

# Knowledge by Constraint

Ben Holguín

January 2019

## 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 Introduction

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 the wrong exam, and he doesn’t know who it was. In fact, all he knows is that their answer sheets give different answer to 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.

Putting ourselves in Peggy’s shoes, here are some natural sounding things we could think or say:

- (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 or not.

And putting ourselves in Roger (the TA)’s shoes, we could think or say:

- (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 were to discover that Peggy is the student with the good answers, we could even think or say:

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

But now we have a puzzle. 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 (likewise for every other question). 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:<sup>1</sup>

**CREDENCE** If S knows that P, then S’s credence in P is  $> .5$ .

**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 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 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

---

<sup>1</sup> Here we 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’. We will revisit this assumption in §2.3.

than .5 (and 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.<sup>2</sup>

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 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.

In short, the uses of ‘know’ raised to salience by *History Exam* cry out for explanation. If we take our intuitions about (1)–(6) at face value, then we have a case that presents a strong challenge to a number of widely accepted epistemological principles. 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.

The fact that we can think and talk in these ways has gone largely unnoticed in the literature on knowledge.<sup>3</sup> This paper aims to fill that lacuna. It will work under the assumption that our intuitive judgments about sentences like (1)–(6) ought to be taken at face value, with the goal of exploring the semantic, meta-semantic, and epistemological consequences of adopting a theory of knowledge that is capable of accommodating them. Though §2 (and to some degree §4) will argue for this assumption at some length, readers should not expect the question of whether it is ultimately tenable to be settled here. As will be made evident throughout the paper’s discussion, the uses of ‘knows’ of interest are really quite puzzling, and it’s not clear that going pragmatic or error-theoretic about them will make them any less so. But those who remain skeptical that the judgments ought to be taken at face value should treat the paper as a test of assumptions, seeing how many of the standard views on knowledge must bend or break to account for these ways of thinking and talking.

With that in mind, the paper is structured as follows. The next section argues again certain

---

<sup>2</sup> As we will see in §5, this possibility may deserve to be taken seriously.

<sup>3</sup> The closest comparison I know of are some cases of “loose knowledge” discussed by Hawthorne (2000, pp. 202-3; 2004, pp. 68-9). However, Hawthorne’s cases do not involve subjects who are aware of their impoverished epistemic circumstances—as Peggy is in *History Exam*—nor do they involve judgments about knowledge ascriptions embedded under conditionals, quantifiers, disjunctions, or modals—as in (1)–(4) respectively. As will be seen in §§3–4, these features of *History Exam* and related cases are of special theoretical interest.

natural debunking explanations of the data (§2). We then develop a contextualist theory of knowledge that can account for the data (§3), and discuss at length some issues concerning its meta-semantics (§4). After that we will examine some contrasts in the data the view has trouble explaining (§5). An amendment to the view will be proposed to account for them, one that complicates the discussion of the BELIEF condition on knowledge. The paper concludes (§6) with a discussion of what to make of principles like CREDENCE, BELIEF, and JUSTIFICATION in light of the theory.

## 2 Against taking the judgments at face value

Before giving a theory of knowledge that is capable of accommodating *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. The worry 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, when thinking about *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:

**Office Meeting** Don and Betty have both been asked to attend a company meeting by their boss, Roger. However, shortly before the meeting they get a group email from him. It says that one of them was given the wrong room number for the meeting, that he doesn't know who it was, but that if they get to the room and no one is there, they should go to reception to get the right one. (As it turns out, Don was correctly told room 101, Betty incorrectly room 110.)

I submit that it would be natural for Don to say or think things like:

- (1<sub>R</sub>) If I got the right room number, then I know where to go for the meeting.
- (2<sub>R</sub>) One of us [Betty or Don] knows where to go for the meeting.
- (3<sub>R</sub>) Either Betty knows where to go for the meeting, or I do.
- (4<sub>R</sub>) I might know where to go for the meeting; depends on to whom Roger gave the wrong room number.

Since *in fact* Don got the right room number, it follows that he knows where to go for the meeting. And once Roger finds out that he screwed up with Betty, not Don, he can think and say things like:

- (6<sub>R</sub>) Don is the one who knows where to go for the meeting.

Second, if we follow Williamson (2000, ch. 1) in taking factive mental stative operators (FMSOs) like ‘remembers that’, ‘sees that’, etc., to entail their ‘knows that’ counterparts, then cases like the following present yet more instances of strange knowledge:<sup>4</sup>

**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.

I submit it would be natural for Joan to say or think things like:

- (1<sub>M</sub>) If I got the placebo, then I remember what I ate for dinner last night.
- (2<sub>M</sub>) One of us remembers what she ate for dinner last night.
- (3<sub>M</sub>) Either Megan got the placebo, or I remember what I ate for dinner last night.
- (4<sub>M</sub>) 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<sub>M</sub>) Joan remembers what she ate for dinner last night.

---

<sup>4</sup>[Redacted] gives an extended discussion of similar cases involving emotive factive verbs like ‘surprise’, ‘regret’.

Examples can be multiplied with ease (see §4 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 the *conditional* knowledge ascriptions—e.g., (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<sub>w</sub>) 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<sub>w</sub>) plus the fact that Peggy's answers are good that Peggy knows what the answer to question 5 is.

