Thinking, Guessing, and Believing

Ben Holguín

January 2020

Abstract

This paper defends the view, put roughly, that to think that p is to guess that p is the answer to the question at hand, and that to think that p rationally is for one's guess to that question to be (in a certain sense) non-arbitrary. Some theses that will be argued for along the way include: that thinking is question-sensitive and, correspondingly, that 'thinks' is context-sensitive; that it can be rational to think that p while having arbitrarily low credence that p; that, nonetheless, rational thinking is closed under entailment; that one can rationally think that p, gain further evidence that p, and as a result become rationally compelled to cease thinking that p; that (rational) thinking does not supervene on (rational) credence; and that in many cases what one thinks on certain matters is, in a very literal sense, a choice. Finally, since it is highly plausible that thinking just is believing, it is highly plausible that each of these claims is true of belief as well.

1 Introduction

There are many questions about belief on which there is little consensus. Is belief closed under entailment? What is the relationship between belief and credence? How about belief and knowledge? Can one believe a proposition without being disposed to assert it or use it as a premise in deliberation? Is what one is permitted to believe sensitive to pragmatic or moral factors?

But if there's one thing that *has* been a matter of consensus—or near enough anyway[1]—it's that it is rational to believe a proposition p only if one's evidence for p is a good amount better than one's evidence for ¬p. The question of *how* much better is debated. But the details aside, the general thought is widely regarded as a simple platitude.

That is until recently. Hawthorne et al. (2016) have argued compellingly that the evidential requirements on belief are much weaker than epistemological orthodoxy has taken them to be.[2] Not only can one believe that p while knowing that one isn't in a position to know that p (or to assert p, or to use p as a premise in deliberation), one can believe that p even while knowing that one's evidence makes it more likely than not that p is false.

[1] Kaplan (1995) is a notable exception. James (1956) too, though I'm not sure he's part of the intended reference class.

[2] See Dorst (2019) and Rothschild (2019) for further developments.

We'll see the details of some of their arguments shortly. But the train of thought underlying many of them is this: (i) belief is what is denoted by the ordinary expression 'believes'; (ii) 'thinks' and 'believes' are synonymous; (iii) the attitude denoted by 'thinks' has very weak evidential requirements; so (iv), belief has very weak evidential requirements.

This paper takes no stance on the soundness of this argument. In particular, it takes no stance on the truth of either (i) or (ii). This is because the aim of the paper is not to defend a theory of belief *per se*, but to defend a theory of *thinking*—i.e., the attitude denoted by 'thinks' in sentences of the form 'S thinks that p'.

Of course if thinking *is* believing—if one thinks that p iff one believes that p—then this paper is also defending a theory of belief. But it will be better if we give our theory of thinking from a position of neutrality on these matters. This is not so much because the arguments for (i) or (ii) are particularly complex or controversial. Whether (i) is true or false seems to be mostly just a matter of stipulation. And though (ii)'s truth or falsity clearly isn't a matter of stipulation, the case for it is very strong.

Here is the short version.[3] If thinking weren't believing, then you'd expect to be able to imagine circumstances in which, for some agent S and proposition p, it would be natural to think or say that S thinks that p without believing it (or vice-versa). But evidently we cannot imagine such circumstances:

(1) a. ✗ I think it's raining, but I wouldn't say I believe it is.
  b. ✗ I'm not sure whether Jane thinks Federer will win Wimbledon, but I know she doesn't believe he will.
  c. ✗ My friends think I'm a good a person, but my mom believes I am.

It is difficult (if not impossible) to recover a coherent interpretation of these sentences. But replace 'believe' with (e.g.) 'is sure' or 'is certain' and they sound perfectly fine. If thinking is believing then it is entirely unsurprising why this is so. But if thinking isn't believing—if one can think that p without believing that p (or vice-versa)—then the existence of these judgments is a mystery. There is thus a strong case to be made that a theory of thinking *just is* a theory of the attitude that is the denotation of 'believe', and that a theory of the attitude that is the denotation of 'believe' *just is* a theory of belief (at least on the intended interpretation of 'belief').

But even still: it can be a fraught matter what those who fashion themselves as theorists of belief take themselves to be theorizing about. Some are clearly happy to help themselves to a stipulative notion and leave the ordinary one aside.[4] But many others seem to be motivated by

[3] Here my presentation follows Rothschild's (2019).

[4] See, e.g. Greco (2015, p. 180), who in defense of the view that "believing" that p requires having credence 1 that p writes:

> If the claim that belief involves maximal confidence is to be worth taking seriously at all, we cannot be working with a conception of belief closely tied to natural language constructions involving 'belief' and 'believe'. Much work in epistemology suggests an alternative conception of belief, more closely

ordinary language judgments—a practice that is difficult to make sense of if there isn't a presumption that the object of study is the one denoted by 'belief'. And although the arguments for the view that thinking is believing seem to me quite strong, it's not as if they are unchallengeable.[5] It is for reasons such as these that I think it better if we approached our central question—What does it take to think that p?—without prejudging whether in doing so we are also giving a theory of belief.

Until the paper's concluding section, then, we will be concerned exclusively with thinking, with the aim of answering the following two questions:

The metaphysical question:

What is it to think that p?

The normative question:

Under what conditions is it rationally permissible to think that p?

For conspicuously absent from either Hawthorne et al. (2016) or Rothschild's (2019) discussion of thinking is anything resembling a full account of its metaphysics or norms. Dorst (2019), to his credit, defends a view here (though he is coy about which of the two questions he takes himself to be answering, or whether he's answering both at once). But we will soon see many reasons to think views like his cannot be correct. So regardless of one's stance on the connections between thinking and believing, or between the attitude that is the denotation of 'believe' and the object of conventional epistemological study, we have a lacuna that deserves to be filled.

And on that matter, this paper defends the view, put roughly, that to think that p is to guess that p is the answer to the question at hand, and that to think that p rationally is for one's guess to be in a certain sense non-arbitrary. Some theses that will be argued for along the way include: that thinking is question-sensitive and, correspondingly, that 'thinks' is context-sensitive; that one can rationally think that p without being in a position to use p as a premise in theoretical or practical reasoning; that it can be rational to think that p while having arbitrarily low credence that p; that, nonetheless, rational thinking is closed under entailment; that one can rationally think that p, gain further evidence that p, and as a result become rationally compelled to cease thinking that p; that (rational) thinking does not supervene on (rational) credence; and that in many cases what one thinks on certain matters is, in a very literal sense, a choice. Finally, if (as it seems) thinking *just is* believing, then all this goes for belief as well.

related to knowledge. . . When belief is understood along some versions of these lines, the simple view that (strong) belief involves credence 1 is once again a live option.

[5] See Williamson (2018).

2 A point about felicity judgments

The main aim of this paper is to give a theory of the attitude that is the denotation of 'thinks', as well as that attitude's norms. We will not be making any prior assumptions about the theoretical roles of thinking, and so the primary source of evidence for our theory will be judgments about natural language sentences involving 'thinks'—judgments that for the most part will simply be taken at face value.

One might wonder how intuitions about natural language sentences containing 'thinks' could reveal facts about the *rational* requirements on thinking. But the idea is entirely familiar (even if not under this guise). There is something obviously strange about an utterance of a sentence like:

(2) ✗ I think it's raining and that it isn't.

This is what the '✗' is meant to indicate. We could try to explain the strangeness by giving a semantics for 'thinks' (i.e., a theory of thinking *simpliciter*) that makes 'S thinks that p' entail 'It is not the case that S thinks that ¬p'. But that is not what the evidence warrants. If prior to asserting (2) I had declared that I tend to have irrational patterns of beliefs about the weather, there would be no impression that (2) expresses something that couldn't be true. Indeed, it would be reasonable to cite the truth of (2) as evidence that I was obviously irrational. So (2)'s badness isn't due to the fact it expresses a proposition that cannot be true. Rather, its badness is due to the fact that we take there to be something *irrational* about thinking that it's raining and that it isn't raining. It's a judgment that reveals something about the norms on thinking, rather than its metaphysics.

By the same token, the fact that a sentence of the form 'S thinks that p' is felicitous in certain circumstances will often be pretty good evidence that S is *rational* in thinking that p in those circumstances. For example, if I know that Jones has an extremely low chance of winning the upcoming lottery and you ask me how I think he'll fare, it would be perfectly appropriate for me to respond:

(3) ✓ I think he'll lose.

What does this tell us about thinking and its norms? Well, one thing it tells us is that it is metaphysically possible to think that p even when one's evidence for p is merely statistical. But another thing it tells us is that it is metaphysically possible to *rationally* think that p even when one's evidence for p is merely statistical. I don't *seem* to be representing myself as irrational in using (3) to answer your question about what I think. And we know that we often do detect when a person making a 'thinks'-report would have to be irrational for the report to be true. So absent strong reasons to think our judgments about sentences like (3) are generally confused, we should think that rational agents can speak from knowledge in asserting it.

More generally, we're going to take the fact that it would be fine for S to utter 'I think that p' in circumstances C to be strong evidence that it is possible for S to rationally think that p in C, and we're going to take the fact that it would seem *not* fine for S to utter 'I think that p' in C to be strong evidence that it is *not* possible for S to rationally think that p in C. *Mutatis mutandis* for our judgments about the felicity of reports of the form 'I don't think that p', 'I neither think that p nor that ¬p', 'I don't have a view on p', etc: if S can felicitously utter these sentences in C, then it must be possible for S to rationally fail to think that p in C (otherwise it must not be).[6]

With that, we now turn to answering our two central questions: What does it take to think that p? And what does it take to think that p rationally? As it turns out, getting clear on the first question is much easier once we've gotten clear on the second. So for the next few sections our concern will be with the theory of rational thinking, rather than thinking *simpliciter*.

3 Thinking is weak

We'll start with a thesis defended in detail by Hawthorne et al. (2016) (as well as Dorst 2019 and Rothschild 2019). It is that thinking is *weak*. Somewhat more precisely:

Weakness

The evidential requirements on thinking that p are weaker than the evidential requirements on asserting that p, using p as a premise in deliberation, or being sure that p.

Whether the evidential requirements on assertion, deliberation, and surety are equivalent is not a question we need to settle here. All that the proponent of WEAKNESS is committed to is that the strength of evidence needed to rationally think a proposition is less than the strength of evidence needed to do any of these other things—i.e., that there are circumstances in which it is appropriate for an agent to think that p, despite the fact that it would be inappropriate for her to be sure of it, assert it, or use it as a premise in deliberation.

Since the arguments for WEAKNESS have already been given a fair bit of attention in the recent literature, our treatment of them here will be brief.

One style of argument focuses on simple judgments about 'thinks'-reports. For starters, it is perfectly felicitous to assert that one thinks p but is unsure whether p:

(4) ✓ I think it will rain, but I'm not sure it will.

(Just imagine, for instance, checking the morning weather forecast and seeing a 70% chance of rain.) It is also perfectly felicitous to assert that one thinks but does not know whether p:

[6] Some theorists (e.g., Stanley 2008, Williamson 2018) have suggested that sentences of the form 'I think/believe that p' have uses on which they don't report the fact that speaker thinks/believes that p. Instead, they serve some other function: say to express (or otherwise make salient) the fact that speaker's evidence for the proposition that p is relatively weak. I find the line of reasoning here opaque. It may well be that first-personal doxastic reports have these sorts of expressive and/or parenthetical functions. But as far as I can tell the best explanation of *why* they have that function is that thinking is a "hedged" (i.e. weak) attitude. Besides, the inference from 'S asserted 'I think that p" to 'If S is neither lying nor self-deluded, then S thinks that p' is as valid-seeming as natural language inferences can be.

