)$, so by (24) we have $\neg[C(Kp|\ q \vee \neg q) = 1]$, so $\neg[C(Kp) = 1]$, so $[C(Kp) < 1]$, so:

(26) $C(\neg Kp) > 0$

Combining (25) and (26) with the ratio formula, $C(p|\neg Kp) = \frac{C(p \wedge \neg Kp)}{C(\neg Kp)} = \frac{C(\neg Kp)}{C(\neg Kp)} = 1$. Applying (24) we have $K(p|\neg Kp)$, and by CONDITIONAL KNOWLEDGE we get $K(\neg Kp \rightarrowtail p)$.

3.2 Defending Stability

My second premise was Stability: if you are sure that p while leaving open q, then (if you are coherent) you are sure that 'If q, p':

Stability: If $C(p) = 1$ and $C(q) > 0$, then $C(q \rightarrowtail p) = 1$

The case for Stability is strong.

First, it follows from a special case of the widely-endorsed Ramseyan thesis that the probability of a conditional is the corresponding conditional probability.[15] For if $C(p) = 1$ and $C(q) > 0$, then by the ratio formula $C(p|q) = 1$, which by the Ramseyan thesis implies $C(q \rightarrowtail p) = 1$.

Second, as discussed in footnote 5, there is no threat of triviality from Stability.

Third, Stability follows from the Stalnaker-Lewis closest-world semantics, under the standard assumption that worlds consistent with your certainties are closer to each other than to worlds inconsistent with your certainties (Stalnaker, 1975, 275).[16]

Fourth, Stability follows from a natural implementation of both the restrictor and the strict semantics for the conditional (Kratzer, 1986; Williams, 2008).[17]

Finally, consider the odd results of denying Stability. Suppose I'm certain that we'll have fun, I leave open that it'll rain, but I'm not certain that 'If it rains, we'll have fun.' Then we could have the following exchange:

You: 'Will it rain?'

Me: 'Maybe.'

You: 'Well, if it rains, will we still have fun?'

Me: 'I'm not sure.'

You: 'Darn.'

[15] See e.g. Ramsey 1931; Stalnaker 1970b; Adams 1975; van Fraassen 1976; Edgington 1995; Bennett 2003; Khoo 2013, 2016; Rothschild 2013; Bacon 2015.

[16] *Proof.* Suppose $C(p) = 1$ and $C(q) > 0$. Since q is consistent with your certainties, for an arbitrary w consistent with your certainties, the set of closest q-worlds to w is a set Y that is also consistent with your certainties. Since $C(p) = 1$, every Y-world is a p-world (ignoring infinitary complications). Thus the closest q-worlds to w are all p-worlds, so $q \rightarrowtail p$ is true at w. Since w was arbitrary, $C(q \rightarrowtail p) = 1$.

[17] Assume that (1) when $C(q) > 0$ at world w the modal base at w picks out the set of worlds consistent with your certainties at w, and (2) your certainties are introspective: $[C(r) = 1] \Rightarrow [C(C(r) = 1) = 1]$ and $[C(r) < 1] \Rightarrow [C(C(r) < 1) = 1]$. *Proof:* Suppose $C(p) = 1$ and $C(q) > 0$ are true at w. By (1), it follows that all q-worlds in the modal base are p-worlds (again ignoring infinitary complications), so $q \rightarrowtail p$ is *true* at w on the restrictor or strict semantics. Now, $C(p) = 1$ at w, and since $C(q) > 0$ at w, likewise $C(\neg q) < 1$. By assumption (2), all worlds consistent with your certainties are ones in which $C(p) = 1$ and $C(\neg q) < 1$ (hence $C(q) > 0$)—so by the above reasoning they are all ones at which $q \rightarrowtail p$ is true. Hence $C(q \rightarrowtail p) = 1$, as desired.

Me: 'Oh don't worry. It might rain; and if it rains we might not have fun. #But I'm sure we'll have fun.'

In short, it is difficult to deny STABILITY. Let me briefly address one way of doing so.

Objection: One way to understand the 'Ramsey test' (Ramsey, 1931) is as follows: the conditional 'If q, p' (as uttered by you) is true only if: upon adding q hypothetically to your knowledge, p follows. In other words, the conditional $q \rightarrowtail p$ is true (if and)[18] only if you have *conditional knowledge* of p given q:

STRONG RAMSEY: '$q \rightarrowtail p$' is true (if and) only if $K(p|q)$

If we combine STRONG RAMSEY with the assumption that you can be certain of p without being certain that you know p, we can refute STABILITY.[19]

Reply: So much the worse for that combination, I say. Since STABILITY follows from (i) the conditional-probability version of the Ramseyan thesis, (ii) the Stalnaker-Lewis semantics, and (iii) natural implementations of the strict and restrictor semantics, we have good reason to reject STRONG RAMSEY. Moreover, we can explain away its appeal.