However, there are at least three 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, 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+<sub>w</sub>) seems false:

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

(1+<sub>w</sub>) ? 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.

Third 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<sub>w</sub>) ? I know that: if Pete doesn't know what the answer to question 5 is, then the answer to question 5 is b.

If Pete has the good answers, then the answer to question 5 can't be b. But supposing there is no strange knowledge, then regardless of whether Pete has the good answers, he doesn't know what the answer to question 5 is. So for all Peggy knows, the antecedent of (7<sub>w</sub>) is true—Pete doesn't know what the answer to question 5 is—while its consequent is false—the answer to question 5 *isn't* b—and so (7<sub>w</sub>) should *not* seem like the kind of thing Peggy could think or say. Thus, wide-scoping does not help explain the good readings of (7).

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

As was first observed in footnote 1, 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:<sup>5</sup>

(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 or not.

(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.

<sup>5</sup> (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).

The claim that *History Exam* presents a counterexample to -WH TO -THAT would be interesting in its own right if it were true.<sup>6</sup> 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.

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<sub>1</sub>: ✓ Yes, she knows that the answer to that question is b.

A<sub>2</sub>: ?? 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.

<sup>6</sup> 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*.



## 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*.<sup>7</sup> 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. 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 one thing, it is not clear that any neatly accounts for the many possible variations on *History Exam*, such as §2.1’s *Office Meeting* and *Memory Experiment*. And as mentioned at the end of §1, I’m more interested in exploring the consequences of modeling the judgments semantically than proving that they ought to be so modeled.

But by way of making a substantive case against the loose speech response, it is worth comparing its present application to an application that is less controversial: namely, accounting for so-called non-factive uses of ‘knows’. As is well known, there are certain contexts in which utterances of ‘S knows that P’ can seem basically fine even when it is common knowledge that P is false. Here are two examples, the first from Hazlett (2010, p. 501):

- (10) Everyone knew that stress caused ulcers, before two Australian doctors in the early 80s proved that ulcers are actually caused by bacterial infection.
- (11) 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 the fact that these sentences sound basically fine to show that our best theory of knowledge should reject the entailment from ‘S knows that P’ to P? I follow orthodoxy in saying probably not.<sup>8</sup> I think they are just instances of loose speech. And the reason I am comfortable dismissing *these* felicitous uses of ‘knows’ as loose speech, while not (1)–(6), is that in this case, the loose speech diagnosis can be motivated on independent and reasonably principled grounds.

Here is what I have in mind. 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.<sup>9</sup> To use just one example: if you ask me why my Flat-Earther friends won’t let their children visit [space.com](http://space.com), I can felicitously

<sup>7</sup> I use the label ‘loose speech’ somewhat loosely, and intend that it cover accounts error-theoretic, pragmatic, and whatever else in between.

<sup>8</sup> Though see Hazlett (2010, 2012) for arguments against orthodoxy.

<sup>9</sup> See, e.g., Schlenker (2004) for background on this phenomenon.

answer ‘Because it’s NASA propaganda’, even when it is clear to both us that no such thing is true. It’s just a convenient way of expressing how the world looks from another perspective. And knowing that there is a general “perspective-shifting” phenomenon, we (as theorists) should expect there to be contexts in which the facts about what agents know and the facts about what those agents *take* themselves to know are treated as interchangeable. 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<sup>10</sup>

I know of no comparably general account of what is driving our intuitive judgments about cases like *History Exam*, *Office Meeting*, and *Memory Experiment*. Until one is offered, we have little reason to think that treating the judgments non-semantically has any explanatory advantage over treating them semantically. I will thus continue to assume that our judgments about these cases ought to be taken at face value, rather than be quarantined as a pragmatic oddity or error in conceptual competence. Again, this is not a proof they need to be treated this way. Nor is it yet an argument for the conclusion that we need to have serious doubts about the principles to which these cases pose *prima facie* counterexamples—e.g., CREDENCE, BELIEF, and JUSTIFICATION (more on which in §6). What it is reason to take seriously is the possibility that our judgments about sentences like (1)–(6) are conceptually competent. And with that, we now turn to developing a theory of knowledge that can make sense of them.

### 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.<sup>11</sup>

Aside from its intuitive implausibility, such an extreme form of invariantism faces a number of empirical problems. For starters, 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. The same can happen when we start asking questions like ‘Does Peggy *really* know what the answer to question 5 is?’; or when we make a big deal of the fact that, though Peggy may *in fact* have the good answers, she doesn’t *know* that she does. So even though it is easy to hear (1)–(6) as having good readings, it is also easy to hear them as

<sup>10</sup> 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.

<sup>11</sup> 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.

having bad readings. This level of malleability in our intuitive judgments is 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. For example, in the context of *History Exam*, Peggy could naturally think or say something along the lines of:

- (12) If I have the good answers, then I know that unbeknownst to me or Pete, Pete is going to fail the exam.<sup>12</sup>

But assuming with the invariantist that ‘knows’ only ever expresses one epistemic relation, (12)’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 the sentence. Instead, the natural reading of (12) must involve a non-uniform interpretation 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 semantics

On a *relevant alternatives* conception of knowledge, to know that P is to eliminate all the relevant alternatives to P.<sup>13</sup> Here is a contextualist-friendly regimentation of this slogan:

<sup>12</sup> Note as concerns the discussion in §2.3 that this is a knowledge-that ascription.