(5) ✓ I think it will rain, but I don't know that it will.

And it is also possible to felicitously assert that one thinks that p but knows that there is a substantial chance that ¬p:

(6) ✓ I think it will rain, but I know there's a substantial chance it won't.

Since there are circumstances in which conjunctions like (4)–(6) can be uttered felicitously, it must be possible for a rational agent to speak truly in asserting them. Since it is possible for a rational agent to speak truly in asserting them, it must be possible for a rational agent to think it will rain while knowing there is a substantial chance that it won't. But of course it is not possible for a rational agent in these circumstances to be sure that it will rain, assert it will rain, or use the proposition that it will rain as a premise in deliberation:

(4*) ✗ It will rain, but I'm not sure it will.

(5*) ✗ It will rain, but I don't know that it will.

(6*) ✗ I'm sure it will rain, but I know there's a substantial chance it won't.

Hence WEAKNESS.

We can make a similar argument for WEAKNESS by thinking about cases rather than sentences. Suppose A has purchased one of the 100 tickets to an upcoming lottery. She has no special information about what the outcome will be. Does A have enough information to be *sure* that she'll lose the lottery? Of course not. Can she assert that she'll lose, or take for granted that she'll lose in (e.g.) deciding whether to sell the lottery ticket for a penny? No—this would be a failure on A's part to give proper weight to her evidence. But does A have enough information to *think* that she'll lose? Certainly. There is nothing irrational about thinking you're not going to win the lottery.

Indeed, examples that illustrate the weakness of thinking are legion. We think things about the weather, upcoming elections, unsolved murders, mathematical conjectures, and so on—even when we know full well that the evidence for our opinions on these is matters is far from decisive. But we do not act as if we are sure of these things. We do not assert them outright, and we do not treat them as the kind of propositions whose possible falsity can be ignored when engaging in practical deliberation. The evidential requirements on thinking are weak.

4 Thinking is extremely weak

Alright—but how weak? The answer this section defends is: *extremely* weak. One can rationally think that p despite the fact that one's evidence makes the probability that p arbitrarily close to zero.

It will be helpful going forward if we allow ourselves some quick technical stipulations. We will model an agent S's evidence (hereafter E_S) as a proposition: in particular, the proposition that is the conjunction of all the propositions S rationally permitted to be sure of. We will associate with each such body of evidence a *rational credence function* (hereafter $C_S(\cdot)$), which takes propositions to real numbers on the unit interval. We will assume (i) that rational credence functions obey the axioms of the probability calculus; and (ii) that the rational credence function associated with a certain agent's evidence takes p to 1 iff that agent's evidence entails p. In saying 'S has rational credence x that p' (i.e., $C_S(p) = x$), we will mean that the rational credence function associated with S's evidence takes p to x, rather than that S *in fact* has credence x that p and is rational for doing so. As a reasonable gloss: S's rational credences are the credences S would have were S rational.[7]

With this terminology in place, we can give a precise characterization of the claim that thinking is extremely weak:[8]

EXTREME WEAKNESS

There is no positive number x such that: necessarily, for any agent S and proposition p, if $C_S(p) \leq x$, then S is not rationally permitted to think that p.

The argument for EXTREME WEAKNESS is fairly simple. But to warm up to it we'll first argue for the following more moderate principle:

SUBSTANTIAL WEAKNESS

Possibly: for some agent S and proposition p such that $C_S(p) < .5$: S is rationally permitted to think that p.

In words: it possible to rationally think p while having rational credence less than .5 that p.

Here is the argument for SUBSTANTIAL WEAKNESS. Suppose I tell you that an upcoming lottery has 100 tickets, that A has purchased 48 of them, and that the remaining 52 have been distributed evenly among 52 other people (B, C, D...etc.). So A has a 48% chance to win, everyone else 1%. Question: Who do you think will win?

[7] Supposing the standards for rational surety are as lax as ordinary language suggests they are ('I'm sure it rained yesterday', 'I'm not sure whether I'll win the lottery, but I'm sure that if I do I won't squander it all on gambling', etc.), there will be many ordinary, contingent propositions in which agents get to have rational credence 1. This means our notion of rational credence may differ in many important respects from the more traditional notions, which tend to make it very hard to have rational credence 1 in contingent propositions. I see nothing non-terminological hanging on this.

[8] Strands of the discussion of this principle are present in (Hawthorne et al., 2016, pp. 1400–1401) and are developed more fully in (Dorst, 2019, pp. 17–18). (Dorst attributes the arguments for the extreme weakness of thinking to Hawthorne et al., who in turn attribute them to Jeremy Goodman.) Note, however, that Hawthorne et al. appear unwilling to defend a principle as strong as EXTREME WEAKNESS. They accept that rationally thinking that p needn't entail having rational credence greater than or equal to .5 that p (i.e., SUBSTANTIAL WEAKNESS—to be introduced in a moment)—but do not seem to be comfortable thinking one could rationally think that p while having (e.g.) rational credence .001 that p. Dorst, on the other hand, embraces this possibility (on grounds more or less identical to those that are about to be presented). See also Windschitl and Wells (1998) and Yalcin (2010) for discussion of analogous principle concerning the semantics of expressions like 'probable' and 'likely'.

Here is a perfectly reasonable answer: A. After all, we know she is *48 times* more likely to win than anyone else. But we also know her chances of winning are a mere .48. So rational thinking does not require rational credence greater than .48.

Similar judgments can be elicited in other domains. Consider questions like:

- Which horse do you think will finish in first?
- Who do you think will get the Democratic nomination?
- How many people do you think were responsible for the "Jack the Ripper" murders?

You might only have credence .35 that horse A will win, or that Biden will get the nomination, or that the Jack the Ripper murders were performed by no more than one person. But this is no principled barrier to offering these up as your answers. The question is about what you *think*, not what you know (or are sure of). And as long as (e.g.) you're more confident that A will win than that any of the other horses will, being 35% certain that A will win seems to suffice for thinking it. Thinking isn't just weak, it's substantially weak.

Of course, an argument for SUBSTANTIAL WEAKNESS is not yet an argument for EXTREME WEAKNESS. But since the argument for EXTREME WEAKNESS is really just a generalization of the argument for SUBSTANTIAL WEAKNESS, it is worth pausing to consider how this intermediate conclusion might be rejected.

As far as I can tell, the most plausible way of resisting the case for SUBSTANTIAL WEANKESS is to deny that our natural answers to questions of the form 'Wh- F do you think Gs?' really report the things we think. I see two ways of pushing this line.

First, one might (correctly) point out that questions of the form 'Wh- F do you think Gs?' are standardly taken to *presuppose* that there is some F that you think Gs. So, for example, in asking 'Who do you think will win the lottery?', I presuppose—possibly incorrectly—that there is some person you think will win. So perhaps in the circumstances of the 48 ticket case, your answering 'A' is merely your best attempt to accommodate my question's presupposition, rather than a genuine attempt to report what you actually think.

Unfortunately this response has some serious problems. For one, the fact that it is so easy to accommodate the presuppositions of the question 'Who do you think will win the lottery?' is itself good evidence that there is no barrier to thinking propositions in which one's rational credence is a mere .48. Why? Because supposing that being in a position to rationally think that p is compatible with having rational credence .48 that p, it is no surprise that the question's presuppositions are easy to accommodate. But supposing these things are *not* compatible, it is much less obvious why we would go about accommodating the question's presuppositions. Notice that the appropriate answer to questions like 'Who is it that you are sure will win the lottery?', 'Who do you have greater than .5 confidence will win the lottery?', etc., is 'No one', not 'A'. This is good evidence that, contrary to the imagined response, we do not blithely represent ourselves as having irrational doxastic states in order to try to accommodate the presuppositions of the questions we have been asked.

A different issue with this response is that one doesn't even have to be asked a question to report oneself as thinking that A will win the lottery. In the circumstances of the 48 ticket lottery scenario it is perfectly acceptable to assert outright: 'I think A will win', or 'I think the winner will be A', etc. Placing stress on 'A' makes the true readings of these sentences crystal clear. Likewise, if you were to overhear someone else say any of these things (knowing they have the same evidence as you), the natural conclusion to draw would *not* be that that person is irrational; rather, the natural conclusion to draw would be that that person is doing the perfectly normal thing of expressing the thought that the overwhelming favorite to win will, well, win. The fact that such reports are so readily elicited by questions is beside the point.

The second way of denying the probative force of the question/answer data focuses on the apparent *optionality* in how one may choose to respond to such questions. More concretely, though it can be appropriate to answer the lottery question with 'I think A will win', it seems it can *also* be appropriate to answer agnostically: say with an 'I don't know', 'I'm not sure', 'There isn't anyone I think will win', or what have you.

Why would the availability of this alternative response cast doubt on the probative force of the data? I myself find it less than perfectly clear. But presumably it would have to involve the view that, for any given rational credence function and proposition p, rationality permits exactly one of the following attitudes: thinking that p, thinking that ¬p, or agnosticism toward p (i.e., neither thinking that p nor thinking that ¬p). It would then follow that at least one of the two seemingly appropriate kinds of answers ('I think A will win' vs. 'I don't have a view') fails to track the underlying facts about rational thinking. And since the details of the case are such that the opinionated 'I think A will win' is more surprising than the agnostic 'There isn't anyone I think will win'—at least from the perspective of conventional epistemology—one might take this all to be reason to regard it with suspicion.

We will return to the issue of optionality at some length in §9; for now our treatment of it will be brief. The view that the laws of rationality associate any given rational credence function with no more than one coarse-grained doxastic attitude is controversial and contestable. Though its theoretical appeal may count as *some* evidence that the natural language judgments ought to be regarded with suspicion, it would be dubious methodological practice to dismiss them outright on such grounds.

It is worth noting, for example, that the exact same phenomenon seems to arise in more mundane lottery cases. Suppose we modify the details of the lottery so that A has 99 of the 100 tickets, rather than 48. Now consider again the question 'Who do you think will win?'. Though I expect many will be inclined to answer 'A', it is not at all clear that it would be irrational to answer agnostically instead: responses like 'I don't know' or 'I have no particular view on who will win' seem just fine. So there is optionality here too. But in this version of the case it seems especially implausible that the presence of optionality implies that we don't report what we think when we say that we think the person who has a 99% chance of winning will win.

In light of these considerations, we ought to take our intuitions about the answers to questions like 'Wh- F do you think is G?' seriously. And since there is nothing intuitively problematic about reporting that one thinks that the person with 48 of the 100 tickets to a lottery will win that lottery, we have good reason to believe that SUBSTANTIAL WEAKNESS is true.

The argument for EXTREME WEAKNESS—the principle that rational thinking is compatible with arbitrarily low rational credence—is a simple generalization of the argument for SUBSTANTIAL WEAKNESS. For any real number $0 < x \leq 1$, we just need to come up with a lottery in which: (i) some person A has the best chance of winning and (ii) that chance is x. For then we will have found a scenario in which it can be rational to think that A will win the lottery despite the fact that one's rational credence in that proposition is no higher than x.[9]

For $x = .01$, just imagine a 1,000 ticket lottery in which A has 10 tickets and the remaining 990 are distributed evenly among 990 other entrants. If asked 'Who do you think will win the lottery?' the answer 'A' remains appropriate. Again, this is not to say that it is mandatory. An 'I don't know' or 'There isn't anyone I think will win' is fine too. But one who answers 'I think A will' needn't be irrational.