First, let me more directly illustrate STRONG RAMSEY's inconsistency with the conditional-probability Ramseyan thesis. Suppose that you're certain that you know that the coin I'm holding is fair:

(27) $C(K(\textit{fair})) = 1$

And suppose that you're certain that your certainties match your knowledge on all relevant claims about the coin:

(28) $(\forall \text{ relevant } q)\ C(Kq \leftrightarrow [C(q) = 1]) = 1$

Finally, suppose you're certain of the (material) conditional that if the coin is fair, you don't have conditional knowledge that it will land heads if flipped:

(29) $C(\textit{fair} \supset \neg K(\textit{heads}|\textit{flip})) = 1$

Now consider how confident you are in the indicative conditional *if the coin is flipped, it'll land heads*. Intuitively—and by the conditional-probability Ramseyan thesis—you are $\frac{1}{2}$ confident of this claim: $C(\textit{flip} \rightarrowtail \textit{heads}) = \frac{1}{2}$. (After all, you know it's a fair coin. Suppose it's flipped. How confident are you that it'll land heads? $\frac{1}{2}$. So you're $\frac{1}{2}$ confident that if it's flipped, it'll land heads.)

[18] We only need the left-to-right direction for this discussion.

[19] *Reductio:* Suppose $C(p) = 1$ and $C(Kp) < 1$. Then $C(\neg Kp) > 0$, so by STABILITY $C(\neg Kp \rightarrowtail p) = 1$. By STRONG RAMSEY, $\neg Kp \rightarrowtail p$ implies $K(p|\neg Kp)$, and therefore we have $C(K(p|\neg Kp)) \geq C(\neg Kp \rightarrowtail p) = 1$. On any reasonable logic for conditional knowledge, you can know p conditional on $\neg Kp$ only if you can know p: $K(p|\neg Kp) \Rightarrow Kp$. So since you're certain of the former, you're certain of the latter: $C(Kp) = 1$, contradicting our initial hypothesis that $C(Kp) < 1$.

However, Strong Ramsey predicts that you have credence *0* in it. First, the intuitive reason: Strong Ramsey predicts that your credence that *if the coin is flipped, it'll land heads* is no higher than your credence that *conditional on the coin being flipped, I know it'll land heads*. But since you're certain that you know the coin is fair, you have credence 0 that you have conditional knowledge of how it'll land. By Strong Ramsey, this means you have credence 0 that *if the coin is flipped, it'll land heads*.

Next, rigorously. From (27) and (29), it follows that you are certain that it's not the case that conditional on the coin being flipped, you know it'll land heads:

(30) $C(\neg K(\mathit{heads}|\mathit{flip})) = 1$

Thus you have credence 0 in the claim that you have conditional knowledge of *heads* given *flip*:

(31) $C(K(\mathit{heads}|\mathit{flip})) = 0$

And by (the left-to-right direction of) Strong Ramsey, $(\mathit{flip} \rightarrowtail \mathit{heads})$ implies $K(\mathit{heads}|\mathit{flip})$; therefore $C(\mathit{flip} \rightarrowtail \mathit{heads}) \leq C(K(\mathit{heads}|\mathit{flip}))$, which by (31) equals 0. Thus:

(32) $C(\mathit{flip} \rightarrowtail \mathit{heads}) = 0$

Upshot: Strong Ramsey predicts that your credence in the indicative conditional *if this coin is flipped, it'll land heads* is *0*. But—knowing it's a fair coin—your credence in this conditional is in fact $\frac{1}{2}$. So Strong Ramsey is false.

Moreover, we can explain the appeal of Strong Ramsey without endorsing its problematic consequences. For it is plausibly true that:

Weak Ramsey: You have warrant to assert 'If q, p' only if $K(p|q)$

The difference, of course, is that whereas Strong Ramsey demands conditional knowledge for the *truth* of a conditional, Weak Ramsey demands conditional knowledge for its *assertability*. The latter is perfectly consistent with my assumptions, as well as with the conditional-probability version of the Ramseyan thesis. In fact, Weak Ramsey follows from the knowledge norm of assertion when combined with the innocuous assumption that you have conditional knowledge of p given q iff you know the indicative conditional 'If q, p': $K(p|q)$ iff $K(q \rightarrowtail p)$. Since Weak Ramsey captures the intuitive thought behind Strong Ramsey without the problematic consequences, we should endorse the former and reject the latter.

I conclude that it is quite difficult to deny Stability or Diligence in a way that avoids the inference from the possibility of a KK failure to the knowability of abominable conditionals. Therefore if KK can fail, Abomination can be true: sometimes you can know that if you don't know p, then p.

4 Opposition: Assertability

KK-deniers cannot avoid the knowability of abominable conditionals, so they cannot endorse the natural hypothesis that such conditionals have the same status as Moorean sentences. That's a cost.