<sup>13</sup> 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).

R.A. SCHEMA ‘S knows that P’ is true at world  $w$  and context  $c$  ( $\langle w, c \rangle$ ) iff: ‘All the relevant worlds are P worlds’ is true at  $\langle w, c \rangle$ .<sup>14</sup>

Allowing ourselves some looseness with use and mention, the intuitive idea behind R.A. SCHEMA is this. In any given context there is a set of worlds against which a claim to knowledge is tested—namely the worlds that are relevant in context. For a world to be relevant in context, it has to be compatible with two bodies of information: first, a body of information associated with *the subject* (more on which in a moment); and second, a body of information associated with *the context* (also more on which in a moment). Only if a world is compatible with both bodies of information is it relevant. And if every world that is relevant is a P world, then by lights of that context the subject knows that P.

What is the body of information associated with the subject? Her **evidence**, which we’ll model as a set of worlds (i.e., a proposition). What is the body of information associated with the context? A **restriction**, which we’ll also model as a set of worlds (i.e., a proposition). What a subject knows in context is whatever is entailed by the intersection of her **evidence** and the contextually supplied **restriction**. Being less loose with use and mention: ‘S knows that P’ is true at  $\langle w, c \rangle$  just in case: all the worlds at the intersection of S’s **evidence** according to  $\langle w, c \rangle$  and the **restriction** supplied by  $\langle w, c \rangle$  are P worlds.

Now to say a bit more about how **evidence** and the **restriction** are working. Officially we will leave **evidence** is an unanalyzed primitive, assuming only that it represents a *reflexive* epistemic relation, and thus that an agent’s **evidence** at a world  $w$  always contains  $w$ . This is 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*.<sup>15</sup> 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—*History Exam*’s Peggy does not \*know\* what the answer to question 5 is.

That covers **evidence**. The idea behind the **restriction** is easiest to grasp if we once again allow ourselves some looseness with use and mention. Intuitively, we want it to represent the conjunction of all the propositions “known for free” in context—i.e., the propositions we get to

<sup>14</sup>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.

<sup>15</sup>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, *ordinary knowledge* is more convenient for our purposes. We will say more about this choice in §6.2.

know without having any **evidence** for. In some contexts the restriction may be something trivial like the tautology  $\top$ —in which case agents will only know what is entailed by their **evidence** (i.e., what they \*know\*). In others it may be something weightier, say the proposition  $\Psi$ , in which case agents will know any  $P$  such that their **evidence** entails the material conditional ( $\Psi \supset P$ ). But the basic idea is that what changes from context to context is how much work an agent’s **evidence** (\*knowledge\*) is put to in establishing what they know.

However, for reasons that will emerge shortly, rather than have the value of the **restriction** be supplied *directly* by context, we will have context supply a *question*—i.e., a partition over possible worlds—from which the **restriction** is derived. More precisely, we will use the mechanisms of *constraint* to generate our contextually variable **restriction**.<sup>16</sup> 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 ‘Is it true that  $\Psi$ ?’, and that  $\Psi$  is true at  $w$  but false at  $w'$ ; then the proposition generated by  $\Psi^?$  at  $w$  is  $\Psi$ , 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 **restriction** is whichever proposition the relevant world says is the true answer to the contextually relevant constraint question.

Combining a \*knowledge\*-based analysis of **evidence** with a constraint-based analysis of **restriction**, we arrive following implementation of R.A. SCHEMA:

**BASIC CONSTRAINT** ‘S knows that P’ is true at  $\langle w, c \rangle$  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.<sup>17</sup>

To help see how the view works, consider the following toy example. 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 BASIC 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

<sup>16</sup> I get the “constraint” label from Dorr and Hawthorne (2013), who use the mechanism to explain various puzzling data involving epistemic modals. 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.)

<sup>17</sup> BASIC 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 BASIC CONSTRAINT does), particularly as concerns standard semantic challenges to epistemic contextualism.

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 at  $\langle w, c \rangle$ , we look at all the worlds compatible with what S \*knows\* at  $w$ , prune off those that are inconsistent with the true answer to the  $c$ -determined constraint  $\Psi^?$ , and see if any of the worlds that remain are ones in which  $\neg P$ . ‘S knows that P’ expresses a true proposition if and only if none of those worlds are  $\neg P$  worlds. A more illustrative but less precise way of thinking about it: if ‘S knows that P’ is true when the  $c$ -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.

### 3.3 Accounting for the data

Now to apply BASIC 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.

Here again are *History Exam*’s (1)–(6):<sup>18</sup>

[Peggy:]

- (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 or not.

[Roger:]

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

According to BASIC CONSTRAINT, ‘Peggy knows that the answer to question 5 is b’ is true at  $\langle w, c \rangle$  just in case every world compatible with what Peggy \*knows\* at  $w$  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

---

<sup>18</sup> I omit (5) because it is redundant.

(to which, at  $w$ ,  $G$  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  $\langle w, c \rangle$  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  $\langle w, c \rangle$ .<sup>19</sup>

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 why, intuitively, Peggy’s has to hedge when she ascribes herself knowledge of the answers by embedding the ascription under a conditional, disjunction, 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 BASIC CONSTRAINT goes through constraint questions to generate its **restricting** propositions. But quickly, some new notation. We will say that ‘knows $_{\Psi^?}$ ’ denotes whatever relation BASIC 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. BASIC CONSTRAINT’s account of why Roger can assert the unembedded (6) while Peggy has to hedge by embedding it in a conditional, disjunction modal, etc., is 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  $\neg G$ . 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. Here is another thing Peggy think or say in the context of *History Exam*:

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

<sup>19</sup> Note that the other student, 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).

