



What it takes to believe

Daniel Rothschild¹ 

© The Author(s) 2019

Abstract Much linguistic evidence supports the view believing something only requires thinking it likely. I assess and reject a rival view, based on recent work on homogeneity in natural language, according to which belief is a strong, demanding attitude. I discuss the implications of the linguistic considerations about ‘believe’ for our philosophical accounts of belief.

Keywords Belief · Credence · Lexical semantics

1 Introduction

Here are two questions:

- What is the meaning of the word ‘believe’?
- What is belief?

The former question tends to occupy semanticists while the latter is the focus of philosophical discussions in epistemology and philosophy of mind.

On the surface of it, the questions seem related. We might expect a theory of the meaning of ‘believe’ to tell us what it takes for a sentence of the form ‘*x* believes that *p*’ to be true. To have this knowledge would seem to help us understand what belief is. However, in practice when semanticists answer questions about the meaning of a term like ‘believe’ they tend to rely on cognate terms in their metalanguage. Philosophers interested in belief are not going to be content with being

✉ Daniel Rothschild
d.rothschild@ucl.ac.uk

¹ University College London, London, United Kingdom

told that for x to believe p is for x to bear the belief-relation to the proposition p .

However, I think considerations from the semantics of ‘believe’ have a non-trivial bearing on how philosophers ought to think about belief. My discussion will mostly center on the claim that belief attributions are in a certain sense *weak* made in ‘Belief is weak’ (Hawthorne, Rothschild, and Spector 2016, henceforth HRS). I’ll first elaborate and defend that claim, responding in detail to an interesting challenge from recent work on homogeneity in natural language. I will move on to discussing the potential implications of the weakness of belief for philosophical accounts of belief. I will argue that while the semantic considerations might naturally support the Lockean view of belief, they can also be leveraged to support other positions.¹

2 Three claims about the weakness of belief

Here are the key claims from HRS I want to revisit:

- (A) The attitude verbs ‘believe’ and ‘think’ semantically express the same attitude.²
- (B) ‘Believe’/‘think’ have a semantically weak reading on which ‘Lilly believes/thinks p ’ is true even when Lilly thinks that p is only likely.
- (C) ‘Believe’/‘think’ never have a very strong lexical meanings, i.e., in no context does ‘Lily thinks p ’ come to mean Lily has ruled out the possibility of not p or is sure/certain that p .

Before going on to the evidence for these claims I want to clarify them in three respects: first, I will say a bit more about what I mean by weakness in (B); second, I will discuss the relationship between these claims and philosophical orthodoxy about ‘belief’; and, third, I will discuss the relationship between these claims and the pragmatics of belief-ascriptions.

In HRS, ‘we argue[d] that believing a proposition merely requires thinking it likely’. I still hold that view, but it requires more clarification. I do not take the view that ‘believing that p is likely’ is *equivalent* to ‘believing that p ’. Rather I hold that insofar as there are probabilistic or evidential requirements on believing that p they do not go beyond the probabilistic or evidential requirements on believing that p is likely. This does not preclude there being other, non-probabilistic requirements on belief.³ (B) simply states that ‘believe’ can be weak in this sense; (C) says it never has a reading with the stronger commitment that believing p requires being certain that p .

¹ Much of the discussion in HRS focused on the relationship between belief and the norm of assertion. I do not address this issue here, though the arguments here all support the conclusions of HRS.

² They are only the same in ‘believe that’/‘think that’-type constructions. In almost every other construction they differ substantially, such as in their nominalizations. They also show a different distribution of adverbial modifications, e.g., ‘fully’ in the ‘that’-constructions.

³ I am thinking particularly of some form of *commitment* or *conviction*, but I don’t have a very good characterization of this quality.

Williamson (2018) argues that the view of the meaning of ‘believe’, above, particularly, (C), is implausible as it requires thinking that generations of epistemologists have simply been confused about the words they have used.

...if believing *p* likely suffices for believing *p*, an obvious question arises as to whether someone who accepts *p* only as far as thinking *p* likely can count as knowing *p*, however well they do on the other conditions. Yet when the most detailed survey of post-Gettier epistemology considers published worries about the belief condition, including several motivated by concerns about ordinary usage, they are all to the effect that it is too strong, not that it is too weak (Shope 1983: 171–192). In a published literature to which hundreds of native speakers of English contributed, quite a few of them ordinary language philosophers alert to nuances of natural language, it would be odd that people did not object to what would arguably be a gross misuse of ‘believe’, if HRS were right.