However, the data that *must* be explained is the *infelicity* of asserting abominable conditionals (in contexts that don't presuppose their consequents). Recall:

(33) #If I don't know Padua's in Italy, it's in Italy.

(34) #Even if I don't know it, Padua's in Italy.

(35) #Whether or not I know it, Padua's in Italy.

Of course, knowability does not in general imply assertability. I know all sorts of things—from the fact that $2+2=4$ to the fact that Uncle Mo has a mole on his left foot—that are generally infelicitous to assert. So KK deniers may grant knowability but offer another explanation for why abominable conditionals are unassertable. Those who have considered my argument have offered a variety of such explanations. Some point to irrelevant antecedents (§4.1), or to Gricean norms (§4.2), or to a localized knowledge-norm (§4.3), or to beliefs about what you know (§4.4), or to other self-effacing speech acts (§4.5). None of them work. This section shows why.

4.1 Irrelevant antecedents

Objection: All I appealed to in order to infer the knowability and (hence) assertability of our abominable conditionals was that our diligent agent was certain of p and left open that $\neg Kp$; that yielded $K(\neg Kp \rightarrowtail p)$. So given *any* q that our diligent agent leaves open—Quebec's in Canada, say—parallel reasoning will yield the conclusion that $K(q \rightarrowtail p)$ is true: she knows that if Quebec's in Canada, Padua's in Italy. But here the antecedent is irrelevant to the consequent—is *that* the reason abominable conditionals are infelicitous?

Reply: No. If Kim knows p and is unsure about q, the conditional 'If q, p' is knowable and—in the right context—assertable. Example: while our group is trying to figure out the location of Padua, Pestering Pete is off-topic:

> Kim: 'Padua's in Italy.'
>
> Pete: 'What about Quebec? I think it's in Canada.'
>
> Kim: 'Whether or not it is, Padua's in Italy. That's what we care about right now.'
>
> Pete: 'But I'm right—right? Quebec's in Canada!'

Kim: 'It doesn't matter, Pete. If it is, Padua's in Italy. If it's not, Padua's in Italy. Stop bugging us about Quebec!'

Kim's reply is perfectly reasonable: conditionals with irrelevant antecedents *can* be known and asserted. Our abominable conditionals cannot—even in a parallel context:

Kim: 'Padua's in Italy.'

Pete: 'Do you know that?'

Kim: #'Whether or not I know it, Padua's in Italy. That's what we care about right now.'

The infelicity of our conditionals does not stem from their irrelevant antecedents.

4.2 Gricean Quantity

Objection: As noted above (§2), you can know the conditional 'If I don't know it, p' only if you in fact can know p. This may motivate a simple Gricean explanation of its infelicity: you are in a position to assert the conditional only if you're in a position to assert p itself—it's infelicitous because you've asserted something inexplicably weak (Grice, 1975).

Reply: Though elegant, this proposal doesn't work. First, just because an assertion is weaker than it could be doesn't mean that it'll be infelicitous. We just saw an example in §4.1: Kim can know 'If Quebec's in Canada, Padua's in Italy' only if she can know that Padua's in Italy. Yet that conditional *is* assertable in response to Pestering Pete's questions.

The Gricean may point to a related curious feature of our abominable conditionals: you are in a position to assert 'If I don't know it, p' only if you know p—i.e. only if the antecedent is false. But—although curious—this feature cannot explain the infelicity of our conditionals. The following conditionals are impeccable:

(36) If no one exists, then I don't exist.

(37) If no one knows anything, then I don't know anything.

By the knowledge norm, these conditionals are assertable only if they're known—only if their antecedents are false. Since they sound fine, this feature cannot explain why our abominable conditionals are infelicitous.

4.3 Local-Knowledge Norm

Objection: I've applied the knowledge norm of assertion straightforwardly to conditionals: it is knowledge of a conditional that warrants assertion of it. But perhaps it's

more perspicuous to think of (33)–(35) as *conditional assertions*, rather than assertions of conditionals. The idea is something like this: when you say 'If q, r' you update the context so that it includes q, and then in that new 'local' context you assert r. On this proposal, our norm should look like this:

> Local Knowledge: You may assert 'If q, r' only if: on the supposition that q, you know that r.

Local Knowledge problematizes our abominable conditionals without appeal to KK. For it implies that I may assert 'If I don't know it, p' only if: on the supposition that I don't know p, I know p. But on the supposition that I don't know p, I *don't* know p! Hence the assertion is infelicitous.

Reply: Though elegant, Local Knowledge is false. Recall that (37) is impeccable:

(37) If no one knows anything, then I don't know anything.

But Local Knowledge predicts it to have the same status as our abominable conditionals, for it implies that I may assert (37) only if: on the supposition that no one knows anything, I know that I don't know anything. But on the supposition that no one knows anything, I *don't* know that I don't know anything! Local Knowledge falsely predicts (37) to be infelicitous.