(13) 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 directly by a proposition—say, G—then (13) 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 **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  $G^?$  is a world in which the answer to question 5 is b—hence (13).

Thus, BASIC CONSTRAINT provides a simple and intuitive semantic account of the behavior of 'knows' in *History Exam's* (1)–(6). 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 \*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 BASIC 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 contextually supplied constraint question is something like 'Does Peggy have the good answers?'. One might wonder whether we can *predict* this feature of the natural readings of the ascriptions, rather than simply reverse-engineer it from our intuitive judgments.

This is a question about the *meta-semantics* of 'knows': namely, the question of what explains when and why context favor some interpretations (i.e. constraint questions) rather than others. With respect to this question, we will argue first that the data considered so far raise significant problems for standard contextualist meta-semantics for 'knows' (§4.1); and second that given various idiosyncrasies in our intuitive judgments, it is hard to see how there could be *any* predictive meta-semantics for 'knows' (§4.2). But against taking this to be reason to dismiss a contextualist treatment of the data along the lines of BASIC CONSTRAINT, we will then argue that the difficulties involved here are problems for everyone, and in fact may be a symptom of conceptual primitiveness of knowledge (§4.3).



## 4.1 Against standard contextualist meta-semantics

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 ‘knows’ are due to shifts in the high-level psychological states of those making the relevant knowledge ascriptions. 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<sup>20</sup> 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.<sup>21</sup> 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’ across contexts are to be explained in terms of changes in the  $X$  facts.

Though PSYCHOLOGISM is vague, it can be shown that any reasonable precisification of it does poorly with the various embedded uses of ‘knows’ raised to salience by *History Exam*.

One reason to find PSYCHOLOGISM problematic is that our judgments about what Peggy can say or think in *History Exam* seem to be exactly the opposite of what we would expect, at least taking the  $X$  facts at face value. Take (4) (‘I might know what the answer to question 5 is; it depends on whether I got the good answers or not’), and consider, in turn, Peggy’s expectations, presuppositions, and interests. Peggy is clearly taking seriously the possibility that she has bad answers—in fact, she is literally announcing its possibility. So it’s not like she is expecting that she will answer the question correctly. For similar reasons it is equally clear that Peggy is not presupposing that she has the good answers, or anything else that would entail the answer to question 5 is b. And that she answers the test questions correctly is a matter of enormous practical significance to her—in fact, we could make it a life or death matter without disrupting the relevant intuitions. So ‘knows’ simply should not be amenable to a reading on which ‘Peggy

<sup>20</sup> Here the theories referenced include those of [Heller \(1995\)](#); [Lewis \(1996\)](#); [Cohen \(1998\)](#); [Rieber \(1998\)](#); [Neta \(2002\)](#); [Blome-Tillmann \(2009\)](#); [Greco \(2010\)](#).

<sup>21</sup> It won’t matter for this section’s arguments whether the  $X$  factors are ascriber, subject, or assessor dependent.

knows what the answer to question 5 is b' expresses a truth in Peggy's mouth, whether embedded under a modal or not.<sup>22</sup> Yet our judgments about (4) suggest that it does.

A more systematic reason to find PSYCHOLOGISM problematic is that the natural readings of many of *History Exam's* ascriptions involve *intra-sentential* shifts in the denotation of 'knows', shifts that cannot plausibly correspond to anything like a change in the X facts. Consider (12, 'If I have the good answers, then I know that unbeknownst to me or Pete, Pete is going to fail the exam') for instance. On pains of (attitudinal) inconsistency, 'knows' must express two different relations within the scope of (12)'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 (12), 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 unbeknownst to me or Pete, Pete is going to fail the exam' could mark 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., what you're attending to, or what your presuppositions, interests, etc., are) are such that 'I know what the answer to 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 (12), 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

<sup>22</sup> It is not like Peggy is unsure whether she is expecting or presupposing that she has the good answers, or whether it's a matter of significant practical interest to her that she passes the exam.

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 (12). 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 BASIC CONSTRAINT is true, the fact that there are true readings of sentences like (1) and (12) shows that PSYCHOLOGISM must be false.<sup>23</sup>

This is an interesting result in its own right. But it also makes the task of pairing BASIC CONSTRAINT with a predictive meta-semantics that 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 ascriptions in cases like *History Exam*?

## 4.2 Making the data even stranger

I am not sure. In fact, I think it is unclear whether we should expect a satisfying answer to this question. This is not just because the judgments about cases like *History Exam* are hostile to PSYCHOLOGISM; more fundamentally, it is because certain patterns in the data seem resistant to being modeled by *any* kind of predictive theory.

