Knowledge by Constraint

Ben Holguín

October 2018

Abstract

This paper considers some puzzling knowledge ascriptions and argues that they present *prima facie* counterexamples to credence, belief, and justification conditions on knowledge, as well as to many of the standard meta-semantic assumptions about the context-sensitivity of 'knows'. A novel version of epistemic contextualism is developed to account for these ascriptions—one that takes the interpretation of 'knows' to be sensitive to the mechanisms of constraint. It is argued that the theory confirms the initial impressions on the principles tying knowledge to credence and justification, but is somewhat equivocal on the principle about belief.

1   Strange knowledge

Start with a case:

History Exam In preparation for the upcoming multiple-choice history exam, Peggy and Pete have both purchased and subsequently memorized answers sheets from Roger, their corrupt teaching assistant. However, moments before the exam Roger shares an unfortunate discovery: one of Peggy or Pete was given answers to an entirely different exam, and he doesn't know who it was. In fact, all he knows is that their answer sheets diverge on every question.

Question 5 is: "In what year did the Berlin Wall fall?". Peggy's answer sheet says b: 1989, Pete's says c: 1991. Peggy is thus the one with the good answers.[1]

Putting ourselves in Peggy's shoes, (1)–(4) all express similar and perfectly natural thoughts:

(1)   If my answer sheet is good, then I know what the answer to question 5 is.

(2)   One of us knows what the answer to question 5 is. (I hope it's me.)

(3)   Either Pete has the good answers, or I know what the answer to question 5 is.

(4)   I might know what the answer to question 5 is; it depends on whether I got the good answers.

And putting ourselves in Roger (the TA)'s shoes, we might think to ourselves:

---

[1] There are affinities between *History Exam* and cases discussed by Hawthorne (2000, pp. 202-3; 2004, pp. 68-9). As will be seen shortly, my analysis of *History Exam* is more detailed, and the lessons I draw are quite different.

(5)   Whoever has the answer sheet that says '5: b' knows what the answer to every question is.

And—still in Roger's shoes—if we discover that Peggy has the good answers, we may even assert something like (6) outright:

(6)   Peggy knows what the answer to question 5 is.

Taking our judgments about (1)–(6) at face value, it looks to follow from the fact that Peggy has the good answers that she knows what the answer to question 5 is (and likewise for the other questions). This kind of knowledge, if genuine, is rather strange. Among other things, it appears to be in tension with the following widely accepted principles about knowing:[2]

CREDENCE If S knows that P, then S's credence in P is > .5.[3]

BELIEF If S knows that P, then S believes that P.

JUSTIFICATION If S knows that P, then S is in a position to justifiably believe that P.

Start with CREDENCE. If Peggy is minimally rational then her credence in the proposition that the answer to question 5 is b can be no higher than .5. She knows that if and only if she got the good answers is the answer to question 5 b, and she knows that the odds she got the good answers are no higher than .5. Granted, Peggy knows that Pete's answer sheet is equally likely to say a, c, or d, and for that reason may be rationally compelled to put b rather than one of the other answers. But on the question of 'Is it b or not-b?', intuitively she can do no better than guess.

The following much weaker principle can be rejected along similar lines:

MINIMAL CREDENCE If S knows that P, then S's credence in P is > .001.

And that is because we can reimagine *History Exam* to make the odds more as less as low as we want without affecting the relevant judgments. Just imagine, for instance, that there are 1,000 possible answers to a given multiple choice question and that there are 1,000 different students taking the test, each having purchased a different answer sheet from Roger, and that one and only one of the students has the good answers (and that Peggy knows all this). Supposing Peggy is the student with the good answers, (1)–(6) remain perfectly acceptable, even though now Peggy knows that the odds that the answer to question 5 is b are no higher than .001.

Now for BELIEF. If we think that principles like CREDENCE (or at least MINIMAL CREDENCE) hold of belief, then the above considerations provide a direct argument against BELIEF: Peggy knows that the answer to question 5 is b, but her credence in that proposition is not greater than .5 (and

---

[2] Readers may notice that (i) each of (1)–(6) uses 'knows *what*' rather than 'knows *that*'; that (ii) CREDENCE, BELIEF, and JUSTIFICATION are all stated in terms of 'knows that'; and that (iii) when '*what* the answer to question 5 is' is replaced with '*that* the answer to question 5 is b', the reports sound much less natural. We will return to these issues in §2.3. For now, however, I will assume that in the context of *History Exam*, if 'Peggy knows what the answer to question 5 is' is true, then so is 'Peggy knows that the answer to question 5 is b'.

[3] See, e.g., Rothschild and Spectre (2016, 2017) for arguments that make important use of this principle.

could be as low as .001). Belief requires credence greater than .5 (or .001), so knowledge cannot require belief.

It is also worth noting that with the exception of the fact that Peggy intends to fill in the answer 'b', Peggy lacks any of the ordinary hallmarks of belief. She would not assert that the answer to question 5 is b, nor would she bet on more than even odds that the answer to question 5 is b, nor would she try to convince anyone that the answer to question 5 is b. If Peggy believes the answer to question 5 is b, it is presumably in a sense that is unfamiliar to most of those interested in the truth of BELIEF.

Next, JUSTIFICATION. If Peggy were to believe that the answer to question 5 is b, would she be justified in doing so? For roughly the reason sketched in the discussion of CREDENCE, the answer seems to me an obvious no. The entirety of Peggy's evidence for the proposition that the answer to question 5 is b—at least on any remotely intuitive way of understanding 'evidence'—is that her answer sheet says as much. But Peggy also knows that her answer sheet has at most a .5 chance of being correct. And remember: we can modify the case to make that number as small as we want without affecting the relevant judgments.

The kind of knowledge raised to salience in *History Exam* is a strange one. If we take our intuitions about (1)–(6) at face value, then we have data that presents a strong challenge to a number of widely accepted epistemological principles.[4] Moreover, supposing the principles are ultimately false, the case seems not to suggest any straightforward explanation of *why* they would have seemed so plausible in the first place. Peggy's knowledge, if genuine, is radically unsupported by what we'd normally call the evidence, and is in most relevant respects detached from anything we'd normally call belief. There is no view of knowledge I am familiar with that can account for what is happening in *History Exam*. The aim of this paper is to offer one that does.

The paper is structured as follows. The next section will argue that certain natural debunking explanations of the data are inadequate (§2). We then develop a contextualist theory of knowledge that can account for the data (§3), and discuss issues concerning its meta-semantics (§4). After that there is a discussion of some over-generation worries for the view (§5), to which an amendment is proposed. The paper concludes (§6) with a discussion of what to make of principles like CREDENCE, BELIEF, and JUSTIFICATION in light of the theory.

---

[4] It also presents a challenge to principles like:

ANTI-LUCK If S knows that P, then if S believes (or were to believe) that P, it is not a matter of luck that S's belief is (or would be) true.

—at least on the assumption that 'luck' is understood intuitively, rather than in a way that would make the principle true by stipulation. I thus take cases like *History Exam* to be hostile to virtue-theoretic analyses of knowledge, like those of Sosa (2007) and Greco (2010).

2 Four objections

Before giving a theory of knowledge that is capable of accounting for *History Exam*'s strange uses of 'knows', it will be helpful to address some sources of skepticism about the value of such a project. This section will discuss four objections to the claim that our intuitive judgments about *History Exam* ought to be taken seriously for the purposes of semantic and epistemological theorizing. First, an objection concerning the generality of the phenomenon. Second, an objection concerning the scope of 'knows' at logical form. Third, an objection concerning the inference from 'knows-wh' to 'knows-that'. And fourth, an objection that says that knowledge ascriptions like (1)–(6) ought to be regarded as mere loose speech.

2.1 First objection: the generality of the phenomenon

A common first blush reaction to *History Exam* is to chalk up the strangeness of (1)–(6) to the distorting effects "exam cases" have on our judgments about knowledge. It is difficult to say exactly what the thought is here, but I assume the basic idea is that exam settings tempt speakers to treat 'knows' as meaning something along the lines of 'can correctly answer'. And since there is no reason to expect 'can correctly answer' to abide by any of CREDENCE, BELIEF, or JUSTIFICATION, we shouldn't be so surprised if, in a case like *History Exam*, we talked about knowledge as if it didn't either.

§2.4 will present some general reasons to be suspicious of these sorts of reactions to the data. For now, I simply want to point out the phenomenon raised to salience by *History Exam* seems to have nothing special to do with exam settings. I will give two cases in illustration.

First we have:

Restaurant Booking Betty has plans to meet Don in the city for dinner. Don made the booking and told Betty it was at Lutèce. Right before leaving for dinner, Betty gets a call from Don. He tells her that he has dinner reservations with two different people this week—one with her and one with his colleague Stan—and he just found out that he told one of the two the name of the wrong restaurant. (As it turns out, he gave Stan the wrong name. Betty's dinner really is at Lutèce.)

I submit that it would be perfectly natural for Betty to entertain thoughts like:

- (1_R)   If Don screwed up with Stan, then I know where I'll be having dinner tonight.
- (2_R)   Either Stan or I knows where our dinner with Don will be taking place.
- (3_R)   Either Don got it right with Stan, or I know where I'll be having dinner tonight.
- (4_R)   I might know where I'll be having dinner tonight; depends on whether Don screwed up with Stan or me.

Since Betty *in fact* got the correct information from Don, it follows that she knows where she'll be having dinner tonight. And once Don finds out that he screwed up with Stan, not Betty, he can think things to himself like:

(6_R)   Betty knows which restaurant she'll be eating dinner at tonight.

Likewise, if we follow Williamson (2000, ch. 1) in taking factive mental stative operators (FM-SOs) like 'remembers that', 'sees that', etc., to entail their 'knows that' counterparts, then cases like the following present yet more instances of strange knowledge:

Memory Experiment Joan and Megan are participating in a trial of a drug whose primary effect is to swamp its subjects with an extraordinary number of fake "memories" of the events of the past 24 hours. One of the subjects will get the drug, while the other will get a placebo. Who gets which is determined by a coin-flip whose result is known only to the experimenters.

During the experiment Joan and Megan are both (separately) asked 'Do you remember what you ate for dinner yesterday?' Joan appears to remember that she ate fish; Megan appears to remember that she ate spaghetti. As a matter of fact it was Joan who got the placebo and Megan who got the drug.

Joan can easily entertain thoughts like:

(1_M)   If I got the placebo, then I remember what I ate for dinner last night.

(2_M)   One of us remembers what she ate for dinner last night.

(3_M)   Either Megan got the placebo, or I remember what I ate for dinner last night.

(4_M)   I might remember what I ate for dinner last night; it depends on whether I got the placebo or the drug.

Since Joan *in fact* got the placebo, it follows that she remembers (and thus knows) what she ate for dinner last night. And since the experimenters know who got the placebo, they can say things like:

(6_M)   Joan remembers what she ate for dinner last night.