4.4 Unreasonable Assertion

Objection: Asserting our conditionals is infelicitous because it's *unreasonable*—even if you can know them, you can't reasonably *believe* you know them. If that's right, then we can explain their infelicity either by endorsing a 'reasonably believe you know' norm of assertion (Brown, 2008), or telling a story about how, given the knowledge norm, it's unreasonable to assert p if you don't believe that you know it.

Why can't you reasonably believe you know an abominable conditional? §1 showed that if you can know p without being able to know that you can, then it follows that you can know an abominable conditional. But to use this to reasoning to argue that you can reasonably *believe* you know the conditional, I'd have to assume that you can reasonably believe the Moorean conjunction 'I can know p but I can't know that I can.' Since belief aims at knowledge, and Moorean conjunctions are unknowable, you can do no such thing. So my argument doesn't establish that you can reasonably believe that you know an abominable conditional.

Reply: First, a minor issue. Plausibly, belief is *weak* in the sense that the the norms for believing something are less strict than the norms for asserting it (Hawthorne et al., 2016; Dorst, 2017). This shows up in the contrast between statements like:

(38) $\left\{ \begin{smallmatrix} \text{I don't know} \\ \text{I'm not sure} \end{smallmatrix} \right\}$ that Padua's in Italy—but I $\left\{ \begin{smallmatrix} \text{believe} \\ \text{think} \end{smallmatrix} \right\}$ it is.

(39) $\#\left\{ \begin{smallmatrix} \text{I don't know} \\ \text{I'm not sure} \end{smallmatrix} \right\}$ that Padua's in Italy—but it is.

If this is right, then—supposing KK fails—you *can* reasonably believe that you know p without knowing that you do. You might express it:

(40) $\left\{ \begin{smallmatrix} \text{I don't know} \\ \text{I'm not sure} \end{smallmatrix} \right\}$ if I know that Padua's in Italy, but I $\left\{ \begin{smallmatrix} \text{believe} \\ \text{think} \end{smallmatrix} \right\}$ I do.

So, if belief is weak, you *can* come to reasonably believe that you know an abominable conditional.

However, those who press this objection will likely have in mind a stronger notion of belief—something like *being sure*. Plausibly, if you are rationally certain that you can't know that p, you can't be rationally certain that p—writing Cq for *you're rationally Certain of* q: $C\neg Kp \rightarrow \neg Cp$. Given such a connection between certainty and knowledge, you cannot reasonably be certain that: you can know p but can't know that you can know it.[20] Since you cannot reasonably be certain of this conjunction, my argument from §1 does not establish that you can reasonably be certain that you know an abominable conditional.

Granted. But the failure of my argument to establish this conclusion does not show that the conclusion is false. In fact, it's true. For although I used an *actual* KK failure to derive knowledge of an abominable conditional, we are now in a position to show that merely the *nearby possibility* of a KK failure is enough to know an abominable conditional. If KK can fail, you can reasonably be certain that there's a nearby case in which it does—and so you can reasonably be certain that you know an abominable conditional.

I'll first run through a schematic version of the argument, and then work through a particular case. The idea rests on Robustness: you know that *if* q, p iff upon minimally weakening your knowledge to leave open q, you likewise know *if* q, p. Recalling that we write K_q for this minimally weakened knowledge state:

Robustness: $K(q \rightarrowtail p)$ iff $K_q(q \rightarrowtail p)$

The case for Robustness is strong (§2), and here we'll only need the right-to-left direction (which, recall, is trivial if $q \rightarrowtail p$ expresses a proposition and is plausible regardless).

First, the schematic argument. Suppose KK can sometimes fail. This means that sometimes your knowledge of p is *more robust* than your knowledge that you know p—the closest worlds in which you don't know p are closer than the closest worlds in which $\neg p$. So if we minimally weaken your knowledge so that you no longer know that you know p, we'll add the closest set of worlds containing $\neg Kp$-worlds. Since these worlds

[20] Suppose $C(Kp \wedge \neg KKp)$. Then CKp; but also $C\neg KKp$ and hence $\neg CKp$. Contradiction.

are *closer* than the closest $\neg p$-worlds (your knowledge of Kp is less robust than your knowledge of p), doing so won't add any $\neg p$-worlds. So this weakened knowledge-state will still know p, and so will know the abominable conditional, 'Even if I don't know it, p'. Thus to show that you can be certain that you know an abominable conditional, all we need is a case in which: you're certain that you know p and also certain that this knowledge of p is more robust than your knowledge that you know p. If you are also certain of ROBUSTNESS, you will be certain that you know 'Even if I don't know it, p.'

Turn to a specific case. Mr. Magoo is staring at a tree some distance off (Williamson, 2000). His eyesight is fairly good, and the tree is in fact 200 inches tall, so he's certain that he knows that it's at least 150 inches tall. However, following KK-deniers, suppose Mr. Magoo is (rationally) certain that his knowledge about the tree's height requires a margin for error—that in any (relevantly similar) case in which the tree is exactly i inches tall, for all he knows it's $i - 10$ inches tall. In particular, he's certain that in any case where the tree is less than 110 inches tall, for all he knows it's less than 100.