One case in point concerns the contrasts (discussed in §2.3) between knowledge-that and knowledge-wh ascriptions. Given the details of *History Exam*, we can be confident that any

<sup>23</sup> Similar worries beset the “question-under-discussion”-theoretic meta-semantics defended by Schaffer and Szabo (2013). (See, e.g., Ginzburg (1996); von Stechow (2004); Roberts (2012) for background on the theoretical significance of QUDs.) Schaffer and Szabo argue for contextualism about ‘knows’ as well their QUD-based meta-semantics on the basis of cases like the following:

Suppose I know Peggy’s house contains a safe full of diamonds and a safe full of rubies. One day I witness Pete break in and leave with the safe full of diamonds. Sometime after, I see Peggy return home to discover that her diamonds are missing. Ask me ‘Does Peggy know *what* Pete stole?’, and I’ll probably reply along the lines of: ‘Yes, Peggy knows Pete stole her diamonds’. But ask ‘Does Peggy know *who* stole her diamonds?’, and I’ll probably reply along the lines of: ‘No, Peggy doesn’t know that Pete stole her diamonds’. (Emphasis on ‘diamonds’ in the first answer and ‘Pete’ in the second helps bring out the relevant readings.)

Schaffer and Szabo observe that the same knowledge ascription (‘Peggy knows that Pete stole her diamonds’) seems to have both a true and false reading—hence a contextualist semantics for ‘knows’. Moreover, which reading it takes seems to depend on the *presuppositions* of the question under discussion: when the QUD presupposes that Pete stole something (and is aimed at figuring out what he stole) the ascription ‘Peggy knows that Pete stole her diamonds’ seems true; when instead the question presupposes that the diamonds were stolen (and is aimed at figuring out who stole them), the ascription seems false—hence the QUD-based contextualist meta-semantics.

I take no stand on whether Schaffer and Szabo’s treatment of these sorts of cases is correct. All I want to observe is that a similar treatment of many of the data points considered so far in this paper seems implausible. To account for (12), the QUD would have to shift mid-consequent. To account for (1), Peggy would have to start the sentence with one QUD, shift midway through to another, and then presumably shift back to the original by the end (so as to explain why she can’t assert or think ‘I know what the answer to question 5 is’ outright). And how the story would even begin to account for sentences that embed the ascription under a quantifier or a modal (e.g. (2) and (4)) is far from clear.

context in which ‘Peggy knows what the answer to question 5 is’ has a true reading is one in which ‘Peggy knows that the answer to question 5 is b’ has a true reading, and vice-versa. This is why (8) and (14) seem like contradictions:

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

But for some reason, the contexts most naturally evoked by knowledge-that ascriptions are ones on which (1)–(6) are false, while the ones most naturally associated with the knowledge-wh are ones on which (1)–(6) are true. Why should there be this difference in the ease with which the two kinds of knowledge ascriptions take constrained readings given the seemingly obvious (contextual) entailments between them?

Another case in point. A natural meta-semantic hypothesis to pair with BASIC CONSTRAINT is that for conditional knowledge ascriptions like (1) (‘If my answer sheet is good, then I know what the answer to question 5 is’), the contextually supplied constraint is just the antecedent asked as a polar question. This would at least explain all the strange conditional ascriptions considered so far. But it doesn’t take much work to show that that can’t be right—at least not in full generality—for consider (1)’s “tautologous” analog:

- (1<sub>⊥</sub>) ? If my answer sheet is good, then I know whether my answer sheet is good.

It is straightforward to see why (1<sub>⊥</sub>) is true in any context in which the constraint question is ‘Does Peggy have the good answers?’. But to my ears this ascription is noticeably degraded. This seems to be part of a general pattern. Consider, e.g.:

- (15) ✓ If Trump has had a fatal heart attack in the past few moments, then I know Pence is now the president of the United States.
- (15<sub>⊥</sub>) ? If Trump has had a fatal heart attack in the past few moments, then I know Trump has had a fatal heart attack in the past few moments.

So perhaps the meta-semantic generalization ought instead be something along the lines of: for conditional ascriptions, the contextually supplied constraint is the antecedent asked as a polar question—*so long as the constraint doesn’t “trivialize” the embedded knowledge ascription.*

Indeed, there is evidence that natural language imposes this kind of meta-semantic constraint on the interpretation of adverbial quantifiers more generally:<sup>24</sup>

- (16) ✓ If I play poker, I always wear my lucky hat.
- (16<sub>⊥</sub>) ? If I play poker, I always play poker.

<sup>24</sup> Cf. Schaffer and Szabo (2013, §4.3).

(17) ✓ If Paul is in Paris, it is certain he is in France.

(17<sub>T</sub>) ? If Paul is in Paris, it is certain he is in Paris.

So that is something. But how exactly are we supposed to understand the relevant notion of ‘triviality’? Taking (17) as our cue, consider:

(18) a. ✓ If Paul is in Paris, then I know he’s in France.

b. ✓ If Paul is in Paris and Paris is not Berlin, then I know he’s not in Berlin.

These both seem fine, even despite the fact that it is about as trivial as it gets that (e.g.) if Paul is in Paris and Paris is not Berlin, he’s not in Berlin. So evidently there is something *particularly* trivial about conditional ascriptions of the form ‘If P, then I know that P’. But it is not clear how to predict this fact given standard theoretical tools.

And still there is the most basic worry for the prospects of a predictive meta-semantics: how are we explain the fact that in the vast majority of contexts of interest to epistemologists, principles like CREDENCE, BELIEF, and JUSTIFICATION seem to express truths, yet when we embed knowledge ascriptions under conditionals, quantifiers, disjunctions, and modals (as in *History Exam’s* (1)–(4)), counterexamples seem to arise with ease?