Examples multiply with ease (see §4.3 for further evidence of this). I contend that the underlying phenomenon brought out by *History Exam* has no essential connection to the idiosyncrasies of our judgments about knowledge in exam settings.

2.2 Second objection: wide-scoping

A popular reaction to conditional knowledge ascriptions like (1) ('If my answer sheet is good, then I know what the answer to question 5 is') is to claim that our intuitions about these sentences rest on non-obvious subtleties concerning the scope of 'knows' at logical form. In particular, the thought

is that the natural readings of (1) are *wide-scope* readings—that is, readings on which the attitude verb takes scope over the whole conditional, rather than just the embedded -wh clause. According to this view, the true proposition we associate with (1) is better represented by:

(1$_W$) I know whether: if my answers are good, then the answer to question 5 is b.

And it is obvious that it does not follow from (1$_W$) plus the fact that Peggy's answers are good that Peggy knows what the answer to question 5 is.

However, there are at least four reasons to think that the subtleties concerning the scope of attitude verbs in conditionals are beside the point.

First, there is no reason to expect scope ambiguities to explain the full range of data. *History Exam* raises to salience a number of sentences that are clearly *not* indicative conditionals that seem to pose exactly the same problem (1) poses on its narrow-scope interpretation: (2) is a quantified knowledge ascription, (3) is a disjunctive knowledge ascription, (4) is a knowledge ascriptions embedded under a modal, and (5)–(6) are plain old unembedded knowledge ascriptions. So even supposing we could in principle explain away the true readings of (1) in terms of something like a scope ambiguity, it is entirely non-obvious how the story is supposed to be extended to any of the rest of the data.

Second, as was just pointed out in §2.1, it is easy to get ascriptions analogous to *History Exam*'s (1)–(6) but with FMSOs like 'remembers' or 'sees' in place of 'knows'. The wide-scope response founders on these examples. If I'm unsure whether I've been given a placebo or a drug that causes auditory hallucinations, I'll be able to assert or think things like 'If I got the placebo, then I can hear whether the radio is on'. But I definitely cannot assert or think things like 'I can hear whether: if I got the placebo, then the radio is on'. The proposition *that if I got the placebo, then the radio is on* is just not the kind of thing I can hear, even if it is the kind of thing I can know.

Third, the wide-scope response also founders on third-person analogs of (1). Consider a slightly altered version of *History Exam* in which Peggy sells her answers to a third student, Ken, who believes falsely that he, Peggy, and Pete are all studying from the same answer sheet. Taking Peggy's perspective in this new setting, (1+) is fine even though its wide-scope analog (1+$_W$) seems false:

(1+) ✓ If Pete got bad answers, then Ken and I both know what the answer to question 5 is.

(1+$_W$) ? Ken and I both know whether: if Pete got bad answers, then the answer to question 5 is b.

(1+) is puzzling in the same ways (1) is, yet the former clearly cannot be explained away in terms of a scope ambiguity.

Fourth and finally, consider a conditional ascription like:

(7) ✓ If Pete doesn't know what the answer to question 5 is, then I do.

(7) seems to me as felicitous as any of (1)–(6). Indeed, it's basically just another way of saying what (2) says ('One of us knows what the answer to question 5 is'). But assuming 'knows' only has its normal readings, the wide-scope analog of (7) is clearly false:

(7_W) ? I know that: if Pete doesn't know what the answer to question 5 is, then the answer to question 5 is b.

Supposing there is no strange knowledge, then (7_W) is plainly false: regardless of whether Pete has the good answers, he doesn't know what the answer to question 5 is. So for all Peggy knows: (i) Pete doesn't know what the answer to question 5 is, and (ii) the answer to question 5 *isn't* b. Thus, Peggy can't know that if Pete doesn't know what the answer to question 5 is, then the answer to question 5 is b. The good readings of (7_W) are not wide-scope readings.

2.3   Third objection: knowledge-wh vs. knowledge-that

As was first observed in footnote 2, each of *History Exam*'s (1)–(6) involves 'knowledge-wh' rather than 'knowledge-that'. The reason these examples were stated in terms of knowledge-wh is because the knowledge-that counterparts of (e.g.) (1), (3), (4), and (6) sound noticeably less natural:[5]

(1*)   ? If my answer sheet is good, then I know that the answer to question 5 is b.

(3*)   ? Either Pete got the good answers, or I know that the answer to question 5 is b.

(4*)   ? I might know that the answer to question 5 is b; it depends on whether I got the good answers.

(6*)   ? Peggy knows that the answer to question 5 is b.

But the principles for which I claim these data raise problems—CREDENCE, BELIEF, and JUSTIFICATION—are all stated in terms of knowledge-that. So perhaps what cases like *History Exam* show is not that one can know without believing or having justification to believe, but rather that it is possible to know *what* the answer to a multiple-choice question is without knowing of the correct choice that *that* is the answer to the question. More generally, these cases might be taken to show that -WH TO -THAT is false:

-WH TO -THAT: In every context c: If 'S knows-wh $Q^?$' is true in c, then there is a contextually salient answer to $Q^?$ in c, P, such that 'S knows that P' is true in c too.[6]

The claim that *History Exam* presents a counterexample to -WH TO -THAT would be interesting in its own right if it were true. But as we will see in a moment, there is reason to think it can't be. Yes, the contrast calls out for an explanation. And unfortunately I lack the space to try and offer one here. But we can be confident in -WH TO -THAT even if we lack an explanation of why it is harder to hear the true readings of the 'knows-that' ascriptions than the 'knows-wh' ascriptions. Here are two reasons why.

---

[5] (2) and (5) are omitted because neither admits of a straightforward 'know-that' translation. I'll also report that (1*) sounds basically fine to me (and to a number of informants).

[6] With the exception of Farkas (2016a,b), I know of no theorist who denies -WH TO -THAT. Indeed Farkas herself claims that she is the only person in the literature who denies it (2016a, p. 3). Note that her putative counterexamples to -WH TO -THAT are not at all like *History Exam*.

First, if -WH TO -THAT were false, then we should expect (8) to have a true uniform interpretation:

(8)   ?? Although Peggy knows what the answer to question 5 is [namely b], she does not know that the answer to question 5 is b.

But instead it just sounds like a contradiction. I do not know how one would try to explain that on the hypothesis that -WH TO -THAT is false.

Second, it is actually not so difficult to hear the true readings of the 'knows-that' ascriptions. Three quick points in illustration:

A) The 'knows-that' analogs of (1)–(6) read naturally when stress is placed on any of 'Peggy', 'question 5', or 'b'.

B) Once we've anchored ourselves into a context where the 'knows-wh' ascriptions sound fine, the 'knows-that' ones start to sound fine too. For example, the first response to (9) seems entirely appropriate, while the second seems bizarre:

(9)   Q: Does Peggy know what the answer to question 5 is?

   A_1: ✓ Yes, she knows that the answer to that question is b.

   A_2: ?? Yes, but she does not know that the answer to that question is b.

C) Finally, the natural answer to 'Which of the students knows that the answer to question 5 is b?' is 'Peggy', not 'Neither of them'. And against thinking that the preference for 'Peggy' is forced because the question *presupposes* that at least one of the students knows that the answer to question 5 is b, notice that the natural answer to questions like 'Which of the students is certain the answer to question 5 is b?' is 'Neither of them', not 'Peggy'.

One who holds the view that the 'knows-that' analogs of (1)–(6) are all false must therefore do more than just explain away the appeal of -WH TO -THAT. She must also explain why it is so easy to get into a frame of mind on which those ascriptions are intuitively true. It is not at all clear what sort of story can be told here if in fact the 'knows-that' ascriptions are context-invariantly false.

2.4   Fourth objection: loose speech

The last objection I will consider (for now) claims that strange ascriptions like (1)–(6) are mere *loose speech*.[7] That is to say: as uttered in context, none of (1)–(6) is literally true. Impressions to the contrary are due to the fact that (i) each can be used to communicate a true proposition other than the one determined by the literal meaning of the sentence, and (ii) our intuitions are primarily tracking these (non-literal) communicated propositions.

---

[7] I use the label 'loose speech' somewhat loosely, and intend that it cover accounts error-theoretic, pragmatic, and whatever else in between.

What is the true but non-literal proposition communicated by, e.g., (6, 'Peggy knows what the answer to question 5 is') in context? Well, that depends on your preferred version of the objection. But here are some popular suggestions: that Peggy knows for the purposes of the exam what the answer to question 5 is; that Peggy will behave as one who knows what the answer to question 5 is; and that Peggy knows more likely than not what the answer to question 5 is.

Now, there are various specific grounds on which one could quibble with these paraphrases. For starters, it is not clear that any neatly accounts for the many possible variations on *History Exam*, such as §2.1's *Restaurant Booking* and *Memory Experiment*. But rather than engage with any particular suggestion, I'm going to make two general points that suggest that no version of this strategy is promising.

First point: anti-skepticism

The first point is *tu quoque* against those who think that ordinary knowledge ascriptions like (10) sometimes express true propositions:

(10)   I know that Trump is the president of the United States.

Here is the point in abstract. Skeptics about knowledge have various arguments that imply that ascriptions like (10) are really just loose speech.[8] If one wants to be anti-skeptical about knowledge, one better have responses to these arguments. But it is hard to see why such responses wouldn't work just as well as responses to the arguments that try and diagnose strange ascriptions like (1)–(6) as loose speech. Thus, anti-skepticism makes the loose speech response to strange ascriptions dialectically precarious.

Here is the worry in more detail. Those who wish to explain away the intuitive judgments about *History Exam* are often eager to point out the ease with which one can hear (1)–(6) as expressing falsehoods. And on that point they are correct: it is just not that difficult to anchor oneself into a context in which the natural readings of these sentences are other than those raised to salience in §1. For instance, if one is asked something like: 'Does Peggy *really* know what the answer to question 5 is, or does she just know what it is conditional on her answers being good?', it will often be perfectly appropriate to answer in the negative. Likewise, speeches like (11) can be as felicitous out of the blue as any of (1)–(6):

(11)   Regardless of whether I *in fact* have the good answers, there's a 50% chance I don't. So there's no way for me to know what the answer to question 5 is.

They also seem to have the further effect of shifting the context so that (1)–(6) can no longer be felicitously asserted.

The issue, however, is that these are exactly the kinds of diagnostics the skeptic can use to dismiss vast swaths of the ordinary usage of 'knows' as loose speech. For instance, if one is asked

---

[8] (Unger, 1975) is the *locus classicus* for this position. See also (Hawthorne, 2004, §§3.1-3.4) for helpful discussion.

something like: 'Does Jones *really* know that she has hands, or does she just know that it looks like she has hands?', it will often seem perfectly appropriate to answer in the negative.[9] Likewise, speeches like (12) can be as felicitous out of the blue as (10):

(12) Regardless of whether Trump is *in fact* still the president of the United States, there's no way for me to know that he hasn't had a fatal heart attack in the past few moments. So there's no way for me to know whether he is the president.