For those who are skeptical, try to imagine how you would press such a person on their answer. What mistake have they made? They're not claiming they know A will win. Nor are they even claiming A's odds of winning are particularly good. They're just saying what they think. And I find it hard to see how one could be irrational in thinking that the winner of the lottery will be the person with the best chances of winning.[10]

The point generalizes. So long as you know A holds a plurality of the tickets, it will be fine to answer 'Who do you think will win?' with 'A', or to assert unprompted 'I think the winner will be A'. This is regardless of whether A's ticket count is 2 rather than 10, or if the total count is 10,000 rather than 1,000. Granted, if you shrink the gap between A's chances of winning and the next highest person's, or if you lower A's absolute chances of winning—or both—then the agnostic response to the question 'Who do you think will win?' may start to seem more

[9] Cf. Williamson (2018, p. 12):

> With questions of the form 'Which N do you believe/think will VP?', it is not clear that there is in principle any lower limit to how likely one may take one's answer to be, provided that one takes it to be more likely than the alternatives on offer.

[10] Here I depart from Hawthorne et al. (2016, p. 1400) (as well as Yalcin (2010), whom they cite), who write:

> In a lottery in which one ticket has a 2% chance of winning and every other ticket a 1% chance of winning, it would seem odd to believe that the 2% ticket will win or to think it probably will win.

I agree that, as a matter of empirical fact, many would deny thinking that the person with a 2% chance of winning will win. But that's not what matters as regards EXTREME WEAKNESS. What matters is whether one who *does* happen to report themselves as thinking it will inevitably seem irrational for doing so. And for the reasons discussed in the main text, this seems not to be the case.

compelling. But again: *making the agnostic response more compelling is not the same thing as making the opinionated response unacceptable*. And I contend that no matter how bad or close the odds happen to be, it will remain permissible to answer the question 'Who do you think will win the lottery?' with 'The person who is most likely to'. So it is possible to rationally think that p even when one's rational credence that p is arbitrarily close to zero.

5 Thinking is non-monotonic

The kinds of cases in which it seems appropriate for S to think that p despite having very low credence that p seem inevitably to be those in which S's evidence makes p more likely to be true than any other of a class of relevant alternatives. This insight will eventually form the basis of our positive theory of rational thinking. But for now I want to put it to a different purpose, which is to argue that the relationship between rational thinking and rational credence is *non-monotonic*:

NON-MONOTONICITY

There can be two agents (or the same agent at different times) S_1 and S_2 such that: $C_{S_1}(p) < C_{S_2}(p)$, yet S_1 is rationally permitted to think that p while S_2 is not.

In other words: increasing the strength of the evidence will sometimes decrease the strength of the conviction. One can start off rationally thinking that p, acquire evidence that increases the likelihood that p, but—because the evidence increases the likelihood of a relevant alternative q even more—become rationally prohibited from thinking that p.

The case for NON-MONOTONICITY looks to be at least as strong as the case for EXTREME WEAKNESS. Consider a horse race with three entrants: A, B, and C. Suppose S_1 has been told by an expert horse bettor that the horses' chances of winning are 40%, 35%, and 25% respectively, while S_2 has been told by a different expert horse bettor that the horses' chances of winning are 45%, 50%, and 5%. What is the rational thing for each of S_1 and S_2 to think about the outcome of the race?

Intuitively, S_1 should think A will win (or perhaps have no opinion) and S_2 should think B will win (or perhaps have no opinion). If you were to ask S_1 who she thinks will win, she would do just fine answering 'A', whereas she would seem to represent herself as irrational in answering either 'B' or 'C'. Likewise, if you were to ask S_2 who he thinks will win, he would do just fine answering 'B', whereas he would seem to represent himself as irrational in answering either 'A' or 'C'. And all of this is despite the fact S_2 has strictly greater rational credence that A will win than S_1 has: S_2 would be less surprised than S_1 if A were to win, would be willing to take worse bets on A's winning, would have reason to think S_1 is under-confident about A's prospects, etc. So we have a case in which there are two agents such that the first is rationally permitted to think that p while the second is not, and yet the second's rational credence that p is higher.

The same point can be made with one agent instead of two. All we need to do is modify the case so that the two sets of rational credences are both S_1's, only relativized to different times. Perhaps the initial assessment of ({40%, 35%, 25%}) is determined by S_1's knowledge of each horse's historical track record, while the later assessment ({45%, 50%, 5%}) is the combination of this background knowledge plus some insider information from a jockey. However the details are spelled out, it is obviously possible for S_1's rational credences to evolve in this way. And if and when they do, S_1 will go from being rationally permitted to think that horse A will win to being rationally prohibited from thinking it—all while becoming strictly more confident that horse A will win. Hence NON-MONTONICTY.

6 Thinking and closure

We have good reason to believe that thinking is extremely weak and that the relationship between rational thinking and rational credence is non-monotonic. Both claims support the idea that rational thinking is about according what one thinks to what's most likely to be true given one's evidence. This section will provide yet further evidence for this claim by considering the vexed question of whether rational thinking is closed under entailment—i.e., whether the following principle is true:[11,12]

CLOSURE

If a set of propositions Γ is such that S is rationally permitted to think every member of it, and Γ entails p, then S is rationally permitted to think that p.

There is a strong case to be made for CLOSURE. For one thing, it is highly plausible that rational agents can come to think that a proposition is true by deducing it from other propositions they already think are true. This seems to be exactly what we do when we try to reason through things. For another, ordinary thought and talk about thinking seems to take CLOSURE for granted. Notice how puzzling it is to speak as if it had false instances:

(7) a. ✗ I think B will be at the party. I also think C will be at the party. But I wouldn't say I think both B and C will be at the party.

b. ✗ I think B will be at the party. And I think that if B will be at the party C will be there too. But it's not fair to say that I think C will be at the party.

(8) Q: Do you think it will rain tomorrow?

A: Yes.

[11] For further discussion see, e.g., Kyburg (1961); Makinson (1965); Foley (1992b); Ryan (1996); Douven (2002); Christensen (2004); Lin and Kelly (2012); Leitgeb (2013, 2014).

[12] Here we leave the notion of 'entailment' unanalyzed. We will assume that obvious instances of conjunction introduction and modus ponens count as entailments, but we won't take a stand on whether (e.g.) all metaphysical entailments do. Thankfully, getting clear on these difficult issues is not essential for any of what follows.

Q: And do you think it will rain the day after?

A: Why yes I do.

Q: So you think it will rain tomorrow and the day after?

A: ✗No, I wouldn't say that.

If CLOSURE were false, we would expect there to be many situations in which sentences like (7) and (8) would seem felicitous. But this seems not to be the case.

As is well known, however, CLOSURE also happens to face a number of putative counterexamples. Take Kyburg's (1961) famous lottery puzzle, for instance. You know that each of the 100 entrants to an upcoming lottery has a 1% chance of winning. Consequently, for each entrant it seems rationally permissible to think is that that entrant won't win. But of course you know that *someone* has to win, and are thus not rationally permitted to think that none of the entrants will. And therein lies the problem: if (i) for each entrant you're rationally permitted to think that that entrant won't win, yet (ii) you are *not* rationally permitted to think that none of the entrants will, then rational thinking cannot be closed under conjunction.

The standard explanation of why CLOSURE is supposed to fail in these sorts of cases is in terms of epistemic risk aggregation. Taken individually, each of the conjuncts in the conjunction *The first entrant won't win* $\wedge$ *The second entrant won't win* $\wedge \ldots \wedge$ *The last entrant won't win* has a very high chance of being true, and is thus rationally thinkable. But each conjunct has *some* chance of being false, and as one conjoins them these chances aggregate, eventually resulting in a proposition that is guaranteed to be false. And this is allegedly why, contrary to CLOSURE, one can rationally think each member of a set of propositions without being entitled to rationally think their conjunction.

One point I think is worth emphasizing is that even if we do take cases like Kyburg's to undermine CLOSURE, we should reject these sorts of risk-theoretic diagnoses of its invalidity. After all, we know from §4's discussion of EXTREME WEAKNESS that there is no principled barrier to rationally thinking a proposition in which one's rational credence is arbitrarily close to zero. So it's not obvious why increasing the number of conjuncts in the relevant lottery proposition should automatically make it less fit as an object of rational thought. More to the point, though: if we're willing to take the "rational thinking is about thinking most likely" slogan seriously, then we should expect to find even more striking putative counterexamples to CLOSURE—ones that have nothing in particular to do with the aggregation of risk.

And indeed such examples are not hard to come by. Here is one. Suppose we're trying to track down James Bond. Our sources tell us there is a 40% chance he is hiding in London and a 20% chance he is hiding in each of Berlin, Frankfurt, and Munich. In light of this information, are we in a position to think that Bond is hiding in London? Well, suppose we are asked 'Which city do think Bond is in?'. By my lights, 'We're not sure, but we think he's in London' is a perfectly appropriate response. So we must be rationally permitted to think that Bond is in London. But are we rationally permitted to think that Bond is in the *United Kingdom*? It's not so clear. If asked

‘Which country do you think Bond is in?’, the most natural response would be ‘We’re not sure, but we think he’s in Germany’, or perhaps just the agnostic ‘We’re not sure’. If instead we were to answer ‘We’re not sure, but we think he’s in the United Kingdom’, it would be reasonable to object that we are significantly more confident that he is in Germany. So despite the fact that we appear to be rationally permitted to think that Bond is in London, we don’t appear to be rationally permitted to think that Bond is in the United Kingdom. And this is despite the fact that we are certain that London is in the United Kingdom, and thus have no less rational credence that he’s in the United Kingdom than we do that he is in London.

What to make of CLOSURE then? At this point it is hard to say. But this much is clear: if the putative counterexamples to the principle are genuine, we are owed a debunking explanation of the theoretical arguments in favor of the principle, as well as the intuitions to the effect that discourses like (7)–(8) are defective. And if the putative counterexamples are merely putative, then we are owed a debunking explanation of at least one of the intuitive judgments driving them.

We will ultimately leverage the “rational thinking is about thinking most likely” slogan to argue in favor of the second approach. In particular, we will take CLOSURE to be valid, and explain away the recalcitrant data in terms of illicit shifts in the underlying class of alternatives against which the relevant ‘thinks’-reports are assessed. But first let’s take stock.

7 Against existing views

So far we’ve argued in favor of the following three principles:

WEAKNESS

The evidential requirements on thinking that p are weaker than the evidential requirements on asserting that p, using p as a premise in deliberation, or being sure that p.

EXTREME WEAKNESS

There is no positive number x such that: necessarily, for any agent S and proposition p, if $C_S(p) \leq x$, then S is not rationally permitted to think that p.

NON-MONOTONICITY

There can be two agents (or the same agent at different times) S_1 and S_2 such that: $C_{S_1}(p) < C_{S_2}(p)$, yet S_1 is rationally permitted to think that p while S_2 is not.

We have also made an equivocal case for:

CLOSURE

If a set of propositions Γ is such that S is rationally permitted to think every member of it, and Γ entails p, then S is rationally permitted to think that p.