I am not arguing that stereotypical instances of belief attributions do not, in fact, involve strong belief, i.e., being sure or certain, or that we cannot in certain contexts use sentences like ‘John believes *p*’ to convey the thought (somehow or other) that John is sure that *p*. All I am arguing is that the verb ‘believe’ does not have such a strong reading.

To give an analogy, I take it that ‘open the car door’ means, roughly speaking, to move the car door in such a way as to allow ingress into the car. However, generally when we talk about people opening the car door, we mean that they do so in the normal manner: using the handle, rather than, say, a chainsaw or a crowbar. But I wouldn’t want to say that the expression ‘open the car door’ ever actually means to open the car door in the usual way. Rather, in many cases it is simply a tacit conversational assumption that if someone opened a car door they did so in the normal way. Similarly while talk of belief often attributes, in some sense, certainty, I think this is not part of the contextual meaning of the attitude verb ‘believe’.⁴ If we convey stronger meanings using ‘believe’ and ‘think’, it is only by post-compositional pragmatic inference.⁵

In the next section I’ll go through the most powerful arguments in HRS for these claims.

⁴ Here I probably differ from any standard pragmatic enrichment theories, such as relevance theory (Sperber and Wilson 1986; Carston 2002) and Recanati’s theory of free enrichment (Recanati 2013). As it happens many of the arguments in HRS can be used to argue against the existence of certain pre-compositional enrichment processes. I discuss the relationship between the arguments in belief is weak and various forms of truth-conditional pragmatics in a handout (Rothschild 2015).

⁵ Williamson (2018) suggests that allowing for stronger readings of ‘believe’ due to strengthening may be implausible since the speaker by not using a stronger term (such as ‘being sure/certain’) would suggest s/he is not in a position to make the stronger claim. This Gricean reasoning, however, only goes through if there is some genuine pragmatic competition between ‘think’/‘believe’ and ‘sure’/‘certain’. Given the radically lower frequency of use of the latter terms—as easily verified, for example on google ngrams <https://books.google.com/ngrams>)—I’m inclined to think there isn’t any such automatic competition.

3 The arguments

- Let's start with (A).⁶ If 'believe' and 'think' express extensionally different attitudes then it should be possible to at once to assert one and deny the other. It is not possible, however, as (1) demonstrates.⁷

(1) ?Even if she thinks John is at the party, it's not fair to say she believes it.

Compare:

(2) Even if she thinks John is at the party, it's not fair to say she's certain of it.

Similarly, we cannot contrast the two attitudes with 'but'.

(3) ? Bill thinks John is at the party, but Heather believes he is.

Compare:⁸

(4) Bill thinks John is at the party, but Heather is sure he is.

- Now let us move on to the arguments for (B). First, observe that 'believe' is *neg-raising*. That is, the following two sentences can be read as equivalent:

- (5) a. I don't believe Jesus wept.
b. I believe Jesus didn't weep.

As Horn (1978, 1989) observes, neg-raising verbs tend to be weak. He writes:

What is common to all NR predicates is the relative slimness of the functional difference between the preraised form with lower negation and the logical form with the upstairs negative taking wide scope. It is the closeness of the external (contradictory) readings of *not likely*, *not believe*, *not advisable to*, *likely not*, *believe not*, *advisable not*, respectively, which renders the negated predicates potential neg-raisers, and the relative distance of *not possible*, *not*

⁶ While I haven't seen many philosophers explicitly argue against (A), a little googling turns up many coherent pieces on the web explaining that 'believe' is stronger than 'think'. Here's an illustrative example:

There is a hierarchy when it comes to the ways we express ourselves and our convictions. There are some things we believe, some things we think, and some things we feel. The terms are hierarchical rather than synonymous and over time we ought to see a progression from feeling to thinking to believing. We should want to elevate more of what we feel into what we think and more of what we think into what we believe. I will grant that there can be fine distinctions here, but there is still value in distinguishing them, at least for our purposes. (<https://www.challies.com/articles/i-feel-i-think-i-believe/>, accessed April 2018)

⁷ These examples control for *neg-raising* which I discuss in a moment.