They also seem to have the further effect of shifting the context so that speeches like (10) can no longer be felicitously asserted. So if, being anti-skeptical, we are committed to thinking that (10) expresses a true proposition in at least some contexts, the fact that we can so easily get into a context in which it seems to express a falsehood (and that we tend to get stuck in such contexts) cannot be reason to categorize it as loose speech. But then why think any differently of the strange ascriptions?

A comparison with so-called non-factive uses of 'knows' might be helpful here. There are certain contexts in which utterances of 'S knew/knows that P' can seem fine even when it is common knowledge that P is false. Here are two examples, the first from Hazlett (2010, p. 501):

(13) Everyone knew that stress caused ulcers, before two Australian doctors in the early 80s proved that ulcers are actually caused by bacterial infection.

(14) Stranded under enemy fire, I knew I was going to die there. But then the air support came in and I somehow made it through.

Should we take these sorts of cases to show that our best theory of knowledge should reject the entailment from 'S knows that P' to P?

Probably not. I think they are just instances of loose speech.[10] And the reason I am comfortable dismissing *these* felicitous uses of 'knows' as loose speech, while not (1)–(6), is that the arguments that make the loose speech diagnosis appropriate in this case are: (i) independently motivated and (ii) discriminating enough to target just these uses.

Here is such an argument. It is well-established that sometimes when we wish to describe or predict the behavior of other agents, we use sentences that portray the way the world appears to those agents, regardless of whether those appearances are veridical.[11] To use just one example: if you ask me why my Flat-Earther friends won't let their children visit space.com, I can felicitously answer 'Because it's NASA propaganda', even when it is clear to both us that this isn't so. Knowing that there is a general "perspective-shifting" phenomenon, we should expect there to be contexts in which the facts about what agents know and the facts about what those agents *take* themselves

---

[9] Of course, one can always dig in and insist that Jones does really know that she has hands: say, by pointing out that she can *see* her hands. But the same is true of Peggy in *History Exam*: it is open to one to point out that she has the good answers, and must therefore *really* know the answer to every question.

[10] Though see Hazlett (2010, 2012) for arguments to the contrary.

[11] See, e.g., Schlenker (2004) for background on this phenomenon.

to know are treated as equivalent. We may thus predict the existence of non-factive uses of 'knows' without impugning the semantic validity of the inference from 'S knows that P' to P.[12]

Second point: the form of the judgments

The second problem with the loose speech response does not presuppose anti-skepticism about knowledge. It is that our judgments about strange knowledge ascriptions are both highly delicate *and* highly mechanical, in ways that make non-semantic treatments of the data look implausible.

As we saw in §2.3's discussion of knowledge-wh versus knowledge-that, our intuitive judgments about strange ascriptions depend sensitively on whether the complement to 'knows' is a wh-clause or a that-clause. This sensitivity is difficult to square with the loose speech theorist's account of our intuitions. According to that account, the reason we find assertions of (1)–(6) felicitous is because we associate certain kinds of non-literal content with those speech acts, and judge that *that* content is true. But if a fine sounding assertion of (6, 'Peggy knows *what* the answer to question 5 is') is really just loose speech used to convey a true proposition other than the one that is (6)'s semantic content, then why does an assertion of (6*, 'Peggy knows *that* the answer to question 5 is b') tend to sound so much worse? Is it really that much harder to convey the original message in the language of knowledge-that? It is difficult to imagine what would explain this. More plausible is the idea that something lexical (or broadly semantic) is responsible for the contrast, rather than something like a glitch in our capacities for mind-reading. And as we will see in §§4–5, there are a number of other ways in which our intuitions about strange knowledge ascriptions are exquisitely sensitive to small changes in content and form. So even if the loose speech theorist has a story about what explains the knowledge-wh/knowledge-that contrast, she'll need new ones to account for these other contrasts too.

And still there is the other side of the problem, which is that our judgments about the data are in certain respects too *mechanical* to plausibly be accounted for by a loose speech story. One point brought out by *History Exam* is that our intuitions about knowledge ascriptions seem to be attuned to the linguistic environments in which those ascriptions are embedded. The rough idea (to be made precise later) is this. When a knowledge ascription is embedded in a conditional, quantifier, disjunction, or modal (as in, e.g., (1)–(4) respectively), it is sometimes interpreted as if a contextually salient proposition were already part of the subject's knowledge—regardless of whether (intuitively speaking) the subject has any evidence for that proposition. For example, on the natural readings of a conditional ascription like (1, 'If I have the good answers, then I know what the answer to question 5 is'), the salient proposition is the conditional's antecedent: for the purposes of assessing (1)'s consequent, we treat Peggy as already knowing that her answer sheet is good. This is why (1) seems true, despite the fact that (intuitively speaking) Peggy has no evidence that her answers are good. It is also why (1') seems false:

(1')   ? If Pete has the good answers, then I know what the answer to question 5 is.

[12] And note that perspective-shifting is *not* a plausible diagnosis of our intuitions about *History Exam*. Peggy's perspective is the same as ours in all the respects that matter.

Adding the proposition that *Pete* has the good answers to Peggy's stock of knowledge does nothing to guarantee that she knows what the answer to question 5 is. In other words, 'knows' seems to interact productively with conditionals: our interpretation of 'Peggy knows what the answer to question 5 is' depends not only on whether it is embedded in the consequent of an indicative conditional, but also on the content of the conditional's antecedent.

Setting aside the exact details of this idea (that will come later), we should not expect a loose speech story to do justice to it. This is because on the loose speech story, whether Peggy *in fact* has the good answers makes no difference to whether she knows what the answer to question 5 is. So the inference from, e.g., 'Peggy said *If I have the good answers, then I know what the answer to question 5 is*' to, e.g., 'Peggy really meant *If the proposition that I have the good answers were added to my stock of knowledge, I would know what the answer to question 5 is*' must be derived purely from general reflection on such things as Peggy's psychological state and background facts about rational communication. I for one find it is hard to see how any such derivation could be both extensionally adequate and theoretically well-motivated.

In light of these concerns, I find the prospects of the loose speech diagnosis dim. Our best theory of knowledge needs to have something to say about the uses of 'knows' raised to salience in *History Exam*. This is not to say that we need to have serious doubts about the principles to which they pose *prima facie* counterexamples, e.g., CREDENCE, BELIEF, and JUSTIFICATION; but it is to say that we should get used to thinking that these sorts of strange ascriptions are part of the competent use of the concept. We will return to the question of the epistemological significance of this fact (§6) once we have a theory of 'knows' in place that can account for these uses. And it is to that task we now turn.

3 Knowledge by constraint

3.1 Contextualism versus invariantism

What must a theory of knowledge look like if it is to account for the strange ascriptions of *History Exam*? For starters, it will have to be some version of *contextualism* about knowledge—the view that 'knows' is associated with different epistemic relations across different occasions of use. The alternative is to adopt a kind of *ultra-liberal invariantism* about knowledge, claiming that knowledge is in every context as liberal as Peggy's knowledge of the answer to question 5.[13]

Aside from its intuitive implausibility, such an extreme form of invariantism faces a number of empirical problems. As was established in the previous section, it is relatively easy to get *bad* readings of just about all of (1)–(6). We saw in §2.3 that changing the complement of 'knows' from a wh-clause to the a that-clause tends to have this effect. And in §2.4 we saw that the same can happen when we start asking questions like 'Does Peggy *really* know what the answer to question

---

[13] Here I set aside relativism about knowledge and ignore the difference between moderate and sensitive invariantism. Complications that might arise from considering these sorts of options will not affect the main points of the discussion.

5 is?'; or when we point out that Peggy seems to lack the requisite background knowledge (e.g., that she has the good answers). The malleability in our intuitive judgments are difficult to explain on the assumption that these ascriptions are context-invariantly true.

There are also some more exotic constructions that pose direct challenges to invariantist treatments of the data. Suppose we add the following detail to *History Exam*: that Peggy knows that by the time the exam starts, Roger will have figured out whether it was she or Pete who got the good answer sheet. During the exam she could perfectly well entertain thoughts like:

(15)   If I have the good answers, then I know that Roger is the only one who knows it.

(16)   If I have the good answers, then I know that unbeknownst to me or Pete, Pete is going to fail the exam.

Each reads naturally. But assuming with the invariantist that 'knows' only ever expresses one epistemic relation, each conditional's consequent entails a sentence of the form 'I know that: P and I don't know that P', which in turn entails the straightforwardly contradictory 'I know that P and I don't know that P'. So this cannot plausibly be the natural reading of either sentence. Instead, the natural readings of (15)–(16) must involve non-uniform interpretations of the two occurrences of 'knows': the outer 'know' expressing the kind of strange knowledge brought out by the earlier examples, the inner 'knows' expressing a more ordinary kind.

Thus, the real question is what form of contextualism about knowledge is best suited to explain what is happening with the data under consideration. Might the contextualist theories of knowledge defended by any of (e.g.) Cohen (1986, 1988, 1999); DeRose (1992, 1995); Lewis (1996); Heller (1999); Neta (2003); Schaffer (2005b, 2007); Blome-Tillmann (2009); Ichikawa (2011); Schaffer and Szabo (2013) be up to the task? The answer to this question depends, among other things, on how strictly these views are held to their original letter. If held strictly enough, I'm skeptical that any of these views will get the data right. But I lack the space to give a proper argument for this; or to explore, for each view, whether and how it needs to be revised to give so as to model the data correctly.

So instead the plan is this. This section will outline a bare-boned, contextualist semantics for 'knows' in the style of Lewis's (1996) contextualist relevant alternatives theory. We'll start with a carbon-copy of his view, but then shortly after identify an important technical innovation needed to give an adequate treatment of the embedded strange ascriptions. Importantly, the discussion will be entirely silent on the *meta-semantics* of the proposed semantic theory—i.e. issues about how, when, and why the denotation of 'knows' shifts. The section that follows this one will take up that problem.

3.2 The view

On a *relevant alternatives* conception of knowledge, to know that P is to eliminate all the relevant alternatives to P.[14] We'll regiment this slogan as follows:

R.A. SCHEMA 'S knows that P' is true at world w and context c ($<w, c>$) iff: 'All the relevant worlds are P worlds' is true at $<w, c>$.[15]

Here is the intuitive idea behind R.A. SCHEMA. In any given context, there is a set of relevant worlds against which a claim to knowledge is tested (i.e., the extension of 'All the relevant worlds' in c). For a world to be relevant in the context, it has to pass two tests: first, it has to be compatible with the subject's evidence (more on which in a moment); and second, it has to be among the set of worlds *the context* deems relevant (also more on which in a moment). Only if a world passes both tests is it relevant. And if every world that passes both tests is a P world, then all the relevant worlds are P worlds, and so by lights of that context the subject knows that P.