And since it will be useful in what's to come, we'll also make explicit the existence of a *consistency* requirement on rational thinking, which is the natural principle underlying the badness of speeches like (2, 'I think it's raining and that it isn't'):[13]

CONSISTENCY

S is not rationally permitted to be such that: S thinks that p and S thinks that ¬p.

The aim now is to find a theory of rational thinking that delivers WEAKNESS, EXTREME WEAKNESS, NON-MONOTONICITY, and CONSISTENCY while accounting for the considerations both for and against CLOSURE. We'll start by ruling out the two main existing theories of rational thinking: Cartesianism and Lockeanism.[14]

7.1 Against Cartesianism

According to the Cartesian conception of rational thinking, the evidential requirements on thinking as the same as those on *full belief*:[15]

CARTESIANISM

S is rationally permitted to think that p iff S is rationally permitted to fully believe that p.

What does it take to fully believe that p? Well, that depends on which Cartesian you ask. But the answers tend to share a family resemblance. On some ways of understanding the notion, S fully believes that p just in case S is disposed to assert p or use p as a premise in deliberation. On others, S fully believes that p just in case S is an internal duplicate of agent who knows that p. And yet on others full belief is understood in terms of epistemological notions familiar from ordinary language: S fully believes that p just in case S is sure (or certain) that p.

These ways of understanding 'full belief' are neither obviously equivalent nor obviously inequivalent. But we needn't worry about settling these matters. So long as one's preferred interpretation of 'full belief' is anywhere in the vicinity of the sorts of interpretations just described, CARTESIANISM is obviously going to be a non-starter. Rational agents who know that there is a substantial chance that ¬p are not disposed to assert p or use p as a premise in deliberation; nor are they the internal duplicates of agents who know that p; nor are they sure (certain) that

[13] One might have worries about whether CONSISTENCY is valid in full generality given the the possibility of identity confusion (and perhaps also the existence of the semantic paradoxes). The uses to which we put CONSISTENCY will not exploit any these sorts of considerations, so we can harmlessly ignore them in what follows.

[14] We will often talk about "existing" or "mainstream" theories of (rational) thinking, despite the fact that few authors have taken up the task of giving a theory of rational thinking *per se*. If thinking is believing (and 'theory of' doesn't generate an opaque context) then theories of rational belief just are theories of rational thinking, and so this way of talking is exegetically appropriate. But since we want the discussion to be compatible with the view that thinking isn't believing, readers should feel free to understand talk of "existing theories of thinking", "so-and-so's theory of thinking", etc., as shorthand for "existing theories of belief transposed into theories of thinking","so-and-so's theory of belief transposed into a theory of thinking", etc.

[15] See, e.g., Hintikka (1962); Stalnaker (1984); Williamson (2000); Buchak (2014); Ross and Schroeder (2014); Greco (2015); Staffel (2016). I should also mention that the label 'Cartesianism' is there for vivacity rather than historical accuracy. But see Chignell (2018, §1.2) for evidence that something like it was indeed Descartes' view on belief.

p. But as we know from §3's discussion of WEAKNESS (and even more dramatically from §4's discussion of EXTREME WEAKNESS), rational agents who know that there is a substantial chance that ¬p can do perfectly well in thinking that p. So the rational norms on thinking are not the rational norms on full belief.

7.2 Against Lockeanism

So much for CARTESIANISM. Next we turn to its main rival:[16]

LOCKEANISM

S is rationally permitted to think that p iff S's rational credence that p is sufficiently high.

How high is 'sufficiently high'? Well, that depends on which Lockean you ask.[17] But on standard versions of LOCKEANISM, sufficiently high rational credence is *at a minimum* rational credence greater than .5. The reason why is simple. Suppose the threshold for rational thinking is no greater than .5—that is: if $C_S(p) = .5$, then S is rationally permitted to think that p. By assumption $C_S(p) = .5$ iff $C_S(\neg p) = .5$. It thus follows that if $C_S(p) = .5$, then S is rationally permitted to think that ¬p too. But CONSISTENCY tells us that rational agents are never permitted to think both that p and that ¬p. So it seems the threshold for rational thinking has to be greater than .5. This gives us:

SIMPLE LOCKEANISM

S is rationally permitted to think that p only if $C_S(p) > .5$.

The obvious problem with SIMPLE LOCKEANISM is that it is incompatible with EXTREME WEAKNESS.[18] Rational agents can do perfectly well thinking that p even when their rational credence that p is arbitrarily close to zero. So any version of LOCKEANISM that validates SIMPLE LOCKEANISM must be false.

But that doesn't mean LOCKEANISM must be false, for it is possible to be a Lockean without being a simple Lockean. Indeed, if one follows Dorst's (2019) example and makes the notion of 'sufficient likelihood' both *context*- and *proposition*-sensitive, then one will have the resources to account for EXTREME WEAKNESS. The Lockean should thus become a *sophisticated* Lockean:

SOPHISTICATED LOCKEANISM

For all contexts *c*: 'S is rationally permitted to think that p' expresses a true proposition in *c* iff $C_S(p) > T_{<c,p>}$.

[16] See, e.g., Foley (1992a); Sturgeon (2008); Foley (2009); Leitgeb (2013); Beddor and Goldstein (2018); Dorst (2019); Moss (2019). I should mention that as it is used in the literature, 'Lockeanism' is something of an umbrella term, covering theories of the norms of belief (thinking), as well as theories of its metaphysics. For our purposes we can lump these views together. But our official target will just be the Lockean theories of *rational* thinking.

[17] We'll ignore those who say that the only sufficiently high rational credence is rational credence 1, as for our dialectical purposes such a view collapses the distinction between LOCKEANSIM and CARTESIANISM. But for defenses of this brand of LOCKEANISM see e.g., Clarke (2013); Greco (2015). (See also the discussion of Moss's (2019) view in footnote 19.)

[18] Note that this argument works just as well with SUBSTANTIAL WEAKNESS in place of EXTREME WEAKNESS

Here '$T_{<c,p>}$' should be read as 'the threshold for rationally thinking that p according to context *c*'. Allowing ourselves some looseness with use and mention, what SOPHISTICATED LOCKEANISM says is that whether S is rationally permitted to think that p depends on whether S's rational credence that p is sufficiently high by the standards context sets *for p*.[19]

Here's how going context- and proposition-sensitive allows the Lockean to accommodate EXTREME WEAKNESS without losing CONSISTENCY. First, we stipulate that for every context *c* and proposition p, $T_{<c,p>} + T_{<c,\neg p>} \geq 1$. Since S's rational credence that p and S's rational credence that ¬p will always sum to 1, this stipulation guarantees that there is no context in which 'S rationally thinks that p and S rationally thinks that ¬p' expresses a true proposition. This in turn guarantees that there won't be any counterexamples to CONSISTENCY. Second, we stipulate that for no proposition p is there a positive number x such that in every context c, $T_{<c,p>} \geq x$. That is to say: for any given proposition p and real number $x > 0$, there is always a context in which the threshold for rationally thinking that p is less than x. This guarantees that the view has the flexibility to account for the cases motivating EXTREME WEAKNESS.

Still, SOPHISTICATED LOCKEANISM faces other significant challenges. In addition to requiring *ad hoc* stipulations about the coordination of the proposition-sensitive thresholds for rational thinking, the view remains powerless to account for NON-MONOTONICITY or for the complexities

[19] As far as I can tell Moss (2019) defends a version of SOPHISTICATED LOCKEANISM. On Moss's view, the inference from 'S thinks that p' to 'S has credence 1 that p' is semantically valid. On the surface, then, her view is the kind of LOCKEANISM that collapses the distinction between it and CARTESIANISM. But Moss is well aware of the evidence that thinking is (extremely) weak (pp. 275–6). To account for this, Moss claims that in many contexts we treat as equivalent 'S has credence 1 that p' and 'S has credence near enough to 1 that p', with the interpretation of 'near enough' shifting between these contexts. And although Moss doesn't explicitly speak to the issue, it will become clear in a moment that the contextually determined extension of 'near' will also have to be proposition-sensitive—lest she predict the existence of contexts that invalidate CONSISTENCY. Consequently, I believe every objection raised against Dorst's view applies just as well to Moss's.

surrounding CLOSURE.[20]

Start with NON-MONTONICITY. Despite going in for context- and proposition-sensitivity, the SOPHISTICATED LOCKEANISM *is* a version of LOCKEANISM, and is thus committed to the core idea that whether one is permitted to rationally think that p depends on whether one's rational credence that p *is sufficiently high*. What counts as 'sufficiently high' might change depending on context and the proposition in question, but hold those two things fixed and you fix the evidential requirements on rational thinking. This means that the sophisticated Lockean (indeed, any Lockean) is inevitably committed to the idea that the relationship between rational credence and rational thinking is monotonic: i.e., that if an agent is rationally permitted to think that p (in context) while having rational credence x that p, then any agent with rational credence $y \geq x$ that p is rationally permitted to think that p (in context) too. The reason why is that holding context fixed, the inference from (i) 'S has sufficiently high rational credence that p' and (ii) 'S* has higher rational credence that p than S' to (iii) 'S* has sufficiently high rational credence that p' is plainly valid. But of course we know from §5 that the relationship between rational thinking and rational credence is non-monotonic. Whether an agent is rationally permitted to think that p depends on more than just whether their rational credence that p exceeds some absolute threshold; it also depends on whether there are any salient alternatives to p in which that agent has higher rational credence. And this is why—contrary to SOPHISTICATED LOCKEANISM—it is possible for 'S* has higher rational credence than S that p', 'S is rationally permitted to think that p', and 'S* is not rationally permitted to think that p' all to express true propositions in a single context.

With respect to CLOSURE, the sophisticated Lockean predicts the existence of many contexts in which the principle fails. This is for the simple reason that whenever 'sufficiently high ratio-

[20] Another worry that I won't be discussing in the main text concerns what the SOPHISTICATED LOCKEAN is supposed to say about the infelicity of *epistemic contradictions* under 'thinks', as in:

(9) a. ✗ I think the house is empty but might not be.
  b. ✗ I think the house is empty. I also think that it might not be.

If we assume that when 'might p' is embedded under 'thinks', it means means something along the lines of 'It's compatible with what I'm sure of that p', then (10a) and (10b) should express propositions that rational agents could knowably assert. On the Lockean view (whether SIMPLE or SOPHISTICATED), both would basically be a way of saying 'I have strong but inconclusive evidence that the house is empty'. But evidently they don't say that, as unlike this sentence neither seems to be assertable.

Whether or not the Lockean has the resources to account for these data is a delicate matter. Beddor and Goldstein (2018) give a semantics for 'might' that lets them predict the badness of (10a) but not (10b). But they also argue that it is less obvious how bad (10b) is. Rothschild (2019) expresses skepticism about the Lockean's prospects, especially in light of the fact that sentences that seem like they ought to imply the truth of (10a) and (10b) are fine, e.g.:

(10) a. ✓ I think the house is empty. But it might not be.
  b. ✓ I think the house is empty. But I know that it might not be.