To add another example to the pile, many theorists have taken Gettier (1963) to have shown decisively that knowledge is incompatible with the truth of one’s belief being a matter of a certain kind of luck. But cases like *History Exam* provide *prima facie* evidence that in at least some contexts, Gettier’s generalization fails. Moreover, we can make the same point using one of the standard Gettier cases:

**Standard Gettier** Smith’s boss Roger tells Smith that Jones is getting a promotion. Smith knows Jones has ten coins in his pocket, so trusting Roger’s testimony he infers that the man who will get the promotion has ten coins in his pocket. But unbeknownst to Smith, Roger was confused: it is Smith, not Jones, who is getting the promotion. Also unknown to Smith is that he himself has ten coins in his pocket. So Smith’s belief that the man who will get the promotion has ten coins in his pocket is true after all.

It is widely thought that in the contexts most naturally evoked by *Standard Gettier*, (19) expresses a falsehood:

(19) Smith knows that the man who will get the promotion has ten coins in his pocket.

But putting ourselves in Smith’s shoes, we can just as well say and think things like:

(20) If Roger confused me with Jones and I happen to have ten coins in my pocket, then I know whether the man who will get the promotion has ten coins in his pocket.

(21) I might know how many coins are in the pocket of the man who will get the promotion; depends on whether I have ten coins in my pocket and Roger confused me with Jones.

Given that Roger confused Smith with Jones and that Smith happens to have ten coins in his pocket, it looks to follow that (19) has true readings after all. How one is supposed to explain the strength of the Gettier intuition given the ease with which constructions like (20)–(21) may be conjured up is beyond my present understanding.

### 4.3 What now?

In light of these considerations, are we in a position to conclude that there can be no non-trivial meta-semantics for ‘knows’? Of course not. There has been basically no work on these issues, so it is anyone’s guess what future theorizing will bring to bear. But *as it stands* we do not have a theory that will tell us in general terms when ‘knows’ will take one interpretation rather than another. And this naturally raises the question of whether we should start to have worries about a contextualist semantics for ‘knows’ along the lines of BASIC CONSTRAINT.

Well, one thing to emphasize is that we definitely ought not think that a contextualist semantics for an expression ought to be accepted only if we’ve found a predictive 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 on this boat—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’.<sup>25</sup>

For the sake of argument, however, let us suppose that we have convinced ourselves that we will never find a predictive meta-semantics with which to pair BASIC CONSTRAINT. *Now* do we have reason to reject the theory? It is hard to see why we would, unless we took the difficulty of finding a predictive meta-semantics to be evidence that the data ought not be modeled semantically at all. But what would justify *that* inference? The patterns in our intuitive judgments about knowledge ascriptions do not suddenly become less mysterious just because we decide that some of them must be in error or the byproduct of a yet to be specified pragmatic mechanism. For there remain the questions of what explains when and why these errors arise; or when the pragmatic mechanism gets put in use, what its rules are, and how it interacts more generally with general features of communication and rationality. And it’s not like we have good reason to believe we will do any better with these questions than with the ones about the contextualist’s meta-semantics. The fact that the knowledge ascriptions surveyed in this paper

---

<sup>25</sup> Ichikawa (2011) makes a similar point in defending Lewis’s (1996) contextualist semantics from various complaints about its predictiveness.

seem resistant to rational explanation is a problem for everyone, not just those who take the judgments at face value. Until we've convinced ourselves that we can give a better explanation of what's happening by not treating them as semantically respectable, we should be prepared to follow our intuitions where they lead. If that means readjusting our expectations with respect to our ability to predict when a knowledge ascription will seem true, then so be it.

Finally, in relation to this point, consider the thesis popularized by [Williamson \(2000\)](#) that knowledge is *conceptually primitive*. 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. Given this assumption, we should also think we cannot know in general and non-circular terms the conditions under which 'S knows that P' expresses a true proposition. 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. 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 clear why we should expect to be able to give a non-trivial meta-semantics for 'knows'. After all, the opacity of its meta-semantics could be what *explains* why we cannot know in general and non-circular terms the conditions under which 'S knows that P' expresses a true proposition. Put another way: those convinced of conceptual primitiveness about knowledge are already committed to the view that our intuitions about knowledge form something of a black box. Unless they have some special argument that it has to be the semantics *and only the semantics* that grounds this fact, they should not be resistant to the possibility that the meta-semantics is at least partially responsible too.

## 5 From BASIC CONSTRAINT to THINKING CONSTRAINT

In the absence of a meta-semantic theory that tells us the conditions under which context supplies one constraint question rather than another, the proponent of BASIC CONSTRAINT places very few necessary conditions on the existence of a true reading of 'S knows that P'. In fact, the only necessary condition is that P be true. For if P is true, then so long as there is a true proposition  $\Psi$  such that  $S$  \*knows\* that  $\Psi \supset P$ , then 'S knows that P' will be true whenever the constraint is  $\Psi$ . And despite having just argued that we should be prepared to have very little to say about when 'knows' takes one reading rather than another, I believe BASIC CONSTRAINT ought to be amended so as to become more predictive. This is because there is a certain class of knowledge ascriptions that seem to have no true readings at all, yet ought to have true readings given the combination of BASIC CONSTRAINT and our non-existent meta-semantics. And there is a simple generalization that can be incorporated into the view so as to avoid this problem—one that raises some interesting questions about the relationship between the concepts of knowing,

thinking, and believing.