Now to be a bit more precise. We will say that the extension of 'All the relevant worlds' at $<w, c>$ is determined by two things: (i) S's **evidence** at w, which is a set of worlds (the **boldface** indicates a technical reading of 'evidence'); and (ii) a proposition supplied by c (henceforth a 'restriction'). The relevant worlds at $<w, c>$ are the worlds you get when you take S's **evidence** and intersect it with the restriction. Thus, 'S knows that P' is true at $<w, c>$ iff the intersection of S's **evidence** (as determined by w) and the restriction (as determined by c) entails P.

On the official view **evidence** is an unanalyzed primitive. All that is assumed is that it is a *reflexive* epistemic relation, and thus that an agent's **evidence** at a world w always contains w (to secure the factivity of knowledge). For the purposes of explaining the view, however, it will help to give a working theory of **evidence** in terms of a more familiar epistemic notion. So in an exploratory spirit we will identify one's **evidence** with one's *ordinary knowledge*.[16] I will use '*knows*' (that is, the *asterisks*) to denote ordinary knowledge and related notions (e.g., *epistemic possibility*, *ignorance*, etc.). The basic idea is that *knowledge* is what the typical moderate invariantist about knowledge thinks knowledge is. I will thus assume that facts about *knowledge* are context-invariant. I will also assume that most people *know* whether they have hands, who the current president of the United States is, where they'll be living in a week, and so on. We can stipulate that *knowledge* satisfies CREDENCE, BELIEF, and JUSTIFICATION. Thus—though also stipulatively, if need be—Peggy does not *know* what the answer to question 5 is.

The idea behind contextually generated restriction is easiest to grasp if we allow ourselves some abuse to use and mention. Intuitively, we want it to represent the conjunction of all the propositions

---

[14] Proponents of the relevant alternatives conception of knowledge include Dretske (1970); Unger (1975); Lewis (1996); Schaffer (2005a,b); Rysiew (2006); Blome-Tillmann (2009); Ichikawa (2011); Schaffer and Szabo (2013).

[15] For the sake of concreteness, we will assume that the objects of knowledge are propositions, and that propositions are sets of possible worlds. This is simply a modeling choice. The important features of the account (as designed to handle strange ascriptions) are compatible with a wide variety of views on the nature of content.

[16] This is a departure from Lewis, who uses something along the lines of *Cartesian certainty* in spelling out the **evidence** role. Since we are not (yet) interested in accounting for skeptical puzzles, a weightier notion like *knowledge* is more convenient for our purposes. We will say more about **evidence** in §6.2.

"known for free" in context—i.e., the propositions we get to know without having any **evidence** for. In some contexts the restriction may be something trivial like $\top$—in which case agents will only know what they *know*. In others it may be something weightier, in which case agents will know whatever they *know* conditional on the restriction. But the basic idea is that what changes from context to context is how much work an agent's *knowledge* (**evidence**) is put to in establishing what they know.

However, for reasons that will emerge in §3.3, on the official view the restriction will be an object with the type of a *question*—a partition over possible worlds—from which a proposition is eventually derived. More precisely, we will use the mechanisms of *constraint* to generate our con-textually variable restriction.[17] Constraints are polar questions (i.e., yes/no questions). Constraints combine with worlds to generate propositions, which can then be used to restrict sets of worlds (like a subject's *knowledge*) via set intersection. The proposition a constraint question generates at a world is just the true answer to that question at that world. So, supposing $\Psi^?$ is the constraint $\ulcorner$Is it true that Ψ?$\urcorner$, and that Ψ is true at w but false at w'; then the proposition generated by $\Psi^?$ at w is Ψ, and at w' is $\neg\Psi$. Thus: context's contribution to the calculation of the value of 'knows' will go by way of a question, not a proposition; the agent's *knowledge* will then be restricted by whichever proposition the world says is the true answer to that question.

Combining a *knowledge*-based analysis of **evidence** with a constraint-based account of re-striction gives us the following implementation of R.A. SCHEMA:

MODERATE CONSTRAINT 'S knows that P' is true at $<w, c>$ iff: every world compatible both with S's *knowledge* (at w) and with the true answer to the c-determined constraint $\Psi^?$ (at w) is a P world.

Here is an example to give a sense of how the view works. Suppose Lexie checked the weather forecast this morning and saw that it predicted rain, and that this suffices for Lexie to *know* that *if* the weather forecast is reliable, then it will rain. Suppose also that the weather forecast is in fact reliable. Now consider 'Lexie knows that it is raining' in a context in which the constraint (i.e., $\Psi^?$) is, 'Is the weather forecast reliable?'. According to MODERATE CONSTRAINT, 'Lexie knows that it is raining' is true in this context just in case every world compatible with what she *knows* and with the true answer to 'Is the weather forecast reliable?' is a world in which it is raining. Since the true answer to 'Is the weather forecast reliable?' is that it is reliable, and since every one of the worlds consistent with what Lexie *knows* and with the proposition that the forecast is reliable is a world in which it is raining (remember: Lexie *knows* that if the forecast is reliable, then it is raining), 'Lexie knows that it is raining' is true in that context—just as we want.

More generally: if we want to know whether 'S knows that P' is true in a context, we look at all the worlds compatible with what S *knows*, prune off those that are inconsistent with the true

---

[17] I get the "constraint" label from Dorr and Hawthorne (2013), who use the mechanism to explain various puzzling data involving epistemic modals. The use of a question-like mechanism to generate a proposition (rather than a proposition itself) will be justified shortly. See also Hawthorne (2007) for more on the kind of modal data that is thought to exhibit the constraint phenomenon. (And see also Gillies (2010) and Moss (2015) for further examples of constraint-*like* accounts of the semantics of epistemic vocabulary.)

answer to the contextually determined constraint $\Psi^?$, and see if any of the worlds that remain are ones in which ¬P. 'S knows that P' expresses a true proposition if and only if none of those worlds are ¬P worlds. A more illustrative but less precise way of thinking about it: if 'S knows that P' is true when the contextually determined constraint is $\Psi^?$, then if S were to come to *know* the true answer to $\Psi^?$, S would thereby be in a position to *know* P too.[18]

3.3 Accounting for the data

Now to apply MODERATE CONSTRAINT to *History Exam*. Recall that we are making no assumptions about the meta-semantics of constraint, and so will help ourselves to whatever resolution of the context is needed to get the data right. The aim here is to illustrate certain structural features a theory of knowledge needs to have if it wants to account for the relevant readings. We will return to the question of whether there might be principled ways of deriving these readings in §4.

Recall *History Exam*'s (1)–(6):[19]

(1)   If my answer sheet is good, then I know what the answer to question 5 is.

(2)   One of us knows what the answer to question 5 is.

(3)   Either Pete got the good answers, or I know what the answer to question 5 is.

(4)   I might know what the answer to question 5 is; it depends on whether I got the good answers.

(6)   Peggy knows what the answer to question 5 is.

According to MODERATE CONSTRAINT, 'Peggy knows that the answer to question 5 is b' is true at $<w, c>$ just in case every world compatible with what Peggy *knows* and with the true answer to the c-determined constraint is a world in which the answer to question 5 is b. Let G be the proposition that Peggy's answers are good, and let $G^?$ be the corresponding constraint question (to which G happens to be the true answer). Assuming that the natural readings of (1)–(6) are those on which the c-determined constraint is $G^?$, we get that: 'Peggy knows that the answer to question 5 is b' is true at $<w, c>$ just in case every G-world compatible with what Peggy *knows* is a world in which the answer to question 5 is b. This is indeed the case. Since Peggy *knows* that if her answers are good then the answer to question 5 is b, there are no G-worlds compatible with what Peggy *knows* in which the answer to question 5 isn't b. Thus 'Peggy knows that the answer to question 5 is b' is true at $<w, c>$.[20]

---

[18] MODERATE CONSTRAINT finds semantic analogies between 'knows' and other expressions that are subject to the mechanisms of constraint. If Dorr and Hawthorne (2013)'s claims are right (in broad outline), then epistemic modals like 'must' and 'certainly' are the closest semantic analogs to 'knows'. This makes 'knows' a kind of adverbial quantifier (like 'always' or 'must') rather than a domain quantifier (like 'all' and 'every'—cf. (Ichikawa, 2011)). See Schaffer and Szabo (2013, §§2.4, 3.1, and 4.5 in particular) for discussion of the advantages modeling 'knows' along the lines of a quantificational adverb (as MODERATE CONSTRAINT does), particularly as concerns standard semantic challenges to epistemic contextualism.

[19] I omit (5) because it is redundant.

[20] Note that Pete does *not* *know* that if Peggy's answers are good, then the answer to question 5 is b. This means that 'Pete knows what the answer to question 5 is' is false in any context in which the constraint is $G^?$ (as we want).

It is straightforward to see how the truth of 'Peggy knows that the answer to question 5 is' entails the truth of all of (1)–(6) (given the details of the case). But there remains the question of how to explain why, intuitively, Peggy's has to hedge when she ascribes herself knowledge of the answers (by embedding the ascription under a conditional, modal, etc.), while Roger can felicitously ascribe Peggy that same knowledge unembedded with (6). As it turns out, the answer to this question will also explain why MODERATE CONSTRAINT uses constraint questions to generate its restricting propositions.

But first: some new notation. We will say that 'knows$_{\Psi?}$' denotes whatever relation MODERATE CONSTRAINT predicts 'knows' denotes when the constraint question is $\Psi^?$. Thus: S knows$_{\Psi?}$ that P iff the intersection of S's *knowledge* and the true answer to $\Psi^?$ entails P.

Now for the explanation. MODERATE CONSTRAINT's account of why Roger can assert (6) outright while Peggy has to hedge by embedding it in a conditional, modal, etc., is actually quite simple. It is that only Roger *knows* that Peggy knows$_{G?}$ what the answer to question 5 is. Why does only Roger *know* that Peggy knows$_{G?}$ what the answer to question 5 is? Because only Roger *knows* that the answer to G$^?$ is G. For all Peggy *knows*, the answer to G$^?$ is ¬G. Thus, for all Peggy *knows*, she knows$_{G?}$ that the answer to the question 5 is *not* b. This is why only Roger can felicitously assert (6).

What Peggy does *know*, however, is that she knows$_{G?}$ what the answer to question 5 is *conditional on G being true*. Thus, she *knows* that *if* G, then she knows$_{G?}$ what the answer to question 5 is. Hence why she can assert (1). Likewise, she *knows* that it *might* be that G, and thus *knows* that it might be that she knows$_{G?}$ what the answer to question 5 is. Hence why she can assert (4). And the same points apply *mutatis mutandis* to (2) and (3). Thus, constraint questions play an essential role in accounting the asymmetry in our judgments about what Peggy and Roger may felicitously *assert* in the context of *History Exam*.

As it turns out, they also play an essential role in accounting for the *truth-conditions* of certain kinds of strange ascriptions. Consider another thing Peggy could utter in the context of *History Exam*:

(17)   There's a 50% chance I know what the answer to question 5 is.