And Mandelkern (2019) argues that his theory of epistemic modality can account for these contrasts independently of the correct account of rational thinking. So the force of the worry here is far from clear.

nal credence' means something other than 'has rational credence 1'—which we know it often will—it will be possible to have sufficiently high rational credence that p and sufficiently high rational credence that q without having sufficiently high rational credence that p *and* q, or to have sufficiently high rational credence that p and sufficiently high rational credence that p ⊃ q without having sufficiently high rational credence that q. This much is fine in a vacuum. The problem is that it's not at all clear why CLOSURE should *seem* valid if SOPHISTICATED LOCKEANISM is true. For according to the sophisticated Lockean, being in a position to rationally think that p *just is* having sufficiently high rational credence that p (even if what counts as 'sufficiently high' varies from context to context and proposition to proposition). So the view seems powerless to explain why speeches like (11) should so often seem infelicitous:[21]

(11) ✗ I think it will rain Monday. I also think it will rain Tuesday. But I wouldn't say I think it will rain both Monday and Tuesday.

I conclude that considerations from NON-MONOTONICITY and CLOSURE provide strong evidence against SOPHISTICATED LOCKEANISM. Indeed, I believe they provide strong evidence that we ought to abandon entirely the thought that rational thinking has anything to do with having sufficiently high rational credence. In its place we should embrace the thought that rational thinking is about having *highest* rational credence. The next section few sections develop this idea in detail.

8 Rational thinking as thinking most likely

Our goal is to give a theory of rational thinking that does justice to the intuitive thought that rational thinking is about thinking true the proposition *most* supported by one's evidence. Since whether a proposition is most supported by one's evidence depends on the alternatives to which it is compared, we're going to have to build alternative-sensitivity into our theory of rational thinking. To do this we'll start by briefly complicating our picture of thinking *simpliciter*.

On the standard picture, thinking (simpliciter) is a two-place relation between an agent and a proposition. 'S thinks that p' expresses a truth just in case S stands in the thinking relation to the proposition denoted by p. We are going to reject the standard picture in favor of one on which

[21] Moss (2019, p. 279) tries to account for the badness of speeches like (11) in terms of shifts in the contextually determined threshold on rational thinking. In particular, she claims that the mere consideration of puzzles like Kyburg's (1961) tends to induce contexts in which 'S is rationally permitted to think that p' entails 'S's rational credence that p is greater than S's rational credence that so-and-so will lose the lottery'. I find this diagnosis of the puzzle implausible. It incorrectly predicts that the contexts in which we grapple with the lottery puzzle are those in which we should find it infelicitous to report ourselves as thinking things about tomorrow's weather. (One's rational credence that so-and-so won't win the lottery will often greatly exceed one's rational credence that it will rain tomorrow.) I also find the diagnosis insufficiently general, for it seems to do nothing to explain what is happening in the James Bond case. By assumption, one's rational credence that Bond is in London is exactly one's rational credence that Bond is in the United Kingdom. Yet 'We think Bond is in London' is assertable in its most natural contexts while 'We think Bond is in the United Kingdom' is not. Are we really to think that asking after Bond's current country of residence somehow invokes a more demanding interpretation of 'thinks'?

thinking is a *three*-place relation between an agent, a proposition, and a *partition*. A partition $Q^?$ is a set of mutually exclusive and exhaustive propositions: conjoin any two of its members and you'll get the contradictory $\bot$, disjoin all of its members and you'll get the tautologous $\top$.[22] If theorists working in the tradition of Hamblin (1958) and Groenendijk and Stokhof (1984) are right—and from here on out we will assume that they are—then partitions of this sort are the meanings of natural language *questions*. For example: the meaning of the question 'Is it true that A will win the race?' (at least on its most natural readings) is the partition {A wins, A doesn't win}, while the meaning of the question 'Who will win the race?' (again on its most natural readings) is the partition {A wins, B wins, C wins}.[23,24] Consequently, we will say that thinking is a three-place relation between an agent, a proposition, and a question. We will use phrases like 'S thinks that p relative to the question $Q^?$' to describe our three-place thinking relation, and we will use the shorthand 'thinks$_Q$ p' to indicate that p is thought relative to $Q^?$.[25]

Although we've said basically nothing about how our three-place thinking relation is working, we already know enough to know that the natural language expression 'thinks' must be context-sensitive. 'Thinks'-reports made in ordinary language take only two arguments at surface form. We say 'S thinks that p', not 'S thinks $Q^?$-ishly that p'. And although we can say things like 'S thinks that p is the answer to $Q^?$', surface form still suggests that we are describing a relation between an agent a proposition (albeit a proposition that happens to be about a question). So if thinking is question-sensitive and the attitude we talk about with 'thinks'-reports is thinking, it must be that the semantic value of 'thinks' is a function from contexts to question-sensitive thinking relations: 'S thinks that p' is true in c iff S thinks$_Q$ that p, for the c-supplied question $Q^?$.

So: agents don't think propositions are true or false *simpliciter*; they think propositions are true or false relative to certain questions. But what is it to think a proposition is true relative to

[22] Just as we took the notion of 'entailment' for granted in discussion of CLOSURE, we'll also be taking for granted the nature of the underlying space of possibilities over which the cells of the partition are supposed to be mutually exclusive and exhaustive.

[23] I say '(on its most most natural readings)' because natural language questions can be associated with different sets of possible answers on different occasions of use. For example, 'Who will win the race?' might in some contexts be associated with the partition $\{a \vee b,\ c\}$ (rather than $\{a,\ b,\ c\}$). If we wanted to be maximally careful we'd make sure to specify which of the candidate meanings we have in mind when we talk about "the" question raised by sentences like 'Who will win the race?', but usually it will be obvious which interpretation of the natural language question is intended.

[24] Observant readers may notice that the proposition that either A wins, B wins, or C wins isn't a tautology on any natural understanding of the notion, and thus that the set containing all and only those disjuncts (i.e., $\{a, b, c\}$) doesn't form a partition. We could get around this problem by identifying the natural readings of the question 'Who will win the race?' with the set $\{a, b, c, \emptyset\}$, where $\emptyset$ is defined as the negation of the disjunction of all the other elements of the set. But since the differences that would arise between associating the question 'Who will win the race?' with $\{a, b, c\}$ versus $\{a, b, c, \emptyset\}$ will not affect anything of substance, we will stick with the simpler (not fully partitional) set $\{a, b, c\}$ in what follows.

[25] Yalcin (2018) also argues for a question-sensitive theory of thinking, though his reasons for doing so—namely to help account for the problem of logical omniscience and explain the nature of concept possession—are quite different from the ones pertinent to this paper's discussion. For those familiar with Yalcin's work, it will soon emerge that the kind of question-sensitivity I claim thinking is subject to is fundamentally different than the kind Yalcin claims it is subject to. This is not to say thinking couldn't be subject to both kinds of question-sensitivity. It's just that it would imply that thinking is doubly question-sensitive.

a question? The answer to that will come in §10. It will be much easier to answer it after we've tried to answer the question of when an agent is *rationally permitted* to think a proposition is true relative to a certain question. So for now we'll simply take the notion of *a*rational question-sensitive thinking for granted.

What does it take to *rationally* think$_Q$ that p? Here's a first stab at it. Let E_S be S's evidence proposition (that is: the conjunction of all the proposition S is rationally permitted to be sure of; that is: the conjunction of all the propositions p such that $C_S(p) = 1$). And let S's *best guess* to $Q^?$ be the answer to $Q^?$ in which S has highest rational credence (if multiple answers are tied for first, S's best guess is their disjunction). We can then say that rational thinking is thinking in terms of one's best guess:

BEST GUESS

S is rationally permitted to think$_Q$ that p just in case: the conjunction of E_S and S's best guess to $Q^?$ entails p.

To both motivate and get a feel for how BEST GUESS is working, consider again the horse race case from §5's discussion of NON-MONOTONICITY. An upcoming horse race has three entrants: A, B, and C. You know their respective chances of winning are 40%, 35%, and 25%. Here are two things I might ask you:

1) Who do you think will win the race?

2) Do you think it is true that A will win the race?

Given your evidence, it would be perfectly appropriate to answer the first question with something like 'I think A will win', perfectly appropriate to answer the second question with something like 'No, I think one of the other horses will', but not at all appropriate to answer either question with something like 'I think both that A will win the race and that A won't win the race'. This is something we should be able to explain.

Here is how BEST GUESS does it. Letting a, b, and c be the propositions that horse A wins, B wins, and C wins respectively, we get two distinct partitions of logical space: $\{a, \neg a\}$ and $\{a, b, c\}$. The first corresponds roughly to the meaning of the ordinary language question 'Is it true that A will win the race?', the second to 'Who will win the race?'. We can use these questions to distinguish between two thinking relations: thinking$_{\{a,\neg a\}}$ and thinking$_{\{a,b,c\}}$. Since your rational credence that ¬a is greater than your rational credence that a, you are rationally permitted to think$_{\{a,\neg a\}}$ that A won't win. And since your rational credence that a is greater than both your rational credence that b and your rational credence that c, you are rationally permitted to think$_{\{a,b,c\}}$ that A will win. So with respect to the proposition that A will win, what you are rationally permitted to think$_{\{a,\neg a\}}$ is the opposite of what you are rationally permitted to think$_{\{a,b,c\}}$. Lastly, since your rational credence that $a \wedge \neg a$ is 0, there is no question $Q^?$ such that the answer to it you assign highest rational credence entails (with the rest of your evidence) the proposition that A will win and not win. Thus, there is no $Q^?$ such that you are rationally permitted to think$_Q$ that A will win and that A will not win.

Here is how we connect these facts about question-sensitive rational thinking to our ordinary language judgments about 'thinks'. In the contexts evoked by considering questions like 'Do you think it is true that A will win the race?', the contextually supplied question tends to be the polar one: $\{a, \neg a\}$. Since you are rationally permitted to think$_{\{a,\neg a\}}$ that A won't win the race, your assertion of 'I don't think A will win the race' is felicitous (i.e., both true and consistent with your being rational). Similarly, in the contexts evoked when considering questions like 'Who will win the race?', the contextually supplied question tends to be the wh- one: $\{a, b, c\}$. Since you are rationally permitted to think$_{\{a,b,c\}}$ that A will win the race, your assertion of 'I think A will win the race' is felicitous. And since there is no $Q^?$ for which you are in a position to rationally think$_Q$ that A will and that A won't win the race, there is no context in which your assertion of 'I think that A will and that A won't win the race' can be felicitous.

Given this analysis of the horse race case, it is straightforward to see how BEST GUESS manages to validate the various principles about rational thinking gathered over the first few sections of the paper.[26] Since p can be your best guess to $Q^?$ even when you know there is a substantial chance that ¬p, it's no surprise that thinking is weak. Indeed, since p can be your best guess to $Q^?$ even when your rational credence that p is arbitrarily close to zero, it's no surprise that thinking is *extremely* weak. The relationship between rational thinking and rational credence is non-monotonic because whether p is your best guess to $Q^?$ depends not on your absolute rational credence that p, but your relative rational credence that p. Just as Bill Gates can go from being the richest person in the world to the second richest all while increasing his fortune (say because Jeff Bezos increased his even more), you can go from having highest rational credence that p to having highest rational credence in a proposition that entails that ¬p all while gaining rational credence that p.

So far so good: BEST GUESS gives an elegant account of a series of principles that the Cartesian

[26] Now that we're working with a contextualist theory of 'thinks', the proper statement of principles like EXTREME WEAKNESS and NON-MONOTONICITY requires some meta-linguistic ascent. I leave these statements to this footnote as I think the intended interpretations should be easy enough to recover:

ML WEAKNESS Possibly, there is an agent S and context c such that: 'S is rationally permitted to think p' is true in c even though 'S is rationally permitted to assert that p', 'S is rationally permitted to use p as a premise in deliberation', and 'S is rationally permitted to be sure that p' are all false in c.