I claim that, given this setup, Mr. Magoo should be certain that he knows the abominable (indicative) conditional, 'Even if I don't know it, the tree's at least 100 inches tall.' The basic reason is that since (by the margin-for-error premise) the worlds where he lacks this knowledge are closer than the worlds where the tree is less than 100 inches tall, weakening his knowledge to leave open the former won't lead it to leave open the latter—so by ROBUSTNESS, he knows the conditional. Since he can be certain of these facts, he can be certain that he knows the conditional.

Now more rigorously. By hypothesis, he's certain that he knows the tree is at least 150 inches tall. Since he's certain that he knows that the tree is taller than 110 inches, he's certain the closest possibilities in which the tree is less than 110 inches tall are (epistemically) closer than any in which it's less than 100 inches tall. Thus he's certain that minimally weakening his knowledge to leave open the possibility that the tree is less than 110 inches tall will not dislodge his knowledge that the tree is at least 100 inches tall. Letting $[< n]$' and $[\geq m]$ label the propositions that the tree is less than n and at least m inches tall, Mr. Magoo is certain that he $[<110]$-knows that the tree is at least 100 inches tall: $K_{[<110]}[\geq 100]$. Since (by definition) he $[<110]$-leaves open that it is less than 110 inches tall, he's certain that he $[<110]$-knows the conditional 'If the tree's less than 110 inches tall, it's still at least 100 inches tall': $K_{[<110]}([<110] \rightarrowtail [\geq 100])$. Since he's certain of (the right-to-left direction of) ROBUSTNESS, he's certain that he *knows* this conditional:

(41) $CK([<110] \rightarrowtail [\geq 100])$.

So he should be able to assert it:

(42) Even if the tree is less than 110 inches tall, it's at least 100 inches tall.

So far, this seems correct.

But using our margin-for-error premise, it's a short hop from (42) to an abominable conditional. Since Mr. Magoo is certain that he knows that his knowledge needs a margin for error, he is certain that he knows that if the tree is less than 110 inches tall, he doesn't know that it's at least 100 inches tall:

(43) $CK([{<}110] \rightarrowtail \neg K[{\geq}100])$

Moreover, since the nearest possibility in which he fails to know that the tree's at least 100 inches tall is simply one where it's shorter, he likewise is certain that he knows that if he doesn't know that it's at least 100 inches tall, then (that's because) it's less than 110 inches tall:

(44) $CK(\neg K[{\geq}100] \rightarrowtail [{<}110])$

If the closest p-worlds are q-worlds, and the closest q-worlds are p-worlds, then the closest p-worlds *are* the closest q-worlds; and therefore what's true at the closest p-worlds is also true at the closest q-worlds.[21] Thus from (43) and (44) it follows that Mr. Magoo is certain that the closest worlds in which he doesn't know that the tree is at least 100 inches *are* the closest worlds in which it is less than 110 inches tall. We've seen from (41) that he's certain that he knows that the closest worlds where it's less than 110 inches tall are ones in which it's at least 100:

(41) $CK([{<}110] \rightarrowtail [{\geq}100])$

Using (43) and (44) to sustitute into the antecedent, it follows that he's certain that he knows that the closest worlds where he doesn't know that it's at least 100 inches tall are ones in which it's still at least 100 inches tall:

(45) $CK(\neg K[{\geq}100]) \rightarrowtail [{\geq}100])$.

Therefore—even if proper assertion that p requires being certain that you know p—he should be able to assert:

(46) #Even if I don't know it, the tree's at least 100 inches tall.

Yet (46) is, once again, abominable.

Upshot: if KK fails, then—like Mr. Magoo—you will often be certain that your knowledge of p is more robust than your knowledge that you know p. This, in turn, means you *can* be certain that you know an abominable conditional.

[21] Precisely, any semantics for conditionals based on an ordering of worlds will validate the inference from $(q \rightarrowtail p) \wedge (p \rightarrowtail q)$ to $(q \rightarrowtail r) \leftrightarrow (p \rightarrowtail r)$.

4.5 Self-Effacing Speech Acts

Cohen and Comesaña (2013) point to similarly bizarre results of KK failures (cf. Sosa, 2009). In particular, suppose Kim knows Padua's in Italy even though she (knows that she) doesn't believe she knows it. By the knowledge norm she should then be able to assert the abominable conjunction:

(47) #Padua's in Italy, but I don't think I know that it is.

As with our abominable conditionals, (47) sounds infelicitous—and KK defenders have a ready explanation. What about KK deniers?

Williamson (2013) offers a reply to Cohen and Comesaña (2013) that can be broken two steps: he first offers an alternative explanation of the oddity of (47); and then suggests that KK gives the wrong sort of explanation, anyhow. I will explain and reply to each move, in turn.