(17) has a true reading. It can't be one on which 'knows' denotes *knowledge*, for Peggy *knows* that Peggy doesn't *know* what the answer to question 5 is. So it must take a restricted reading of some sort. However, if the restriction were generated by a proposition—say, G—then (17) would still be false. And that's because there is a 100% chance (not 50%) that all the worlds consistent both with what Peggy *knows* and with the proposition that Peggy has the good answers are worlds in which the answer to question 5 is b. To get it so that there is only a 50% chance that the proposition expressed by 'I know what the answer to question 5 is' is true, it needs to be that Peggy is only 50% certain about whether she stands in the relevant epistemic relation to the proposition that the answer to question 5 is b. And that is exactly what we achieve by using a question like G$^?$ to generate the relevant restriction. Peggy is 50% certain that the answer to G$^?$ is G. She is thus

50% certain that every world compatible with what she *knows* and with the true answer to that question is a world in which the answer to question 5 is b—hence (17).[21]

Thus, MODERATE CONSTRAINT provides a simple and intuitive semantic account of the behavior of 'knows' in *History Exam*'s (1)–(6).[22] Following in the tradition of Lewis (1996), the core idea is that 'knows' expresses a relation between an agent and all the propositions that are entailed by a contextually determined subset of that agent's **evidence**, which in this case is her *knowledge*. The relevant subset is determined by the mechanisms of constraint—the flexibility of which explains why *History Exam*'s Peggy can be ascribed knowledge of propositions for which she possess at best dubious evidence, and the structure of which explains why Roger may ascribe Peggy this knowledge unconditionally even when Peggy cannot.

4 The meta-semantics of constraint

We know that MODERATE CONSTRAINT has the resources to account for the true readings of the strange ascriptions of *History Exam*. However, in giving the account we had to assume that the natural readings of the *History Exam* ascriptions are those in which the constraint is something like 'Does Peggy have the good answers?'. One might wonder what, if anything, explains why these are the natural readings of these ascriptions.

Unfortunately, since I don't know what the explanation is, I will not offer one here. But what I will do is argue that none of the familiar styles of explanation is up to the task, and that given the unruly nature of the data, we probably shouldn't expect our meta-semantics for 'knows' to be much more than a collection of imperfect heuristics. I will then argue that the question of the true semantics for 'knows' needs to be sharply distinguished from the question of its meta-semantics, and that the difficulties in giving a systematic theory of the latter may be symptoms of the concept's unanalyzability

4.1 Rethinking a standard assumption

There is a certain meta-semantic assumption common to most existing forms of contextualism about knowledge. Put roughly, it is the assumption that contextually driven shifts in the interpretation of

---

[21] MODERATE CONSTRAINT also provides an elegant account of (7) ('If Pete doesn't know what the answer to question 5 is, then I do'). Let P stand for the proposition that Pete knows$_{G?}$ what the answer to question 5 is. Peggy doesn't *know* whether P. But she does *know* that if P, then ¬G. (If Pete knows$_{G?}$ what the answer to question 5 is, then he must have the good answers.) So she *knows* that if P, then she knows$_{P?}$ what the answer to question 5 is—hence (7).

[22] For completeness: it captures the natural readings of (15)–(16) as follows:

(15$_G$)   If I have the good answers, then I know$_{G?}$ that Roger is the only one who *knows* it.

(16$_G$)   If I have the good answers, then I know$_{G?}$ that *unbeknownst* to me or Pete, Pete is going to fail the exam.

That is to say: in each case the outer 'know' takes a constrained reading—the one associated with the question 'Does Peggy have the good answers?'—while the inner 'knows' takes its ordinary reading. It is straightforward to see why, given MODERATE CONSTRAINT, both (15) and (16) would come out true on this interpretation.

'knows' are due to shifts in the high-level psychological states of speakers. See, e.g., Cohen (1999):

> ...[S]entences of the form 'S knows P' can, at one time, have different truth-values in different contexts. Now when I say 'contexts', I mean 'contexts of ascription'. So the truth-value of a sentence containing the knowledge predicate can vary depending on things like *the purposes, intentions, expectations, presuppositions, etc.*, of the speakers who utter these sentences. (p. 1; my emphasis)

As well as Rysiew's (2016) survey article:

> ...[T]here is a strong degree of consensus among the theories under discussion[23] that context itself is to be understood in terms of such things as *the interests, purposes, expectations, and so forth*, of the knowledge attributor. (§3.3; my emphasis).

Let *X* denote one's preferred admixture of the factors Cohen and Rysiew allude to.[24] We can then state the core meta-semantic assumption common to most existing forms of contextualism about knowledge as follows:

PSYCHOLOGISM   Changes in the semantic contribution of 'knows' are to be explained in terms of changes in the X facts.

Though PSYCHOLOGISM is extremely vague, it can be shown that any reasonable precisification of it has got to be false. The basic problem is that it cannot account for the various embedded uses of 'knows' in *History Exam*. The reason why is that the natural readings of these constructions involve *intra-sentential* shifts in the denotation of 'knows', shifts that cannot plausibly correspond to anything like a change in the X facts.

Consider for instance (15, 'If I have the good answers, then I know that Roger is the only one who knows it'). On pains of (attitudinal) inconsistency, 'knows' must express two different relations within the scope of (15)'s consequent. But if PSYCHOLOGISM is true, then changes in the denotation of 'knows' can occur only when, for the relevant parties, the X facts change. To get the natural reading of (15), then, we would have to hold that changes in what's relevant, salient, at stake, presupposed etc., occur *mid-consequent*. This is not a desirable result. To name just one implausible consequence: it would seem to suggest that one of the words in the string 'that Roger is the only one who knows it' could marks the occasion whereby the X facts shift from those amenable to a strange reading of 'knows' to a normal one.

A similar point can be made about simpler embedded cases like (1, 'If I have the good answers, then I know what the answer to question 5 is'). Think of things from Peggy's perspective. Either the high-level features of your psychology responsible for settling the interpretation of 'knows' (e.g., your attentions, presuppositions, interests, etc.) are such that 'I know what the answer to

---

[23] Here the theories referenced include those of Heller (1995); Lewis (1996); Cohen (1998); Rieber (1998); Neta (2002); Blome-Tillmann (2009); Greco (2010).

[24] It won't matter for this section's arguments whether the X factors are ascriber, subject, or assessor dependent.

question 5 is' expresses a true proposition in your mouth, or they are such that it expresses a false proposition. If they are such that it expresses a *true* proposition, then what explains your first-order uncertainty about what the answer to question 5 is? *Some* feature of your psychology is causing you to (subjectively) assess it to be no more than .5 likely that the answer to question 5 is b. Shouldn't this be exactly the sort of feature that, given standard forms of contextualism about knowledge, induces a reading of 'knows' on which 'I know what the answer to question 5 is' expresses a falsehood?

Let us assume so: that is to say, let us conclude from the foregoing that the sentence 'I know what the answer to question 5 is' expresses a *false* proposition in your mouth prior to the utterance of (1). Now a new problem arises. (1) is true. So is its antecedent. So the proposition expressed by 'I know what the answer to question 5 is' in consequent position must be other than the one it expresses unembedded. Since 'I know what the answer to question 5 is' expresses a false proposition in your mouth prior to the utterance of the conditional, it follows that the context shifts mid-conditional so as to make it express a true proposition in consequent position. But, as in the case of (15), one wonders what kind of psychological mechanisms could realistically underly such a shift.

We do better not making the assumption that causes these problems. They arise because standard contextualist theories of 'knows' (and their corresponding meta-semantics) were designed to account for the pull of skepticism in the face of lottery puzzles, bank cases, and skeptical paradoxes. These are cases in which the X facts can quite plausibly be taken to explain the shifts in their intuitive judgments about the extension of 'knows'. But these theories were not designed for constructions like (1)–(4) or (15)–(16). What these constructions show is that 'knows' can undergo fine-grained, intra-sentential shifts in denotation—shifts that are not exclusively responsive to changes in the high-level psychological states of those making or assessing claims that contain it. Thus, setting aside the question of whether MODERATE CONSTRAINT (or some variant on it) is true, the fact that there are true readings of sentences like (1) and (15) shows that PSYCHOLOGISM must be false.

This is an interesting result in its own right. But it also makes the task of giving a meta-semantics for constraint look much more difficult. If not facts about attention, stakes, presuppositions, etc., then what *does* explain why the constraint takes the value it does on the natural readings of the strange knowledge ascriptions?

4.2   Against QUDs

Well, one idea worth considering is that the contextually determined constraint is always just the context's "question under discussion" (QUD), in the style of (e.g.) Roberts (2012).[25] Supposing the workings of constraint could be reduced to the workings of QUDs, then so long as the constraint questions posited in §3's explanation of (1)–(6) could plausibly be the relevant QUDs associated with those contexts, MODERATE CONSTRAINT's account of the data would start to look more predic-

---

[25] See, e.g., Ginzburg (1996); von Fintel (2004); Schaffer and Szabo (2013) for further discussion.

tive.[26]

Unfortunately, this proposal fares no better than PSYCHOLOGISM. For the story to work for *History Exam*, it would have to be that the QUD associated with the natural readings of (1)–(6) is 'Does Peggy have the good answers?' (i.e., $G^?$). And that in isolation is fine. But issues arise when we consider slight variants on the original ascriptions. Compare (1) to its "tautologous" analog $(1_\top)$:

(1)   ✓ If I have the good answers, then I know what the answer to question 5 is.

$(1_\top)$? If I have the good answers, then I know whether I have the good answers.

$(1_\top)$ is significantly worse than (1). But if the contextually determined constraint were always just the salient question under discussion, we would not expect to see this contrast. Here is why. Given MODERATE CONSTRAINT and the facts of *History Exam*, $(1_\top)$ is true whenever the contextually determined constraint is $G^?$.[27] This means that the contexts associated with the natural readings of (1) are those in which $(1_\top)$ expresses a true proposition. So the difference in our judgments about (1) and $(1_\top)$ must be due to the differences in the kinds of contexts we associate with their natural readings. And therein lies the problem for the constraints-as-QUDs theory. By hypothesis, the QUD associated with the natural readings of (1) is $G^?$. To account for the fact that $(1_\top)$ sounds false, it would have to be that the QUD associated with the natural readings of $(1_\top)$ is something other than $G^?$. But this is not at all what we should expect: the most straightforward difference between (1) and $(1_\top)$ is that $(1_\top)$ makes salient the question of whether Peggy has the good answers—i.e., $G^?$—while (1), by contrast, merely makes salient the question of whether Peggy knows what the answer to question 5 is. That is to say: if one of (1) and $(1_\top)$ were more likely to be associated with the QUD of 'Does Peggy have the good answers?' it would be $(1_\top)$—exactly the opposite of what is suggested by the data. We thus have some reason to expect a question to be less likely to be used as a constraint when it is simultaneously the salient question under discussion.

In fact, the inverse correlation between a question's working as a constraint and its being a QUD may be what *explains* the general badness of "tautologous" constructions of the form:

(18)   ? If P, then I know whether P.