ML EXTREME WEAKNESS There is no positive number x such that: necessarily, if $C_S(p) \leq x$, then in every context c: 'S is not rationally permitted to think that p' is true in c.

ML NON-MONOTONICITY There can be two agents S_1 and S_2 (or the same agent at different times) such that $C_{S_1}(p) < C_{S_2}(p)$ and a context c such that: 'S_1 is rationally permitted to think that p' is true in c while 'S_2 is rationally permitted to think that p' is false in c.

ML CLOSURE In every context c: if 'A certain set of propositions Γ is such that S is rationally permitted to think every member of it' is true in c, and Γ entails p, then 'S is rationally permitted to think that p' is true in c too.

ML CONSISTENCY In every context c: 'S is not rationally permitted to be such that: S thinks that p and S thinks that ¬p' is true in c.

and Lockean theories of rational thinking do much less well with. But what about CLOSURE?

Well, given our question-theoretic interpretation of the context-sensitivity of 'thinks', the claim that rational thinking is closed under entailment is just the claim that for any question $Q^?$, one is rationally permitted to think$_Q$ what is entailed by the rest of what one is rationally permitted to think$_Q$. And as it turns out, BEST GUESS makes rational thinking$_Q$ closed under entailment. The reason why, in abstract, is that BEST GUESS is a Hintikkan (1962) analysis of rational thinking$_Q$, and any Hintikkan analysis of a propositional attitude will get you closure under entailment for it. More concretely, let BG_S be the set of worlds that results from intersecting E_S with S's best guess to $Q^?$. According to BEST GUESS, then, 'S is rationally permitted to think that p' expresses a true proposition in context just in case every world in BG_S is a p-world.

This makes it easy to see why BEST GUESS gets closure under entailment more or less for free.[27] For any two propositions p and q and world w, if p entails q, then w is a p-world only if it is a q-world. (If w were a p∧¬q-world, then p wouldn't entail q.) It thus follows that 'S is rationally permitted to think that p' is true only if every world in BG_S is a q-world. But if every world in BG_S is a q-world, then holding context fixed, 'S is rationally permitted to think that q' expresses a true proposition too. Therefore, 'S is rationally permitted to think that p' expresses a true proposition in context only if 'S is rationally permitted to think that q' expresses a true proposition in context. Since q is an arbitrary entailment of p, we have CLOSURE.

With respect to the intuitive evidence *against* closure under entailment, BEST GUESS's diagnosis is that the putative counterexamples invoke non-uniform resolutions of the context-sensitivity of 'thinks', and are thus not genuine counterexamples to the principle.

Consider the putative counterexample to the closure of rational thinking under material implication. S has rational credence .4 that Bond is in London, .2 that he is in each of Berlin, Frankfurt, and Munich respectively, and rational credence 1 that London is in the United Kingdom and that each of Berlin, Frankfurt, and Munich is in Germany. In any context in which 'S is rationally permitted to think Bond is in London' expresses a true proposition, 'S is rationally permitted to think Bond is in the United Kingdom' expresses a true proposition. And in any context in which 'S is rationally permitted to think Bond is in Germany' expresses a true proposition, 'S *isn't* rationally permitted to think Bond is in London' expresses a true proposition. The impression that one can be in a position to rationally think that Bond is in London but not that he is in the United Kingdom is simply due to the fact that what one is rationally permitted to think relative to the question 'Which city is Bond in?' is distinct from what one is rationally permitted to think relative to the question 'Which country is Bond in?'. *Mutatis mutandis* for Kyburg's (1961) lottery puzzle. For each entrant n, relative to the question 'Will entrant n win the lottery?' S is rationally permitted to think n won't win the lottery. But relative to the question 'Will someone win the lottery?' S is not rationally permitted to think of any individual entrant that that entrant will not win. Thus, there is no single context in which 'S is rationally permitted

[27] Here we just focus on single-premise closure, but generalizing to the multi-premise case is easy enough.

to think someone will win' and, for each *n*, 'S is rationally permitted to think entrant *n* won't win' all express true propositions.

In short: by combining the idea that thinking is a question-sensitive attitude whose norms are a function of agents' best guesses to those questions, BEST GUESS captures all of the puzzling properties of rational thinking outlined in the paper so far. By my lights, this makes BEST GUESS by far the best theory of rational thinking available.[28]

9 Optionality and cogency

But we can still do better. This is because BEST GUESS isn't well positioned to account for the phenomenon of *optionality* raised to salience in §4's discussion of EXTREME WEAKNESS. (We will also soon see that the existence of this phenomenon places some particularly strong constraints on the space of possible theories of thinking *simpliciter*.)

To help reintroduce the phenomenon let's focus on a horse race with four entrants—A, B, C, and D—where S's rational credences that each will win are $[a\colon .35,\ b\colon .30,\ c\colon .20,\ d\colon .15]$.

Suppose S is asked 'Who do you think will win the race?'. We know she does just fine in answering in terms of her best guess:

(12) ✓ A.

We also know (back from the discussion in §4) that she also does just fine in answering agnostically, as in:

(13) a. ✓ I don't know.
b. ✓ I'm not sure.
c. ✓ There isn't any horse in particular I think will win.
d. ✓ One of A, B, C, or D.

It thus appears that the norms of rationality leaves S with options. If S goes the opinionated way she can defend herself by pointing out that A has the best chances of winning. And if instead she goes the agnostic way she can defend herself by pointing out that that none of the horses is sure to win. In neither case would S seem to be making a mistake.

So far none of this is a problem for BEST GUESS, for BEST GUESS tells us that rational agents are *permitted* to think$_Q$ whatever is entailed by their best guess to $Q^?$. It doesn't say that they

[28] What about footnote 20's epistemic contradictions data? If we adopt the view defended by (e.g.) Yalcin (2007) and Mandelkern (2019) that 'S is rationally permitted to think that it might be that p' is true (on its most natural readings) only if there is a p-world compatible with what S is rationally permitted to think, then we get that 'S is rationally permitted to think that it might be that p' is true in *c* iff for the *c*-determined question $Q^?$, it is not the case that S is rationally permitted to think$_Q$ that ¬p. Consequently, the proponent of BEST GUESS predicts that relative to no question $Q^?$ can a rational agent think that both halves of an epistemic contradiction. That is to say: if 'S is rationally permitted to think that p' expresses a true proposition in *c*, then 'S is rationally permitted to think that it might be that ¬p' must express a false proposition in *c*.

must think it. The problems for BEST GUESS emerge when we consider some of the more subtle patterns in the possible answers to questions about what we think.

We can imagine the possible answers to the question 'Who do you think will win the race?' as falling on a spectrum. On one end of the spectrum we have the four maximally opinionated answers: 'A', 'B', 'C', and 'D'. On the other we have the maximally agnostic answers: 'I don't know', 'One of A, B, C, or D', etc. Between these two extremes we have "mixed" answers that are opinionated in some respects, agnostic in others: 'I'm not sure about A, but I do think it'll be A or B', 'B or D', etc. Given that S's rational credences are $[a\colon .35,\ b\colon .30,\ c\colon .20,\ d\colon .15]$, it is felicitous for her to give voice to some but not all of the possible mixed answers. These, for instance, are felicitous:

(14) ✓ There is no horse in particular I think will win, but I do think it will be either A or B.[29]

(15) ✓ All I can say is that I think it won't be D.[30]

These, for instance, are not:

(16) ✗ I think B or C will win.

(17) ✗ There is no horse in particular I think will win, but I do think it will be either A or C.

Intuitively both (16) and (17) have the feel of wishful thinking on S's part. (16) flies in the face of the fact that S knows that A has the best chance of winning, while (17) flies in the face of the fact that S knows that B has a better chance of winning than C. (If S is only willing to narrow down her view on who will win to two horses, she should be thinking that the winner will be A or B, not that it will be A or C.)

Our theory of rational thinking should be able to account for these sorts of judgments. Given BEST GUESS, it's no surprise why (14) and (15) are fine: S is rationally permitted to think$_{\{a,b,c,d\}}$ whatever is entailed by her best guess to the question 'Who will win the race?'. And both of these answers are indeed entailed by her best guess. It is also no surprise why (16) is infelicitous: the proposition that B or C will win the race is *not* entailed by her best guess to the question of who will win the race, and so S is not rationally permitted to think$_{\{a,b,c,d\}}$ it.

The kind of judgment that poses a problem for BEST GUESS is the one raised to salience by (17). S's best guess to the question of who will win the race is that A will. This proposition entails that A or C will. So BEST GUESS says S is rationally permitted to think$_{\{a,b,c,d\}}$ that A or C will win. And although the proposition that A will win is obviously entailed by itself, this doesn't mean that S is rationally required to think$_{\{a,b,c,d\}}$ that A will win the race. Again, BEST GUESS only tells us when an agent is rationally permitted to think that a proposition is true, not when they must think it. So for all BEST GUESS is concerned, S is perfectly rational in thinking$_{\{a,b,c,d\}}$

[29] Equivalently: 'There is no horse in particular I think will win, but I think it won't be C or D.'
[30] Equivalently: 'I think it will be one of A, B, or C.'

that the winner will be one of A or C, while not having a view$_{\{a,b,c,d\}}$ on whether A in particular will win. So BEST GUESS wrongly predicts that there should be nothing wrong with (17).

As I see it, the lesson to take from the infelicity of (17) is that what one is rationally permitted to think relative to a question $Q^?$ depends not only on the distribution of one's rational credences in its answers, but also on what one *in fact* thinks the answer to that question is. If S happens to think$_{\{a,b,c,d\}}$ that A will win (in accordance with her best guess), then she is indeed rationally permitted to think$_{\{a,b,c,d\}}$ that A or C will win. (The latter can be deduced from the former after all.) But if S *doesn't* think$_{\{a,b,c,d\}}$ A will win—which is what she says in uttering (17)—then S isn't rationally permitted to think$_{\{a,b,c,d\}}$ that A or C will win. At most she can rationally think$_{\{a,b,c,d\}}$ that which is entailed by the proposition that A or B will win.

Now to turn this idea into a theory. We'll start by defining S's **guess** to $Q^?$ as the strongest answer to $Q^?$ such that S thinks$_Q$ that answer is true.[31] By assumption, then, '**guess**' denotes a three-place relation between agents, questions, and unions of answers to those questions. If the question is $\{a, b, c, d\}$, S's **guess** might be maximally strong, as when it's a, b, c, or d; or it could be maximally weak, as when it's $a \vee b \vee c \vee d$; or it could be of middling strength, as when it's something like $a \vee b$ or $b \vee c \vee d$.

With this notion of **guessing** in place, we can give a schematic theory of rational thinking:

RATIONAL THINKING SCHEMA

S is rationally permitted to think$_Q$ that p just in case: (i) the conjunction of E_S and S's **guess** to $Q^?$ entails p; and (ii) S's **guess** to $Q^?$ is rational.

To turn RATIONAL THINKING SCHEMA into an actual theory, we just need to spell out the details of (ii).

The simplest idea would be to say that S's **guess** to $Q^?$ is rational if and only if it is either S's best guess to $Q^?$ (i.e. the answer in which S has highest rational credence) or is the trivial guess to $Q^?$ (i.e., the disjunction of all the answers). But this view wrongly predicts that speeches like (14) should be infelicitous—i.e., that rational agents cannot be such that the strongest thing they think about an upcoming horse race is that either first- *or* second-favorite will win.