## 5.1 Two contrasts

Recall the basic details of *History Exam*. Roger sells answer sheets to Peggy and Pete. One of them gets bad answers, but neither can figure out who it is in time for the test. All they know is that their answer sheets differ on every question. Eventually Roger discovers it was Pete who got the bad answers, and upon doing so thinks to himself:

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

We saw in §3.3 why, given the constraint question  $G^?$  ('Does Peggy have the good answers?') BASIC CONSTRAINT can account for the true reading of (6): (Roger \*knows\* that) Peggy \*knows\* that if she has the good answers, then the answer to question 5 is b; Peggy has the good answers; so all the worlds compatible with what Peggy \*knows\* and with the true answer to  $G^?$  are worlds in which the answer to question 5 is b; so she knows $_{G^?}$  what the answer to question 5 is.

So far so good. But now observe that Roger cannot seem to felicitously think or say (22) (at least not unembedded):

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

We have been given no reason to expect this. According to BASIC CONSTRAINT, whenever the constraint question is something like 'Is the answer to question 5 b?' (hereafter  $5^?$ ), (22) should have a true reading. And this is because (Roger \*knows\* that) Peggy \*knows\* that if the answer to question 5 is b, then she has the good answers. And since the answer to question 5 is in fact b, it follows that all the worlds compatible with what Peggy \*knows\* and the true answer to  $5^?$  are worlds in which Peggy has the good answers, and thus that Peggy knows $_{5^?}$  who has the good answers. So either there is some mysterious feature of the meta-semantics of 'knows' that explains why  $G^?$  but not  $5^?$  is a natural constraint question in the contexts associated with *History Exam*, or the contrast between (6) and (22) is evidence that BASIC CONSTRAINT is in need of amendment.

That's the first unexplained contrast. To see the second, let's add the following detail to *History Exam*: the reason Roger doesn't know who got the good answers is not because he forgot to whom he gave which set, it's because he himself doesn't know which set of answers is the good one. That is to say: Roger knows he gave Peggy answer sheet 1 and Pete answer sheet 2, and he knows what each answer sheet says, but he has no idea which is the good one. And let us also stipulate that Peggy knows all this.

On this version of the case, Peggy can continue to felicitously think or say things like:

(1) ✓ If I have the good answers, then 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 or not.



But observe that Peggy cannot seem to felicitously think or say things like:

- (23) ? If I have the good answers, then Roger knows what the answer to question 5 is.  
(24) ? Roger might know what the answer to question 5 is; it depends on whether I got the good answers or not.

This is surprising. Given BASIC CONSTRAINT, there should be no contrast at all between (1)/(4) (on the one hand) and (23)/(24) (on the other). Roger \*knows\* that if Peggy's answers are good, then the answer to question 5 is b. He thus knows<sub>G?</sub> what the answer to question 5 is. Thus, the contexts in which natural readings of (1)/(4) arise ought to be exactly those in which the natural readings of (23)/(24) arise. But the data seem to suggest otherwise. So again: either there is some mysterious feature of the meta-semantics of 'knows' that explains why G<sup>2</sup> is a natural constraint question to associate with (1)/(4) but not (23)/(24), or the contrast between the Peggy-centric and Roger-centric ascriptions is evidence that BASIC CONSTRAINT is in need of amendment.

## 5.2 Knowing and thinking

I suggest we go with the latter option: amend BASIC CONSTRAINT to make it capable of predicting our two contrasts. Out of concerns of space, however, I will not be able to do full justice to the amendment I find most promising. So what follows will have to be taken as a suggestive sketch of the idea, rather than a fully worked out view.<sup>26</sup>

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' is just fine. The conjecture is just that if 'S knows that P' is *true* at  $\langle w, c \rangle$ , 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:

- (25) ✓ If I have the good answers, then I think that the answer to question 5 is b. [(1)]  
(26) ✓ One of us thinks that the answer to question 5 is b. [(2)]  
(27) ✓ Peggy thinks the answer to question 5 is b. [(6)]  
(28) ✓ If Paul is in Paris, then I think he's in France. [(18a)]

<sup>26</sup> See [redacted] for an extended discussion of this line of response.

So: supposing the conjecture relating ‘knows’ and ‘thinks’ 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  $\langle w, c \rangle$  that ‘S thinks that P’ be true at  $\langle w, c \rangle$  too. This would give us:

THINKING CONSTRAINT ‘S knows that P’ is true at  $\langle w, c \rangle$  iff:

- (i) ‘S thinks that P’ is true at  $\langle w, c \rangle$ .
- (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 (22), (23), and (24) (repeated here)—

- (22) ? Peggy knows who has the good answers.
- (23) ? If I have the good answers, then Roger knows what the answer to question 5 is.
- (24) ? Roger might know what the answer to question 5 is; it depends on whether I got the good answers or not.

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

- (29) ? Peggy thinks she has the good answers.
- (30) ? If I have the good answers, then Roger thinks the answer to question 5 is b.
- (31) ? Roger might think the answer to question 5 is b; it depends on whether I got the good answers or not.

And so we seem to have a predictive generalization connecting our judgments about ‘knows’ to our judgments about ‘thinks’. I thus submit that BASIC 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 explanation of the unavailability of the good readings of (22)–(24).