The rough thought here is that when the QUD is $P^?$, then knowledge ascriptions that are constrained by the true answer to that question are likely to seem trivial. Hence the preference against resolving the context that way.[28]

Without avail of QUDs I am not optimistic about the general strategy of repurposing more familiar contextualist machinery in the pursuit a meta-semantic theory of the workings of constraint.

---

[26] With one minor adjustment: the assumption that constraint questions are always polar questions would have to be relaxed.

[27] The argument is trivial. If Peggy has the good answers and the constraint question is $G^?$, then obviously every *epistemic possibility* compatible with the true answer to that question is one in which Peggy has the good answers.

[28] This is not to say that 'S knows whether P' is *false* (or undefined or what have you) when the constraint is $P^?$. Given MODERATE CONSTRAINT, it has to be true. It's just that it's going to be hard to hear its true readings.

And between that and the failures of PSYCHOLOGISM, I am skeptical any general, non-trivial theory of the meta-semantics of constraint is forthcoming.

4.3   Patchwork meta-semantics

But pessimism about the existence of a general theory aside, there are patterns in the data strong enough to warrant some predictive generalizations. For example: the natural readings of sentences of the form

(19)   If Ψ, then S knows whether P.

(20)   Either $\neg\Psi$, or S knows whether P.

(21)   It might be that S knows whether P; depends on whether Ψ.

tend to be those in which the contextually determined constraint is $\Psi^?$. That is to say, when 'knows' is embedded in certain linguistic environments—conditionals, disjunctions, and certain uses of epistemic modals (those with continuations that make a question salient)—the content of the constraint question tends to be supplied by other parts of the environment: the conditional's antecedent, the negation of the other disjunct, whatever follows the 'depends on whether...' construction, etc. This is exactly what we see with *History Exam* (and §2.1's *Restaurant Booking* and *Memory Experiment*). Moreover, examples may be multiplied with ease:

(22)   If Trump was born in the 1930s, then we know whether he is lying about his age.

(23)   Either the directions my friend gave us are inaccurate, or I know how to get to the concert.

(24)   We might know where to find our missing cat; depends on whether she escaped through the front door or the back window.

Given that the generalization is robust, it is reasonable to appeal to it in account for the data in the way MODERATE CONSTRAINT does. Thus, with two caveats, it is a prediction of MODERATE CONSTRAINT that sentences of the form of (19)–(21) will have true readings whenever S *knows* that P entails Q.

Here are the two caveats. First, the generalization fails for tautologous conditional knowledge ascriptions along the lines of (18) ('If P, then I know whether P'). Second, although the generalization will more or less invariably hold when S, the subject of the ascription, also happens to be the speaker (i.e., for first-person ascriptions), there are counterexamples when the speaker and subject are distinct (i.e., for third-person ascriptions). §5 will discuss this observation in more detail.

For present purposes neither caveat is of concern. None of (1)–(6) is a tautologous conditional knowledge ascription, and all the ascriptions that fit the pattern of (19)–(21) have the speaker coinciding with the subject. Thus, with regards to the question of why the constraint associated with the natural readings of (1)–(6) in the context of *History Exam* is 'Does Peggy have the good answers?', the answer is that most of those sentences fit the pattern of (19)–(21).

But what about those that don't fit the pattern? The generalization does little to explain why the existentially quantified (2, 'One of us knows what the answer to question 5 is') gets the reading it does. And it seems to do even less to explain why the *unembedded* (6) ('Peggy knows what the answer to question 5 is') seems true when uttered by Roger, an onlooker who *knows* that Peggy has the good answers.

What sentences like (2) and (6) show is that sometimes the content of the constraint has to be supplied by features of the extra-linguistic context.[29] There is lots of vague stuff to be said about the circumstances in which we might expect these sorts of ascriptions to sound more natural. For instance: ascriptions like (6) tend to sound better when the ascriber's interest is in how the subject will *behave*—e.g., whether Peggy will write 'b' for question 5—rather than with the quality of the subject's *evidence*—e.g., how good Peggy's reasons are for believing that the answer to question 5 is b. Unfortunately, however, I doubt we can do much better than that. The data just seems to be too unruly to be amenable to mechanical explanations in the style of the one offered for (19)–(21).

So this is where we are on the issue of the meta-semantics of constraint. Setting aside the two caveats from above, we have some general rules for predicting which constraint question will be operative in context *for knowledge ascriptions embedded in conditionals, disjunctions, and modals*. And for the rest of the ascriptions we have vague heuristics and just-so stories that appeal to salient features of the circumstances in which the ascriptions occur. One might take this to be evidence that MODERATE CONSTRAINT is problematically incomplete. But against rushing to this hasty conclusion it is worth stressing three points.

First, *no one* has a tidy story that explains why (e.g.) (2) or (6) should seem true given the facts of *History Exam*. Peggy lacks most all of the features philosophers associate with knowledge. So it is no strike against MODERATE CONSTRAINT in particular that it must appeal to extra-linguistic features of context to account for it. And unlike most other views, MODERATE CONSTRAINT can at least accommodate the data.

Second, consider the thesis, popularized by Williamson (2000), that the concept of knowledge is unanalyzable. Let us suppose it is correct: though there may be various interesting generalizations we can make about the proper use of the concept of knowledge, we cannot know in general and non-circular terms the conditions for its application. From there it follows that we cannot know in general and non-circular terms the conditions under which 'S knows that P' expresses a true proposition.[30] However, if we cannot know in general and non-circular terms the conditions under which 'S knows that P' expresses a true proposition, then it is not reasonable to demand that a contextualist theory of 'knows' come with a general and non-trivial meta-semantics. And that's because the opacity of the meta-semantics of the context-sensitivity of 'knows' could very well be part of what *explains* why we cannot know in general and non-circular terms the conditions under

---

[29] See Dorr and Hawthorne (2013, p. 886) for a similar observation about the meta-semantics of their constraint-based semantics for epistemic modals.

[30] Else we would be in the curious position of knowing in general and non-circular terms the conditions under which 'S knows that P' expresses a true proposition, but not knowing in general and non-circular terms the necessary and sufficient conditions for knowing. What could be missing?

which 'S knows that P' expresses a true proposition. That is to say: if one is accustomed to thinking that we're not going to be able to articulate the rules for distinguishing knowledge from ignorance, one should also be prepared to think that we're not going to be able to articulate the rules for when and why 'knows' undergoes shifts in interpretation.

Third and finally, setting aside whether we ought to think that there are *in principle* limits on the existence of a non-trivial meta-semantics for the context-sensitivity of 'knows', we definitely ought not think that a contextualist semantics for an expression ought to be accepted only if we've found a meta-semantics for it. That is to say, we should not think that we need a theory of the mechanisms that determine when, how, and why context shifts with regards to some expression to know (i) that that expression is context-sensitive and (ii) that it undergoes certain kinds of shifts. Take quantifier domain restriction, for instance. We can know that the domains over which expressions like 'everyone' and 'the sailors' quantify vary from context to context—sometimes the universe, sometimes just the people in some room—even if we don't have a general theory that tells us when these shifts happen and why. To reject a contextualist semantics for quantifiers because it has not been paired with a sufficiently predictive meta-semantics for domain restriction is to confuse importantly distinct theoretical questions. One should not make a similar mistake in assessing the prospects of contextualist accounts of 'knows'.[31]

5   Some over-generation worries

In the absence of a meta-semantic theory that tells us the conditions under which context can yield one constraint rather than another, the proponent of MODERATE CONSTRAINT puts very few necessary conditions on the existence of a true reading of 'S knows that P'. Really the only necessary condition is that P be true. For if P is true, then so long as there is a true proposition Ψ such that S *knows* that if Ψ, then P, then 'S knows that P' will be true whenever the constraint is $\Psi^?$.[32]

This might seem like an alarming prediction. But it can't be *too* far from the truth given the ease with which (19)–(21) take true readings—at least in the first-person (more on which momentarily):

(19)   If Ψ, then S knows whether P.

(20)   Either $\neg\Psi$, or S knows whether P.

(21)   It might be that S knows whether P; depends on whether Ψ.

And more generally, given the unruly nature of the data surveyed in this paper, we should probably learn to live with the fact that no empirically adequate theory of 'knows' is going to be all that predictive.

Still, it's one thing for a theory to predict that 'S knows that P' has both true and false readings (depending on how the context is resolved), yet not be able to tell us when we should expect to

---

[31] Ichikawa (2011) makes a similar point in defending Lewis's (1996) contextualist semantics.

[32] Having $\Psi^?$ just be $P^?$ would achieve this straight away, but as we know from §4.2 there may be a prohibition on resolving the context-sensitivity of 'knows' in this "tautologous" way.

hear those readings; it's another thing to predict that it has both true and false readings when we can't seem to hear one of those readings at all. Theories with the first feature have the drawback of only being able to explain so much; theories with the second look empirically inadequate. This section will discuss some knowledge ascriptions that should have true readings given MODERATE CONSTRAINT, but in fact seem only to have false ones. It will then suggest a revision to MODERATE CONSTRAINT based on observation about the peculiar behavior of 'thinks'.

5.1 Two objections

As first noted in §4.3, it is in general much easier to access true readings of first-personal instances of (19)–(21) than third-personal ones. In fact, when the subject term denotes someone who is not the speaker and who has no special evidence or beliefs about the antecedent and consequent, these sentences seem only to have false readings. For example, suppose Peggy *knows* that Roger doesn't *know* which of the two answer sheets is correct, but does *know* what each of them says. Now compare:

(1)   ✓ If I have the good answers, then *I* know what the answer to question 5 is.

(25)   ? If I have the good answers, then *Roger* knows what the answer to question 5 is.

(25) seems not to have a true reading. Given MODERATE CONSTRAINT, this fact comes as a complete surprise. Roger *knows* that if Peggy's answers are good, then the answer to question 5 is b. He thus knows$_{G^?}$ what the answer to question 5 is. Why, then is the 'knows' of (25) seemingly incapable of being interpreted this way?[33]

The next objection concerns the report 'Peggy knows who has the good answers'. MODERATE CONSTRAINT predicts it is true whenever the contextually determined constraint is 'Is the answer to question 5 b?' (henceforth 5$^?$). This is because Peggy *knows* that if the answer to question 5 is b, then she has the good answers. One might thus expect there to be contexts where 'Peggy knows who has the good answers' reads naturally unembedded, much in the same way there are contexts where (6, 'Peggy knows what the answer to question 5 is') does. But holding the details of *History Exam* fixed, there seem to be no contexts in which Roger can felicitously assert (26):

(26)   ? Peggy knows who has the good answers.

This is especially strange considering that Peggy can get 'knows' to denote knowledge$_{5^?}$ just fine:

(27)   ✓ If the answer to question 5 is b, then I know who has the good answers.

And again, nothing about MODERATE CONSTRAINT or any of the meta-semantic heuristics we have appealed to can explain these facts.