A better idea is to say that S's **guess** to $Q^?$ is rational if and only if that guess is *cogent*:

COGENCY

For any agent S, question $Q^?$, and proposition p that is a union of answers to $Q^?$: p is *cogent* with respect to $Q^?$ and S iff: if there is an answer to $Q^?$ q such that q doesn't entail p, then there is no other answer to $Q^?$ q* such that: (i) q* entails p but (ii) $C_S(q) \geq C_S(q^*)$.

Although COGENCY is a mouthful, the intuitive idea behind it is fairly simple. What it says, in words, is that if p is S's **guess** to $Q^?$, then p better be such that: for each of the answers to $Q^?$ included in p, S has higher rational credence in those answers than in any of the answers

[31] I put '**guess**' in **boldface** to make it clear that it is taking this particular technical interpretation.

excluded from p. So, for example, if S's **guess** to the question $\{a, b, c, d\}$ is $a \vee b$, then S's **guess** is cogent iff S's rational credences are such that: $C_S(a) \sim C_S(b) > C_S(c) \sim C_S(d)$. More abstractly: the claim that S's **guess** to $Q^?$ is cogent is essentially the claim that, from the perspective of S's rational credence function, S's **guess** to $Q^?$ is non-arbitrary.

Having defined our cogency property, we can now give our final theory of rational thinking:

Cogent Guess

S is rationally permitted to think$_Q$ that p just in case: (i) the conjunction of E_S and S's **guess** to $Q^?$ entails p; and (ii) S's **guess** to $Q^?$ is cogent.

That is to say: S is rationally permitted to think that p relative to $Q^?$ just in case p is entailed by the strongest thing S thinks relative to $Q^?$, and the strongest thing S thinks relative to $Q^?$ is cogent.

COGENT GUESS correctly predicts that 'thinks'-reports like (17) should be infelicitous. Given that S's rational credences in the answers to the question 'Who will win the race?' are $[a$: .35, b: .30, c: .20, d: .15$]$, only the following **guesses** are cogent for S: a, $a \vee b$, $a \vee b \vee c$, and $a \vee b \vee c \vee d$. This means that relative to the question 'Who will win the race?', S is rationally permitted to be in all and only the following states: thinking that A will win, thinking merely that A or B will win, thinking merely that A, B, or C will win, or thinking merely that some horse or other will win. Thus, the only way for S to rationally think$_{\{a,b,c,d\}}$ that A or C will win is if she rationally thinks$_{\{a,b,c,d\}}$ that A will. So there is no way S can speak truly in uttering (17) while being rational.

With respect to our earlier principles of interest, COGENT GUESS is exactly like BEST GUESS. One's best guess to $Q^?$ will be cogent no matter one's absolute rational credence in it, so COGENT GUESS predicts EXTREME WEAKNESS. Likewise, the relationship between the cogency of one's possible **guesses** to $Q^?$ and one's rational credences is non-monotonic, so the view predicts NON-MONOTONICITY. Since COGENT GUESS is a Hintikkan analysis of 'is rationally permitted to think', it gets CLOSURE for free. And since it evokes the same kind of question-sensitivity as BEST GUESS, it gets the same debunking explanation of the putative counterexamples to CLOSURE. Finally, since cogent **guesses** never entail contradictions, COGENT GUESS is guaranteed to preserve CONSISTENCY.

We thus have our answer to:

The normative question:

Under what conditions is it rationally permissible to think that p?

One is rationally permitted to think that p relative to a question $Q^?$ just in case (i) p is entailed by one's **guess** to $Q^?$ (i.e., the strongest thing one thinks relative to $Q^?$); and (ii) that **guess** is cogent. And since for many agents S and questions $Q^?$ there are a range of possible cogent

guesses to $Q^?$, we predict that the norms of rationality are permissive with respect to thinking:[32]

PERMISSIVISM

It is not the case that: for any body of evidence E and proposition p, there is a unique doxastic attitude toward p that is consistent with being perfectly epistemically rational and having E as one's evidence.

—which, given the optionality in how one chooses to answer questions about one thinks, is exactly what we should expect to be true. Rational thinking is about thinking cogently. Thinking whatever is most likely to be true is one way to do this, but it is not the only way.[33]

10 Thinking *simpliciter*

Having answered our normative question, we can finally turn to:

The metaphysical question:

What is it to think that p?

We'll start with a schematic answer. Where E_S is the conjunction of all the propositions S is rationally permitted to be sure of (i.e. S's evidence), we'll say that B_S be the conjunction of all the propositions S is *actually* sure of. This gives us:

[32] For some recent discussion of issues related to PERMISSIVISM, see, e.g., Greco and Hedden (2016) and Schoenfield (2019).

[33] One issue I've chosen to suppress until now concerns what our theory of rational thinking should say about *pure* guesses. Suppose I show you a fair coin and ask you how you think it will land. How are you entitled to answer? If COGENT GUESS is correct—indeed, if any mainstream theory of belief interpreted as a theory of rational thinking is correct (save perhaps that of James 1956)—then there is only one correct answer: with agnosticism. Intuitively, you have no (evidential) grounds to favor heads to tails or vice-versa, so the only rationally permissible thing to do is to have no particular view on the matter.

The issue is that there is evidence from natural language that rational agents *are* entitled to think propositions on the basis of a pure guess:

(18) Q: I'm about to flip a fair coin. How do you think it will land?
A: Heads.
Q: Why do you think it will land heads?
A: I don't know; it's just a guess.

It is not at all clear that A's responses here are inappropriate. In fact, I think we often speak this way in forced-choice situations (game shows, multiple-choice tests, and so on).

If we wanted our theory of rational thinking to account for this fact, we'd have to adjust the COGENCY requirement so that it allows **guesses** to be made arbitrarily in the case of rational credal ties. This isn't a dramatic change to the theory, but it might cause one to worry about the methodology of giving substantial weight to ordinary language judgments about 'thinks'-reports. Should we *really* think that a theory of rational thinking must be able to account for these uses? Perhaps not. But supposing we don't, we run into difficult questions about where to draw the line between the uses that are semantically respectable and those that are not. I leave further investigation of the answers to these difficult questions for other work.

Thinking Schema

S thinks$_Q$ that p just in case: the conjunction of B_S and S's **guess** to $Q^?$ entails p.

Having earlier defined S's **guess** to $Q^?$ as the strongest proposition S thinks relative to $Q^?$, THINKING SCHEMA is, of course, a circular analysis of thinking. This is not to say it is uninformative. It entails that thinking is question-sensitive. It also entails that holding context fixed, the inference from 'S is sure that p' to 'S thinks that p' is valid. And it also entails that thinking has some nice closure properties.[34] But if possible it would be good to have an independent characterization of what it is about S that makes it so her **guess** to $Q^?$ is p rather than p*. This is the question the rest of the section will try to answer.

It will be helpful to start by saying something negative: whatever it is that determines whether one's **guess** to $Q^?$ is p rather than p*, it isn't one's credences. To see why, suppose that S_1's actual and rational credences in the answer to the question 'Who will win the race?' are $[a\colon .35,\ b\colon .30,\ c\colon .20,\ d\colon .15]$. Suppose also that S_2 is a credal duplicate of S_1 (i.e., that S_1 and S_2 have the same actual and rational credences), and that all this is common knowledge between them. (Maybe they both talked to the same expert horse bettor together.) We know from the previous section that a person whose rational and actual credences are $[a\colon .35,\ b\colon .30,\ c\colon .20,\ d\colon .15]$ does just fine in answering 'Who do you think will win the race?', in any of the following ways:

(19) a. ✓ A.
b. ✓ One of A or B.
c. ✓ One of A, B, or C
d. ✓ I have no idea.

What I want to suggest is that because someone with these credences *could* answer in any of these ways, it seems as if nothing guarantees S_1 and S_2 *will* answer in the same way. S_1 might answer with (19a), S_2 with (19d). And since S_1 and S_2 can (and in fact might) answer in different ways despite being credal duplicates, the natural hypothesis is that they can *think* different things despite being credal duplicates.

This point can be sharpened by imagining things from S_1's perspective. You know what your credence function says about the various possible outcomes of the horse race, and you know that S_2 has the same credence function as you. You also know how you will answer the question 'Who do you think will win the race?'. But do you know how S_2 will answer it? I don't see how you could. For all you know she could answer with any of (19a)–(19d).

[34] If one takes these closure properties to be a bug rather than a feature of THINKING SCHEMA, one could replace it with THINKING SCHEMA*. Letting $q^?$ be S's **guess** to $Q^?$:

Thinking Schema*
S thinks$_Q$ that p just in case: S is sure that $q^? \supset p$.

More generally, knowing a person's credences just doesn't seem to suffice for knowing how they'll answer questions like 'Who do you think will get the Democratic nomination?' or 'Where do you think Bond is hiding?'. They might answer in accordance with their best guess. But they might also answer agnostically. And until they answer you simply won't know. So knowing what a person's credences are doesn't suffice for knowing what they think. We thus have good reason to think that the facts about what we think do not supervene on the facts about our credences.

But if not our credences, then what *does* determine the facts about what we think? Is it our behavioral dispositions? That seems unlikely. Regardless of the fact that S_1 and S_2 might think different things about who will win the race, neither would use the proposition that A will win the race as a premise in theoretical or practical reasoning. Nor would either be willing to assert that A will win the race. Nor would either take a bet on the outcome of the race that the other wouldn't. Nor would either be more surprised than the other to find out that horse A had won the race. Indeed, it seems that the only relevant behavioral difference that could arise between S_1 and S_2 concerns the first-personal 'thinks'-reports they'd be willing to make about themselves. But clearly *that* can't be what grounds the differences in what they think. They make the 'thinks'-reports they do because of what they think, not the other way around.

In light of these difficulties, I want to suggest a rather different picture of what determines what we think. In particular, I want to suggest that the facts about what we think are determined by our *choices*—i.e., by certain kinds of pure acts of the will. It is because S_1 *chose* as her answer *A will win* and S_2 *chose* as her answer *One of the entrants will win* that S_1 thinks$_{\{a,b,c,d\}}$ that A will win while S_2 merely thinks$_{\{a,b,c,d\}}$ that some horse or other will.

I doubt an analysis can be given of the notion of 'choice' involved here, but the basic (and admittedly vague) idea can be grasped through some paradigm cases. In particular, the kind of choice involved in deciding what to think about a question seems to me akin to the kind of choice involved in picking whether to go left or right at the fork, or in figuring out how to compose the next sentence of an email, or in deciding which part of one's visual experience to attend to, or—perhaps most relevantly to the present discussion—in choosing between heads or tails when asked to guess how a coin will land. It's a kind of making up of one's mind. This is not to say choosing is always a conscious process, to be clear. One can make up one's mind without deliberating consciously about it. Choosing an answer to a question can be as automatic and subconscious as choosing which parts of one's environment to attend to, or how to execute the various steps in a complex physical or cognitive task—say, serving in tennis, playing a video game, choosing one's gestures or words in speech, and so on. But it is a choice nonetheless.

Although we won't say more to substantively characterize the kind of choices involved in thinking, we will assume they abide by three structural constraints. First, they are question-directed. Second, when presented with a question $Q^?$, what one chooses is the union of a subset of $Q^?$'s answers. And third, one's default choice of answer to a question is always the trivial one (i.e. the union of all that question's answers). The hypothesis, then, is that for p to be one's

guess to $Q^?$ (i.e., for p to be the strongest thing on thinks relative to $Q^?$) is for p to be one's choice of answer to $Q^?$. This gives us a proper (even if somewhat tentative) answer to the metaphysical question:

Doxastic Choice

S thinks$_Q$ that p just in case: the conjunction of B_S and S's choice of answer to $Q^?$ entails p.