Objection: (47) sounds bad because 'the second conjunct blatantly undermines the normally intended effect of the first conjunct on the audience, by giving reason not to rely on it' (Williamson 2013: 82). Since asserting 'I don't think I know p' gives your audience good reason to think you don't know p, the second conjunct gives them good reason not to listen to the first conjunct. Hence asserting (47) is self-defeating—it ensures your audience will not listen.

Reply: Grant that this explains the infelicity of (47). For it to further explain our abominable *conditionals*, these conditionals must be self-defeating—the speaker's assertion of the conditional would have to give the audience good reason for not listening to them. But this is simply not true: placing 'I don't know p' in the antecedent of a conditional does *not* give the audience positive reason for thinking you don't know p. This is even more pronounced with 'Even if...' conditionals, which positively suggest that the antecedent is *im*plausible. For instance:

(48) Even if I live forever, I'll never go to Kansas.

Yet, as we've seen, 'Even if...' versions of our conditionals are still abominable:

(34) #Even if I don't know it, Padua's in Italy.

Upshot: abominable conditionals are infelicitous *despite* the fact that they don't give the audience reason to accept the antecedent. The Williamsonian explanation cannot account for them.

Objection: Williamson claims KK gives the *wrong sort* of explanation of the linguistic data. Note that (given the knowledge norm), asserting 'I don't think I know p' amounts to asserting 'I don't think I have warrant to assert p.' Thus (47) amounts to

(49) #Padua's in Italy, but I don't think I have warrant to assert that it is.

Clearly this sounds bad—for exactly the same reason as (47). KK explains it by making knowledge—the thing that warrants assertion—a luminous condition: when it holds, you can know that it does.

But Williamson claims that this is the wrong strategy. For (49) is an instance of a broader phenomenon: *in general* performing a speech act A while expressing doubts that one has the authority to perform A is odd—even when such entitlements are clearly not luminous. For instance:

(50) #Stand at attention!—But I don't think I have the authority to command you. (Williamson, 2013, 82)

(51) #I pronounce you married—but I don't think I have the authority to do so.

(50) and (51) are infelicitous, despite the fact that you can have the authority to command a soldier or perform a marriage without knowing that you do (you can forgot the chain of command, or whether you've been ordained). Since having such authority is not a luminous condition, appealing to KK cannot explain the data. Similarly, we can add, with abominable *conditional* speech acts:

(52) #Even if I don't have the authority to do so, I pronounce you married.

Since Williamson claims KK cannot explain the oddity of (50)–(52), he concludes that it gives the wrong sort of explanation of our abominable conditionals and conjunctions.

Reply: Contra Williamson, the luminosity of knowledge *can* explain (50)–(52). Granted, having the *authority* to command or marry is not a luminous condition—it depends on external facts about chains of command or marriage laws. But having the authority to ϕ is not the same as being warranted to ϕ. Of course, it's a *necessary* condition: if you don't have the authority to command me, then you're not warranted in doing so. But it's insufficient. Suppose you're second-in-command, and we're all merrily eating dinner. Unbeknownst to anyone, our commander just died. Three seconds later—for no reason at all—you start barking orders at us: 'Johnson: mop the floors!' 'Smith: you're relieved from duty!' By pure luck your orders are authoritative—but they're far from warranted or felicitous!

Here, then, is how KK explains the infelicity of (50)–(52). First, having the authority to perform a speech act A is necessary but not sufficient for you to be warranted in performing A. Thus the oddity of (50)–(52) is parasitic on the oddity of, e.g.:

(50′) #Stand at attention!—But I don't think I have warrant to command you.

(52′) #Even if I don't have warrant to do so, I pronounce you married.

Second, having *warrant* to perform a speech act is equivalent to a particular sort of knowledge—and thus, given KK, will be a luminous condition.

Notice that we can perform speech acts by making assertions. In standard contexts, it's appropriate to command, 'Stand at attention!' iff it's appropriate to assert, 'You must now stand at attention' (cf. Portner, 2007); and it's appropriate to say, 'I pronounce you married' iff it's appropriate to assert 'You are now married.' More generally, here's a hypothesis:

> Assertive Acts: Given a context, for any speech act A there's a sentence p_A such that you're warranted in performing A iff you're warranted in asserting p_A.

Examples could be multiplied. You're warranted in declaring, 'I name this ship Voyager' iff you're warranted in asserting 'This ship will be called Voyager.' You're warranted in pronouncing, 'I'll end class here' iff you're warranted in asserting, 'Class is now over.' You're warranted in announcing, 'I confer this degree upon you' iff you're warranted in asserting 'You have now graduated.' And so on. Although the proposition may vary for different contexts, there's good reason to think we'll always be able to find one. Assertive Acts is a natural hypothesis.