---

[33] I'll note in passing that this contrast is yet further evidence against the wide-scoping response to the data of §2.2. If the explanation of why (1) is fine is that 'knows' is taking wide-scope over the entire conditional, then there is no reason (25) shouldn't be fine too.

5.2 Knowing and thinking

One way of responding to these challenges is to appeal to the black box that is the meta-semantics of constraint. This would be to accept that there *are* true readings of (e.g.) (25) and (26) in the relevant circumstances, but claim that for reasons having to do with how the context-sensitivity of 'knows' is typically resolved, it is very difficult (perhaps even impossible) to hear them.

Though I have no proof that this isn't the best account of the judgments, it seems to me quite unattractive. To repeat an earlier point: it's one thing being able to find true readings of a knowledge ascription without being able to say what makes them favored in certain circumstances rather than others; it's another not being able to find them at all. Being in the latter situation seems to me pretty good reason to believe that something has gone wrong with one's semantic theory of 'knows', rather than with one's meta-semantic theory of the underlying context-sensitivity.

So instead I will pursue a different idea, which is to revise MODERATE CONSTRAINT so that it *predicts* the absence of the relevant readings. What follows is a brief sketch of the revision I find most promising.[34]

First, a conjecture: holding context fixed, whenever 'S knows that P' has a true reading, 'S thinks that P' will too. Note: this is not to say that, holding context fixed, whenever 'S knows that P' is *felicitous*, 'S thinks that P' will be too. If one is in a position to *know* whether 'S thinks that P' is true, then assertions of the form 'It might be that S thinks that P' can and often will be infelicitous (because too weak), even when the corresponding 'It might be that S knows that P' are just fine. The claim is just that if 'S knows that P' is true at $<w,c>$, then so is 'S thinks that P'.

Second, an observation: the conjecture holds of all the data considered in this paper. Keeping in mind that the claim here is about truth rather than felicity, I leave readers to check this themselves. But here are a few examples of some 'thinks' analogs to the knowledge ascriptions we've been investigating:[35]

(28)   ✓ If I have the good answers, then I think that the answer to question 5 is b. [(1)]

(29)   ✓ One of us thinks that the answer to question 5 is b. [(2)]

(30)   ✓ Peggy thinks the answer to question 5 is b. [(6)]

(31)   ✓ If the answer to question 5 is b, then I think that I have the good answers. [(27)]

So: supposing the conjecture is true, we would have reason to add a *thinking* requirement to our theory of knowledge. That is to say, we would have reason to make it a necessary condition on the truth of 'S knows that P' at $<w,c>$ that 'S thinks that P' is true at $<w,c>$ too. This would give us:

THINKING CONSTRAINT 'S knows that P' is true at $<w,c>$ iff:

   (i) 'S thinks that P' is true at $<w,c>$

---

[34] See Holguín (2018) for a full defense of the response.

[35] Note: the claim is not that these 'thinks' ascriptions *always* have true readings. They definitely have false ones too. The claim is just that they *sometimes* have true readings—namely whenever the corresponding knowledge ascriptions do.

(ii) Every world compatible both with S's *knowledge* (at w) and with the true answer to the c-determined constraint $\Psi^?$ (at w) is a P world.

And with THINKING CONSTRAINT in place, we could then leverage judgments about the falsity of a 'thinks' report into an explanation of the falsity of the corresponding 'knows' report.

Applied to the false-seeming (25) and (26) (repeated here)—

(25)   ? If I have the good answers, then Roger knows what the answer to question 5 is.

(26)   ? Peggy knows who has the good answers.

—the idea seems to work. As it turns out, their 'thinks' analogs *also* seem false:

(32)   ? If I have the good answers, then Roger thinks the answer to question 5 is b.

(33)   ? Peggy thinks that she has the good answers.

I thus submit that MODERATE CONSTRAINT is semantically incomplete. It predicts that there can be true readings of 'S knows that P' even when there are no true readings of 'S thinks that P'. THINKING CONSTRAINT remedies that mistake, and in so doing provides a principled answer to the over-generation challenges outlined at the beginning of this section.

5.3   What about 'thinks'?

Now, I imagine that some readers will be less than fully satisfied with this response, even taking into consideration the necessity of its abbreviated presentation. One might complain that all that we've done is taken a problem for a theory of 'knows' and replaced it with a problem for a theory of 'thinks'. Yes, we know why (25) and (26) seem to have no true readings: it's because (32) and (33) seem to haven no true readings. But what explains *that* fact? Although I lack the space to give a proper answer to this question, the following should give a sense of what relation 'thinks' could be tracking.

Our best guess to a question is the answer we think most likely to be correct. *History Exam* is filled with questions for which its characters can have best guesses. The question 'What is the answer to question 5?' admits of four possible answers: a, b, c, and d. Peggy's credences in those answers are $\frac{1}{6}$, $\frac{1}{2}$, $\frac{1}{6}$, and $\frac{1}{6}$ respectively. Thus, her best guess to that question is b. By contrast, supposing Roger doesn't know whose answers are good, Roger's credences are 0, $\frac{1}{2}$, $\frac{1}{2}$, and 0 respectively. And so it's not the case that Roger's best guess is b, for the simple reason that he has just as much reason to guess c. Similarly, the question 'Who has the good answers?' admits of only two possible answers: Peggy and Pete. And Peggy's credences in each are $\frac{1}{2}$. And so it is not the case that her best guess to that question is that she has the good answers—again for the simple reason that she has no reason to guess one way or the other.

So here is a hypothesis: the 'thinks' judgments discussed above are to be explained in terms of the underlying facts about best guesses.[36] It is because Peggy's best guess to 'What is the answer to

---

[36] See, e.g., Dorst (2017) and Holguín (2018) for further discussion of the connection between best guesses and the semantics of 'thinks'.

question 5?' is b that 'Peggy thinks that the answer to question 5 is b' is true, and hence why Roger may truly assert (6, 'Peggy knows what the answer to question 5 is'). But it is because Peggy's best guess to 'Who has the good answers?' is *not* Peggy that 'Peggy thinks that she has the good answers' is false, and hence why Roger may *not* truly assert (26, 'Peggy knows who has the good answers'). Supposing this hypothesis is correct, it follows that a necessary condition on knowing that P is that P is your best guess to a contextually salient question.

6 Epistemology by constraint

6.1 Revisiting the principles

At last we are ready to return to the question with which the paper opened: What are we to make of the principles that §1 claimed to be in tension with the strange knowledge ascriptions of *History Exam*? Since we've gone contextualist about knowledge, what we're now really asking about is the truth of:

'CREDENCE' If 'S knows that P' is true at $<w,c>$, then 'S's credence in P is $> .5$' is true at $<w,c>$.

'BELIEF' If 'S knows that P' is true at $<w,c>$, then 'S believes that P' is true at $<w,c>$.

'JUSTIFICATION' If 'S knows that P' is true at $<w,c>$, then 'S is in a position to justifiably believe that P' is true at $<w,c>$.

We'll start with 'BELIEF' as the dialectic surrounding it is more complex than for either 'CREDENCE' or 'JUSTIFICATION'. If THINKING CONSTRAINT is true, then 'S knows that P' entails 'S thinks that P'. Whether 'BELIEF' is true then just depends on whether 'S thinks that P' entails 'S believes that P'. And I am more than happy to follow the existing literature in assuming that it does.[37] So if THINKING CONSTRAINT is true, then—contra the arguments of §1—'BELIEF' is as well.

What then to make of those arguments? Well, there were two of them: first, that *History Exam* shows us that 'S knows that P' is sometimes compatible with 'S has arbitrarily low credence in P'; and second, that *History Exam* shows us that 'S knows that P' is sometimes compatible with 'S is rationally unwilling to assert or bet that P'. And floating in the background was the thought that 'S believes that P' is incompatible with these things. But this assumption can be rejected. And one who likes THINKING CONSTRAINT should be willing to do so.

What about 'CREDENCE' and 'JUSTIFICATION'? Here I think the decision between MODERATE and THINKING CONSTRAINT is immaterial: given the existence of cases like *History Exam*—cases in which 'S knows that P' expresses a true proposition even when both 'S has $> .001$ credence that P' and

---

[37] See, e.g., Hawthorne et al. (2016), Dorst (2017). To give an example of an argument for the entailment, observe the abominableness of conjunctions like:

(34)　??Andrew thinks it's raining, but it's not that he believes it is.

'S is in a position to justifiably believe that P' express false ones—neither principle can plausibly be true in full generality. This should also be unsurprising given the general form of theories like THINKING CONSTRAINT. These theories are designed to accommodate the fact that for just about any P compatible with what S *knows* (i.e., any true proposition), there will contexts in which 'S knows that P' expresses a true proposition—give or take, perhaps, the facts about what S's best guesses are. But intuitively neither credence nor justification works like this. It simply does not follow from the fact that P is true (and perhaps that P is S's best guess to some salient question) that there is a true reading of 'S has credence > .5 that P' or 'S is in a position to justifiably believe that P'. So absent some decree that ties 'S has credence > . 5 that P' and 'S is in a position to justifiably believe that P' by fiat, we should learn to live with the falsity of both 'CREDENCE' and 'JUSTIFICATION'. And even if we were to go the route of tying 'knows' to 'credence' and 'justifiably believe' by fiat, it would not follow that either 'CREDENCE' or 'JUSTIFICATION' is true *in the ways epistemologists care about*. The fact that one can conjure up a constraint-based semantics for (e.g.) 'credence' and 'is justified in believing' is no reason to think that the kinds of generalizations that motivated epistemologists to posit principles like BELIEF and JUSTIFICATION are in good standing.

6.2 Revising the principles?

However, even if principles like 'CREDENCE' and 'JUSTIFICATION' are false in full generality, there may have restricted interpretations on which they are true. And perhaps it is these interpretations that epistemologists have in mind when they study and defend connections between knowledge, credence, and justification. What are the relevant interpretations? A natural thought is that they the ones that arise when 'knows' goes *un*constrained.[38] That is to say: they are principles about **evidence**:

CREDENCE* If S's **evidence** entails that P, then S's credence in P is > .5.

BELIEF* If S's **evidence** entails that P, then S believes that P.

JUSTIFICATION* If S's **evidence** entails that P, then S is in a position to justifiably believe that P.

To the extent we find this line of thinking plausible, THINKING CONSTRAINT looks epistemologically innocuous. And that is because on this view, one's **evidence** is one's *knowledge*—an epistemic relation that is putatively familiar, and that we know satisfies each of the above principles. A proponent of THINKING CONSTRAINT thus has a story in which we properly use the word 'knows' in all the strange ways witnessed in *History Exam*, but in which those uses tell us little about the worldly relation of concern to epistemologists: *knowledge*.