And combining DOXASTIC CHOICE with COGENT GUESS, we get the following tidy picture of thinking and its norms: to think that p is for p to be one's choice of answer to the question at hand; to rationally think that p is for one's choice to be cogent.[35]

11 Thinking and believing

Let's take stock. Thinking is a question-directed attitude. There are two ways one can come to think a proposition p relative to a question $Q^?$. One can either be sure that p, or one can choose as one's answer to $Q^?$ a proposition that together with the rest of what one is sure of entails p.

Since thinking is a choice, it follows that a strong form of doxastic voluntarism is true.[36] It also follows that two agents with identical credence functions can think different things relative to the same question. This means that (rational) thinking does not supervene on (rational) credence. It also explains why two rational agents with the same evidence can come to different conclusions about whether p without either of them having made a mistake.

Here's what else we know about thinking: it is weak. You can rationally think that p while being unwilling to assert p or use p as premise in deliberation. In fact, we know that it is *extremely* weak. You can rationally think that p despite being arbitrarily close to certain that p is false. We also know the relationship between rational thinking and rational credence is non-monotonic: you can start off rationally thinking that p, become rationally strictly more confident that p than you were before, and as a result no longer be rationally permitted to think that p. Despite all this, we also know that rational thinking is closed under entailment. Whenever p entails q and you're rationally permitted to think that p, you will automatically be rationally permitted to think that q too.

Of course, if thinking is believing—if 'S thinks that p' is true in context iff 'S believes that p' is—then everything we've said about thinking is true of belief as well. And this would mean that belief cannot play many of the theoretical roles with which philosophers have associated it. If I

[35] Does DOXASTIC CHOICE imply that what one thinks about a question is *always* a voluntary matter? It does not. DOXASTIC CHOICE says that we can't help but think any proposition we're sure of. And it's an open empirical question whether people can ever choose to think non-cogently. Speaking for myself, I know that as much as I might want it to be true that I will win the lottery, I can't bring myself to think it. But the possibility of wishful thinking suggests that some agents are capable of choosing (i.e. thinking) non-cogently.

[36] I am not the first person to defend the view that there are cases in which agents exercise direct doxastic control over what they think. See, e.g., Nickel (2010); McHugh (2015); Roeber (2019a,b) for some recent defenses of that claim. However, I know of no author besides James (1956) whose view permits direct doxastic control in as wide a range of cases as DOXASTIC CHOICE does.

know that A has only a 10% chance of winning the upcoming race, then it is not rational for me to take for granted that A will win the race or to assert that A will win the race; nor is it rational for me to take an even-money bet that A will win the race (let alone a 3:1 or 4:1 bet). But it can be rational for me to believe that A will win the race. So belief is not the attitude we hold toward the propositions we rely on in theoretical or practical reasoning; nor is it the attitude we hold toward the propositions we are willing to assert; nor is it even the attitude we hold toward propositions that we find highly likely to be true.

More generally: if an epistemological theory assigns *belief* a theoretical role that obviously cannot be played by *educated guessing*, then we should be highly confident that that theory is false. Why? Because (i) COGENT GUESS is the correct theory of rational thinking; (ii) COGENT GUESS says that the norms on thinking are much like the norms on educated guessing;[37] (iii) thinking is believing; so (iv) the norms on belief must basically be the norms on educated guessing.

The lesson to take from this is that if thinking is believing, then just about every existing theory of belief is false. Or, more conservatively, *if* existing theories of belief are theories of the attitude that is the denotation of the ordinary expression 'believe', then just about every existing theory of belief is false.

Many theorists of "belief" might happily accept this conditional while denying its antecedent. Perhaps all along they've taken themselves to be giving theories of the distinct attitude *opining*, where (as a matter of stipulation) to opine a proposition p is to do some number of the following: treat p as a premise in deliberation; be willing to assert p; be the internal duplicate of one who knows p; be disposed to feel surprise upon learning that ¬p; and so on. It is worth noting that many of the questions we've been asking in this paper—Is rational thinking closed under entailment? Does thinking supervene on credence? Can thinking ever be voluntary?—retain much of their theoretical interest when 'thinking' is replaced with 'opining'. So it's not clear how much is lost conceding that one's theory of opining is a theory of something other than belief.[38] And for those who think that having a home in natural language is a precondition on an attitude's being a worthwhile object of epistemological study, note that all the claims just made about acceptance seem to lose neither plausibility nor theoretical interest when stated in terms of *being sure*. So perhaps that's the attitude epistemologists have been theorizing about all along.[39]

Defenders of traditional conceptions of belief who are uncomfortable with the idea of restating their theories in terms of opining or surety are in a more difficult situation. If thinking is choosing but believing isn't, then thinking can't be believing. But the evidence that thinking is believing is very strong. Again, if there are coherent interpretations of sentences like

(1a) ✗ I think it's raining, but I wouldn't say I believe it is.

[37] See Horowitz (2019) for more on the norms of educated guessing.
[38] Though for opposition to this view see (Moss, 2019, pp. 273-4) and Williamson (2018, §6).
[39] Cf. Ayer (1956); Chisholm (1957).

(1b) ✗ I'm not sure whether Jane thinks Federer will win Wimbledon, but I know she doesn't believe he will.

(1c) ✗ My friends think I'm a good a person, but my mom believes I am.

they appear to be elusive. And readers can see for themselves whether their judgments about the paper's example sentences are affected in any substantial way when 'believes' is substituted for 'thinks'. I myself detect no meaningful difference in my reactions to the examples (except perhaps that they sound less colloquial). And this raises the question: If thinking isn't believing—indeed, if thinking isn't anything *like* believing—then why does natural language seem to treat the two as if they were the same?

Whether those who defend traditional conceptions of belief can give a satisfying answer to this question is not something we will resolve here. But the conditional point remains: if thinking is believing, then believing isn't about being sure, or even about being sufficiently sure. It's about choosing. And your evidence only settles what you ought to choose when it is decisive on the matter. Otherwise the choice is yours. So long as you choose cogently, you can take comfort knowing that your choice will be rational.

References

Ayer, A. J. (1956). The problem of knowledge. *British Journal for the Philosophy of Science*, 9(34):159–163.

Beddor, B. and Goldstein, S. (2018). Believing epistemic contradictions. *Review of Symbolic Logic*, (1):87–114.

Buchak, L. (2014). Belief, credence, and norms. *Philosophical Studies*, 169(2):1–27.

Chignell, A. (2018). The ethics of belief. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2018 edition.

Chisholm, R. M. (1957). *Perceiving: A Philosophical Study*. Cornell University Press.

Christensen, D. (2004). *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford University Press.

Clarke, R. (2013). Belief is credence one (in context). *Philosophers' Imprint*, 13:1–18.

Dorst, K. (2019). Lockeans maximize expected accuracy. *Mind*, 128:175–211.

Douven, I. (2002). A new solution to the paradoxes of rational acceptability. *British Journal for the Philosophy of Science*, 53(3):391–410.

Foley, R. (1992a). The epistemology of belief and the epistemology of degrees of belief. *American Philosophical Quarterly*, 29(2):111–124.

Foley, R. (1992b). *Working Without a Net: A Study of Egocentric Epistemology*. New York: Oxford University Press.

Foley, R. (2009). Beliefs, degrees of belief, and the lockean thesis. In Huber, F. and Schmidt-Petri, C., editors, *Degrees of Belief*, pages 37–47. Springer.

Greco, D. (2015). How i learned to stop worrying and love probability 1. *Philosophical Perspectives*, 29(1):179–201.

Greco, D. and Hedden, B. (2016). Uniqueness and metaepistemology. *Journal of Philosophy*, 113(8):365–395.

Groenendijk, J. and Stokhof, M. (1984). *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, University of Amsterdam.

Hamblin, C. L. (1958). Questions. *Australasian Journal of Philosophy*, 36(3):159–168.

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

Hintikka, J. (1962). *Knowledge and Belief*. Ithaca: Cornell University Press.

Horowitz, S. (2019). Accuracy and educated guesses. In Gendler, T. S. and Hawthorne, J., editors, *Oxford Studies in Epistemology 6*. Oxford Scholarship Online.

James, W. (1956). *The Will to Believe and Other Essays in Popular Philosophy. Human Immortality; Two Supposed Objections to the Doctrine*. Dover Publications.

Kaplan, M. (1995). Believing the improbable. *Philosophical Studies*, 77(1):117–146.

Kyburg, H. (1961). *Probability and the Logic of Rational Belief*. Wesleyan University Press.

Leitgeb, H. (2013). Reducing belief simpliciter to degrees of belief. *Annals of Pure and Applied Logic*, 164(12):1338–1389.

Leitgeb, H. (2014). The stability theory of belief. *Philosophical Review*, 123(2):131–171.

Lin, H. and Kelly, K. (2012). A geo-logical solution to the lottery paradox, with applications to conditional logic. *Synthese*, 186(2):531–575.

Makinson, D. (1965). The paradox of the preface. *Analysis*, 25(6):205–207.

Mandelkern, M. (2019). Bounded modality. *Philosophical Review*, 1(128):1–61.

McHugh, C. (2015). The illusion of exclusivity. *European Journal of Philosophy*, 23(4):1117–1136.

Moss, S. (2019). Full belief and loose speech. *Philosophical Review*, pages 1–35.

Nickel, P. J. (2010). Voluntary belief on a reasonable basis. *Philosophy and Phenomenological Research*, 81(2):312–334.

Roeber, B. (2019a). Evidence, judgment, and belief at will. *Mind*, pages 1–18.

Roeber, B. (2019b). Permissive situations and direct doxastic control. *Philosophy and Phenomenological Research*, pages 1–17.

Ross, J. and Schroeder, M. (2014). Belief, credence, and pragmatic encroachment1. *Philosophy and Phenomenological Research*, 88(2):259–288.

Rothschild, D. (2019). What it takes to believe. *Philosophical Studies*, pages 1–19.

Ryan, S. (1996). The epistemic virtues of consistency. *Synthese*, 109(2):121–141.

Schoenfield, M. (2019). Permissivism and the value of rationality a challenge to the uniqueness thesis. *Philosophy and Phenomenological Research*, pages 1–17.

Staffel, J. (2016). Beliefs, buses and lotteries: Why rational belief can't be stably high credence. *Philosophical Studies*, 173(7):1721–1734.

Stalnaker, R. (1984). *Inquiry*. Cambridge University Press.

Stanley, J. (2008). Knowledge and certainty. *Philosophical Issues*, 18(1):35–57.

Sturgeon, S. (2008). Reason and the grain of belief. *Noûs*, 42(1):139–165.

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

Williamson, T. (2018). Knowledge, credence, and the strength of belief. Unpublished manuscript.

Windschitl, P. D. and Wells, G. L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology*, 75(6):1411–1423.

Yalcin, S. (2007). Epistemic modals. *Mind*, 116(464):983–1026.

Yalcin, S. (2010). Probability operators. *Philosophy Compass*, 5(11):916–37.

Yalcin, S. (2018). Belief as question-sensitive. *Philosophy and Phenomenological Research*, 97(1):23–47.