If we accept it, KK explains the oddity of (50′) and (52′), and hence of (50)–(52). For you're warranted in commanding 'Stand at attention!' iff you're warranted in asserting, 'You must now stand at attention', iff (by the knowledge norm) you know the addressee must now stand at attention. Thus we can substitute the latter in for the former to see that (50′) is warranted iff (50″) is:

(50″) #You must now stand at attention—but I don't think I know that you must.

Likewise, (52′) is warranted iff (52″) is:

(52″) #Even if I don't know that you are now married, you are now married.

Yet (50″) and (52″) *just are* the abominable conjunction and conditional with which we began. As we've seen, KK elegantly explains their infelicity. Since they are felicitous iff (50′)–(52′) are, which in turn must be felicitous for (50)–(52) to be felicitous, it follows that KK *does* explain the infelicity of these self-effacing speech acts. Contra Williamson, KK gives a unified explanation of the data.

5 Coda

I have argued that there are robust linguistic patterns that support KK. Many who have heard this argument have agreed—yet remained incredulous: *How* could KK be true? The fact that Padua's in Italy doesn't imply that you know that it is: $p \not\Rightarrow Kp$. So how could being in a position to know the former always put you in a position to know the latter?

Although this can seem puzzling, it's really not so odd. Often we are in a position to know a proposition p only if we are in a position to know a stronger proposition p^+. An extreme case: *it's raining* does not entail *it's raining and someone exists*; but you can know the former only if you can know the latter. Mundane cases exist, too. *Padua's in Italy* does not entail *Padua's in Italy and I have some reason to think that it is*; but plausibly you can know the former only if you can know the latter. *The dog is outside* does not entail *the dog must be outside*; but plausibly you can know (or assert) the former only if you can know (or assert) the latter. *It's raining or snowing* does not entail (the indicative conditional) *if it's not raining, it's snowing*; but plausibly you can know (or assert) the former only if you can know (or assert) the latter (Stalnaker, 1975).

In each case there is a gap between p and p^+, but constraints on the structure of knowledge and its connection to semantic content prevent us from accessing that gap. The gap is a sort of blindspot (Sorensen, 1988). Likewise, says a KK-defender, for the gap between p and Kp: although there's a gap between Padua being in Italy and my knowing that it is, the structure of knowledge prevents me from accessing that gap. *Why* exactly this is so remains a difficult question. But—given these precedents—I do not think it is a hopeless one.

6 Conclusion

We began with a puzzle: if KK fails, abominable conditionals of the form 'If I don't know it, p' should be knowable and assertable (§1). Manifestly, they are not. If KK holds, we can explain why (§2). We then considered a variety of attempts to avoid this result—none of which held water (§§3–4). Upshot: we have new evidence for KK.

It is inconclusive, of course. But the evidence is mounting. We have one central reason to doubt KK (Williamson, 2000), and a variety of reasons for unease (cf. Sorensen, 1988) . But we also have half-a-dozen reasons to *believe* KK.[22] And we now know that there is robust linguistic evidence for it.

Make that half-a-dozen-plus-one.[23]

[22]Cf. McHugh (2008); Cresto (2012); Cohen and Comesaña (2013); Greco (2014, 2015, 2017); Stalnaker (2015); Das and Salow (2016).

[23]Thanks to Kai von Fintel, Justin Khoo, Matthew Mandelkern, Ginger Schultheis, Jack Spencer, Bob Stalnaker, and Roger White—as well as audiences at MIT, the 2018 Eastern APA, and two anonymous reviewers—for much helpful feedback and discussion.

References

Adams, Ernest, 1975. *The Logic of Conditionals*, volume 86 of *Synthese Library*. Springer Netherlands.

Alchourrón, Carlos E, Gärdenfors, Peter, and Makinson, David, 1985. 'On the logic of theory change: Partial meet contraction and revision functions'. *Journal of Symbolic Logic*, 50(2).

Bacon, Andrew, 2015. 'Stalnaker's Thesis in Context'. *The Review of Symbolic Logic*, 8(1):131–163.

Bennett, Jonathan, 2003. *A Philosophical Guide to Conditionals*. Clarendon Press.

Bradley, Richard, 2000. 'A preservation condition for conditionals'. *Analysis*, 60(267):219–222.

Brown, Jessica, 2008. 'The Knowledge Norm for Assertion'. *Philosophical Issues*, 18(1):89–103.

Cohen, Stewart and Comesaña, Juan, 2013. 'Williamson on Gettier Cases and Epistemic Logic'. *Inquiry*, 56(1):15–29.

Cresto, Eleonora, 2012. 'A Defense of Temperate Epistemic Transparency'. *Journal of Philosophical Logic*, 41(6):923–955.

———, 2017. 'Lost in Translation: Unknowable Propositions in Probabilistic Frameworks'. *Synthese*, 194(194):3955–3977.

Das, Nilanjan and Salow, Bernhard, 2016. 'Transparency and the KK Principle'. *Noûs*.