Granted, there would still be some residual worries. For one thing, there remains the question of why we should find the distinction between constrained and unconstrained uses of 'knows' to be of particular epistemological interest once we've convinced ourselves that the full range of readings

---

[38] Alternatively: they are the interpretations that arise when the constraint question is the trivial $\top$?.

predicted by THINKING CONSTRAINT are part of the proper use of the concept. For another, one might start to wonder about the epistemology of *knowledge* itself. Much of our thinking about *knowledge* is shaped by thought experiments. For the results of those thought experiments to be probative, it better be that the relevant intuitions involve unconstrained interpretations of 'knows'. But given that unconstrained uses can arise even on unembedded uses (as witnessed by (e.g.) (6)), one wonders about how reliably we can distinguish the unconstrained interpretations from the constrained ones.

But there is also a deeper worry lurking, one that emerges once we place THINKING CONSTRAINT in a larger dialectical context. THINKING CONSTRAINT is somewhat atypical as contextualist theories of knowledge go. Allowing ourselves to be sloppy with use and mention for the moment: THINKING CONSTRAINT sometimes makes it very *easy* to know things, whereas most of the familiar brands of contextualism are designed to make it sometimes very *hard* to know things.[39] What does it mean for knowledge to be easier or more difficult in this way? Well, if we assume *knowledge* corresponds to the "normal" level of difficulty, then we can say that a theory of 'knows' makes it easy to know things just in case: according to that theory, there are contexts in which 'S knows that P' is true even when it's not the case that S *knows* that P. Likewise, we can say that a theory of 'knows' makes it hard to know things just in case: according to that theory, there are contexts in which 'S knows that P' is false even when S *knows* that P.

What is important to observe is that given a body of **evidence**, the mechanisms of constraint can only make knowledge easier to get—never harder. And that's because the mechanisms of constraint always *shrink* the domain of worlds over which 'knows' quantifies. So, for instance, if our preferred theory of **evidence** is one that ties it to *knowledge*, then on the resulting implementation of R.A. SCHEMA (namely THINKING CONSTRAINT), there will never be a context in which S *knows* that P but 'S knows that P' is false.

Now given the data considered so far, this is all fine and well: *History Exam*'s knowledge is about as easy as it gets. But the literature on contextualism about knowledge is filled with evidence suggesting that there are contexts in which 'S knows that P' is false even though it is uncontroversial that S *knows* that P. Consider the following skeptical puzzle, for instance:[40]

(35)   I know that I have hands.

(36)   If I know that I have hands, then I know that I am not a handless BIV.

(37)   I do not know that I am not a handless BIV.

(35) and (37) are both seem to have true readings, while (36) follows from a highly plausible principle like CLOSURE (together with innocuous background assumptions):

CLOSURE   If S knows that P and that P entails Q, then S is in a position to know that Q.[41]

---

[39] I do not want to oversell the point here. Some epistemic contextualists explicitly try to accommodate "loose" readings—e.g., Schaffer (2005b, 2007); Schaffer and Szabo (2013). While others—e.g., Lewis (1996) and DeRose (2002)—seem to be open to the possibility of extending their contextualist theories to account for such readings.

[40] See, e.g., Klein (2015) for an introduction to the vast literature on skepticism.

[41] For an introduction to some of the many issues involved with CLOSURE, see, e.g., Hawthorne (2005).

Supposing we want to take both intuitions and CLOSURE seriously, we're going to have to countenance contexts in which 'S knows that P' expresses a false proposition even while 'S *knows* that P' expresses a true one.[42] Consequently, if we want our contextualist theory of knowledge to predict skeptical uses of 'knows', we're going to need the **evidence** to be something other than *knowledge* in at least some contexts.

So now a new question arises: is the **evidence** relation context-sensitive or invariant? Supposing it is context sensitive, 'knows' would be subject to at least two sources of context-sensitivity: context-sensitivity about what is part of one's **evidence**, and context-sensitivity about what the constraint question is. The question of whether CREDENCE* and JUSTIFICATION* are true would then turn on the question of whether every way of resolving the context-sensitivity of **evidence** leaves intact its connections to the other relations.[43]

But supposing instead we give it an invariantist analysis, which is the option I find more attractive, then '**evidence**' will have to context-invariantly denote some extremely weak epistemic relation—something like *factive Cartesian certainty*. By extension, the mechanisms of constraint will have to do all the work in getting 'knows' to denote a relation like *knowledge*. This would then put the defenders of the **evidence**-centric principles CREDENCE* and JUSTIFICATION* in an awkward position. For although Cartesian certainty plausibly does satisfy each of CREDENCE and JUSTIFICATION, the manner in which it does so is less than fully satisfying. Now slogans like 'knowledge requires high credence', 'knowledge requires justification', etc., are true only in virtue of the fact that (e.g.) the relation of *being incapable of rational doubt* requires high credence and justification. I imagine most epistemologists would have hoped for more.

Unfortunately, however, we lack the space to explore these issues any further. As such, we will end our discussion of CREDENCE and JUSTIFICATION on a somewhat equivocal note. What is clear is that strange ascriptions show that our ordinary ways of thinking and talking about knowledge are often not responsive to the facts about credence or justification. And that is because giving a proper account of these uses requires a particular kind of contextualism about knowledge—one on which just about any true proposition can count as known given the right setup. It is thus no surprise that these uses of 'knows' should look so strange from the perspective of conventional epistemology. What remains to be seen is whether there are principled grounds on which the epistemologist can quarantine them.

---

[42] In case it isn't clear why: If (37) expresses a true proposition in *c*, then given CLOSURE it follows that (35) expresses a false proposition in *c*. But there is no *c* such that 'I *know* that I have hands' expresses a false proposition.

[43] Given the relative obscurity of contextualist theories of knowledge that deny principles like CREDENCE, BELIEF, and JUSTIFICATION, it is not implausible that a good contextualist theory of **evidence** would validate them.

References

Blome-Tillmann, M. (2009). Knowledge and presuppositions. *Mind*, 118(470):241–294.

Cohen, S. (1986). Knowledge and context. *Journal of Philosophy*, 83(10):574–583.

Cohen, S. (1988). How to be a fallibilist. *Philosophical Perspectives*, 2(n/a):91–123.

Cohen, S. (1998). Contextualist solutions to epistemological problems: Scepticism, gettier, and the lottery. *Australasian Journal of Philosophy*, 76(2):289–306.

Cohen, S. (1999). Contextualism, skepticism, and the structure of reasons. *Philosophical Perspectives*, 13(s13):57–89.

DeRose, K. (1992). Contextualism and knowledge attributions. *Philosophy and Phenomenological Research*, 52(4):913–929.

DeRose, K. (1995). Solving the skeptical problem. *Philosophical Review*, 104(1):1–52.

DeRose, K. (2002). Assertion, knowledge, and context. *Philosophical Review*, 111(2):167–203.

Dorr, C. and Hawthorne, J. (2013). Embedding epistemic modals. *Mind*, 122(488):867–914.

Dorst, K. (2017). Lockeans maximize expected accuracy. *Mind*, pages 1–37.

Dretske, F. I. (1970). Epistemic operators. *Journal of Philosophy*, 67(24):1007–1023.

Farkas, K. (2016a). Know-wh does not reduce to know-that. *American Philosophical Quarterly*, 53(2):109–122.

Farkas, K. (2016b). Practical know-wh. *Noûs*, 50(2).

Gillies, A. (2010). Iffiness. *Semantics and Pragmatics*, 3(4):1–42.

Ginzburg, J. (1996). Interrogatives: Questions, facts and dialogue. In Lappin, S., editor, *The Handbook of Contemporary Semantic Theory*, pages 385–422. Blackwell Reference.

Greco, J. (2010). *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge University Press.

Hawthorne, J. (2000). Implicit belief and a priori knowledge. *Southern Journal of Philosophy*, 38(S1):191–210.

Hawthorne, J. (2004). *Knowledge and Lotteries*. Oxford University Press.

Hawthorne, J. (2005). The case for closure. In Steup, M. and Sosa, E., editors, *Contemporary Debates in Epistemology*, pages 26–43. Blackwell.

Hawthorne, J. (2007). Eavesdroppers and epistemic modals. *Philosophical Issues*, 17(1):92–101.

Hawthorne, J., Rothschild, D., and Spectre, L. (2016). Belief is weak. *Philosophical Studies*, 173(5):1393–1404.

Hazlett, A. (2010). The myth of factive verbs. *Philosophy and Phenomenological Research*, 80(3):497–522.

Hazlett, A. (2012). Factive presupposition and the truth condition on knowledge. *Acta Analytica*, 27(4):461–478.

Heller, M. (1995). The simple solution to the problem of generality. *Noûs*, 29(4):501–515.

Heller, M. (1999). The proper role for contextualism in an anti-luck epistemology. *Philosophical Perspectives*, 13(s13):115–129.

Holguín, B. (2018). Believing is guessing. Unpublished manuscript.

Ichikawa, J. (2011). Quantifiers and epistemic contextualism. *Philosophical Studies*, 155(3):383–398.

Klein, P. (2015). Skepticism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2015 edition.

Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy*, 74(4):549–567.

Moss, S. (2015). On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics*, 5(8):1–81.

Neta, R. (2002). S knows that p. *Noûs*, 36(4):663–681.

Neta, R. (2003). Contextualism and the problem of the external world. *Philosophy and Phenomenological Research*, 66(1):1–31.

Rieber, S. (1998). Skepticism and contrastive explanation. *Noûs*, 32(2):189–204.

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69.

Rothschild, D. and Spectre, L. (2016). A puzzle about knowing conditionals. *Noûs*.

Rothschild, D. and Spectre, L. (2017). At the threshold of knowledge. *Philosophical Studies*, pages 1–12.

Rysiew, P. (2006). Motivating the relevant alternatives approach. *Canadian Journal of Philosophy*, 36(2):259–279.

Rysiew, P. (2016). Epistemic contextualism. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition.

Schaffer, J. (2005a). Contrastive knowledge. In Gendler, T. S. and Hawthorne, J., editors, *Oxford Studies in Epistemology 1*, page 235. Oxford University Press.

Schaffer, J. (2005b). What shifts? : Thresholds, standards, or alternatives? In Preyer, G. and Peter, G., editors, *Contextualism in Philosophy: Knowledge, Meaning, and Truth*. Oxford University Press.

Schaffer, J. (2007). Knowing the answer. *Philosophy and Phenomenological Research*, 75(2):383–403.

Schaffer, J. and Szabo, Z. G. (2013). Epistemic comparativism: A contextualist semantics for knowledge ascriptions. *Philosophical Studies*, (2):1–53.

Schlenker, P. (2004). Context of thought and context of utterance: A note on free indirect discourse and the historical present. *Mind and Language*, 19(3):279–304.

Sosa, E. (2007). *A Virtue Epistemology*. Oxford University Press.

Unger, P. K. (1975). *Ignorance: A Case for Scepticism*. Number 105. Oxford University Press.

von Fintel, K. (2004). A minimal theory of adverbial quantification. In Kamp, H. and Partee, B., editors, *Context-dependence in the analysis of lingusitic meaning*, pages 137–175. Amsterdam: Elsevier.

Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.