### 5.3 What about ‘thinks’?

The move from BASIC to THINKING CONSTRAINT raises two salient questions. First, if ‘S knows that P’ entails ‘S thinks that P’, then what is going on with §1’s arguments that our judgments about *History Exam* present *prima facie* counterexamples to principles like BELIEF? Second and relatedly, what can we say about the propositional attitude that is the denotation of ‘thinks’ on the relevant readings?

We'll answer the first question in the next section. With respect to the second, what we're basically asking for is an explanation of why, e.g., there seems to be a true reading of 'Peggy thinks that the answer to question 5 is b' yet no true reading of 'Peggy thinks she has the good answers'. And although it's not clear we *need* to be able to answer this question to be confident in the judgments or of their use in explaining the two contrasts outlined at the beginning of this section, it would certainly be better if we could. So on that I offer the following.

One's best guess to a question is the answer one thinks is most likely to be correct. *History Exam* is filled with questions to 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, the question 'Who has the good answers?' admits of only two possible answers: Peggy and Pete. Peggy's credences in each are  $\frac{1}{2}$ . So it is not the case that her best guess to that question is that she has the good answers, for she has just as much reason to guess that Pete does. Likewise, consider the version of *History Exam* where Roger knows that Peggy's answer sheet says "5:b" and Pete's "5:c", but doesn't know which is correct. Here Roger's credences in the four possible answers to 'What is the answer to question 5?' are 0,  $\frac{1}{2}$ ,  $\frac{1}{2}$ , and 0 respectively. It is thus not the case that Roger's best guess to 'What is the answer to question 5?' is b, for he has just as much reason to guess c.

So here is a hypothesis: our judgments about 'thinks' are tracking the underlying facts about best guesses.<sup>27</sup> It is because Peggy's best guess to 'What is the answer to 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'). And it is because it is *not* the case that Peggy's best guess to 'Who has the good answers?' is Peggy that 'Peggy thinks that she has the good answers' is false, and hence why Roger may *not* truly assert (22) ('Peggy knows who has the good answers'). Similarly, it is because it is *not* the case that Roger's best guess to 'What is the answer to question 5 is b?' that 'Roger thinks the answer to question 5 is b' is false, and hence why Peggy may *not* truly assert (23) or (24).

I lack the space to further explore the hypothesis that our intuitive judgments about 'thinks' track facts about best guesses. But I submit that connecting 'knows' to 'thinks' and giving a tentative analysis of the latter in terms of best guesses is a simple, empirically motivated way of helping make sense of some data for which our current understanding is quite limited.

---

<sup>27</sup> See Dorst (2019) as well as [redacted] for paper length defenses of the view that there are intimate connections between the facts about thinking and the facts about best guesses.

## 6 Epistemology by constraint

### 6.1 Revisiting the principles

Having explored the semantic and meta-semantic issues raised by the strange knowledge ascriptions of cases like *History Exam*, we are now ready to return to the epistemological principles to which §1 claimed they posed a threat. Since we've gone contextualist about knowledge, what we're now really asking about is the truth of:

**'CREDESCENCE'** If 'S knows that P' is true at  $\langle w, c \rangle$ , then 'S's credence in P is  $> .5$ ' is true at  $\langle w, c \rangle$ .

**'BELIEF'** If 'S knows that P' is true at  $\langle w, c \rangle$ , then 'S believes that P' is true at  $\langle w, c \rangle$ .

**'JUSTIFICATION'** If 'S knows that P' is true at  $\langle w, c \rangle$ , then 'S is in a position to justifiably believe that P' is true at  $\langle w, c \rangle$ .

We'll start with 'BELIEF' as the dialectic surrounding it is more complex than for either 'CREDESCENCE' 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.<sup>28</sup> 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.<sup>29</sup>

What about 'CREDESCENCE' 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 '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 THINKING CONSTRAINT. This theory is 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' seems to express 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

<sup>28</sup> See, e.g., Hawthorne et al. (2016), Dorst (2019), [redacted]. To give an example of an argument for the entailment, observe the abominableness of conjunctions like:

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

<sup>29</sup> Cf. Dorst (2019), [redacted].

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 be narrower 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, belief, and justification. What are the relevant interpretations? A natural thought is that they are the ones that arise when 'knows' goes *unconstrained*.<sup>30</sup> 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 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

<sup>30</sup> Alternatively: they are the interpretations that arise when the constraint question is the trivial T<sup>?</sup>.

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.<sup>31</sup> 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 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 loose 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:

(33) I know that I have hands.

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

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

(33) and (35) are both seem to have true readings, while (34) 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.

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 though S \*knows\*

<sup>31</sup> 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.

that P.<sup>32</sup> 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.

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.

---

<sup>32</sup> In case it isn’t clear why: If (35) expresses a true proposition in *c*, then given CLOSURE it follows that (33) expresses a false proposition in *c*. But there is no *c* such that ‘I \*know\* that I have hands’ expresses a false proposition.

## 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. (2019). Lockean 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).
- Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23(6):121–123.
- 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. (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.
- Ichikawa, J. (2011). Quantifiers and epistemic contextualism. *Philosophical Studies*, 155(3):383–398.
- 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.
- 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.
- 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 linguistic meaning*, pages 137–175. Amsterdam: Elsevier.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford University Press.