DeRose, Keith, 1995. 'Solving the Skeptical Problem'. *The Philosophical Review*, 104(1):1–52.

Dorst, Kevin, 2017. 'Lockeans Maximize Expected Accuracy'. *Mind*, To Appear.

Edgington, Dorothy, 1995. 'On Conditionals'. *Mind*, 104:235–329.

Gärdenfors, Peter, 1981. 'An Epistemic Approach to Conditionals'. *American Philosophical Quarterly*, 18(3).

Greco, Daniel, 2014. 'Could KK be OK?' *Journal of Philosophy*, 111(4):169–197.

———, 2015. 'Iteration and Fragmentation'. *Philosophy and Phenomenological Research*, 91(3):656–673.

———, 2017. 'Cognitive Mobile Homes'. *Mind*, 126(501):93–121.

Grice, Paul, 1975. 'Logic and Conversation'.

Hansson, Sven Ove, 2017. 'Logic of Belief Revision'. In Edward N Zalta, ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 201 edition.

Hawthorne, John, Rothschild, Daniel, and Spectre, Levi, 2016. 'Belief is Weak'. *Philosophical Studies*, 173(5):1393–1404.

Hawthorne, John and Stanley, Jason, 2008. 'Knowledge and Action'. *Journal of Philosophy*, 105(10):571–590.

Ichikawa, Jonathan Jenkins, 2017. *Contextualising Knowledge: Epistemology and Semantics*. Oxford University Press.

Khoo, Justin, 2013. 'Conditionals, Indeterminacy, and Triviality'. *Philosophical Perspectives*, 27(1):260–287.

———, 2016. 'Probabilities of Conditionals in Context'. *Linguistics and Philosophy*, 39(1):1–43.

Kiparsky, P and Kiparksy, C, 1970. 'Fact'. In Karl E. Bierwisch and M. Heidolph, eds., *Progress in Linguistics: A Collection of Papers*, 143–173. Mouton Publishers.

Kratzer, Angelika, 1986. 'Conditionals'. In *Proceedings of the Chicago Linguistics Society*, volume 22, 1–15.

Levi, Isaac, 1977. 'Subjunctives, dispositions and chances'. *Synthese*, 34(4).

Mandelkern, Matthew and Khoo, Justin, 2018. 'Against Preservation'. *Analysis*, To appear.

McHugh, Conor, 2008. 'Self-knowledge and the KK principle'. *Synthese*, 173(3):231–257.

Moss, Sarah, 2012. 'On the pragmatics of counterfactuals'. *Nous*, 46(3):561–586.

Portner, Paul, 2007. 'Imperatives and modals'. *Natural Language Semantics*, 15(4):351–383.

Radford, Colin, 1966. 'Knowledge: By Examples'. *Analysis*, 27(1):1–11.

Ramsey, Frank, 1931. *The Foundations of Mathematics and Other Logical Essays*. Kegan Paul, Trench, and Trubner & Co.

Rothschild, Daniel, 2013. 'Do Indicative Conditionals Express Propositions?' *Noûs*, 47(1):49–68.

Salow, Bernhard, 2017. 'Elusive Externalism'. *Mind*, to appear.

Sorensen, Roy A., 1988. *Blindspots*. Clarendon Press, Oxford.

Sosa, David, 2009. 'Dubious Assertions'. *Philosophical Studies*, 146(2):269–272.

Stalnaker, Robert, 1970a. 'Pragmatics'. *Synthese*, 2(1-2):272–289.

———, 1970b. 'Probability and Conditionals'. *Philosophy of Science*, 37(1):64–80.

———, 1975. 'Indicative conditionals'. *Philosophia*, 5(3):269–286.

———, 1984. *Inquiry*. Cambridge University Press.

———, 1998. 'On the representation of context'. *Journal of Logic, Language and Information*, 7(1).

———, 2006. 'On the Logics of Knowledge and Belief'. *Philosophical Studies*, 128(1):169–199.

———, 2015. 'Luminosity and the KK Thesis'. In Sanford Goldberg, ed., *Externalism, Self-Knowledge, and Skepticism*, 1–19. Cambridge University Press.

van Fraassen, Bas, 1976. 'Probabilities of Conditionals'. In William Harper and Clifford Hooker, eds., *Foundations of probability theory, statistical inference, and statistical theories of science.*, 261–308. Reidel.

Willer, Malte, 2017. 'Lessons fro Sobel Sequences'. *Semantics and Pragmatics*, 10(4).

Williams, J. Robert G., 2008. 'Conversation and Conditionals'. *Philosophical Studies*, 138:211–223.

Williamson, Timothy, 2000. *Knowledge and its Limits*. Oxford University Press.

———, 2013. 'Response to Cohen, Comesaña, Goodman, Nagel, and Weatherson on Gettier Cases in Epistemic Logic'. *Inquiry*, 56(1):77–96.