⁸ Focus on 'thinks' and 'is sure' is probably necessary to get the right reading of (4), but such focus does not rescue (3).

realize, not obligatory from *possible not, realize not, obligatory not* which removes these from contention. (Horn 1978, p. 196)

The point is that if ‘believe’ is strong, then its negation is not going to be close to not believing. But if ‘believe’ has a weak reading (close to believing likely) then we get the slim functional difference Horn describes as characterizing the class of neg-raisers generally.

It is worth noting that this not just a useful generalization about neg-raisers but also provides a clue to understanding why there is neg-raising to begin with. For instance, if ‘likely p ’ just means ‘ p is more likely than not,’ to infer ‘likely not p ’ from ‘not likely p ’ all we have to assume is that p and not p don’t have exactly the same likelihood. If the speaker can be taken to be making this assumption, then an assertion of ‘not likely p ’ will come to mean ‘likely not p ’. Thus the slim functional difference Horn describes fits in well with the view that neg-raising is a kind of conversational implicature. However, there is no consensus in the literature that neg-raising is the result of an implicature.⁹

Another piece of evidence for the weakness of belief is the presuppositional behavior of ‘believe’ under questions:

- (6) Who does Tom think/believe won the election?
 $\not\rightarrow$ There is someone Tom is sure won the election.

Normally a question presupposes that there is one true answer. Compare:

- (7) Who is Tom sure won the election?
 \rightsquigarrow There is someone Tom is sure won the election.

The natural explanation of why (6) carries a weak presupposition (or none at all) is simply that ‘think’/‘believe’ is weaker than ‘be sure’. If it simply means ‘think likely’ then (6) would presuppose only that Tom thought one candidate more likely to have won than another, which is surely the most it could be felt to presuppose.

Another argument for a weak reading of ‘believe’ is the coherence of the following sentence:

- (8) William thinks/believes it’s raining, but he’s not sure.

If ‘believe’/‘think’ only had a strong reading this would not be coherent. Direct evidence for the idea that believing p means thinking p likely comes from cases of prediction. An assertion of:

⁹ See Horn and Bayer (1984), Horn (1989), and Romoli (2013) for development of a scalar-implicature view of neg-raising. Note that one of the main challenges to these views is that the verbs that neg-raise vary considerably across languages: a prominent example is ‘hope’ (not a neg-raiser in English) vs. ‘hoffen’ (a neg-raiser in German). I explore an alternative, semantic explanation of neg-raising in the next section.

(9) Lily thinks/believes Jose won the race.

seems to require only that Lily thinks Jose is the most likely winner, not that she is sure Jose won.¹⁰ As HRS point out, this type of statement can be used to ascribe belief in cases where the believed probability of the proposition is less than .5 (though greater than all the relevant alternatives). This suggests that meaning of ‘believe’ or ‘think’ is really much weaker than that of ‘be sure’.

- I will now turn to evidence for the stronger claim (C), that ‘think’ and believe’ lack strong readings. The way I argue for this claim, as in HRS, is to show that even in circumstances in which a strong reading of ‘think’ or ‘believe’ would be necessary for a sentence to be coherent, that reading is still not available. (10) is an example of this type.
- (10) ? It’s not that Tim thinks John will win, nor that he doesn’t think John won’t win, but he thinks it is likely John will win.

If we read ‘think’ as strong (10) should make perfect sense, but it doesn’t.¹¹
A related example is:

(11) ? While I SUSPECT Tim is at the party, it’s not that I THINK/BELIEVE he is.

For a control, contrast this with:

(12) While I SUSPECT Tim is at the party, it’s not that I am SURE he is.

Another contrast:

- (13) a. ? Lilly only SUSPECTS that Bill stole the painting, but Bill THINKS/
BELIEVES he did.
b. Lilly only SUSPECTS that Bill stole the painting, but Bill is SURE he did.

If ‘thinks/believes’ had a possible reading in which they attribute certainty we would expect that reading to be accessed in (13-a). But the infelicity of (13-a) suggests we cannot access this stronger reading.¹²

Here is another example that suggests that ‘to believe *p*’ can never mean to be sure that *p* (this comes from Levi Spectre, p.c.):

¹⁰ For clarity assume the winner is not known in the relevant context, so the reading is not available where Lily is sure Jose won the race, but is mistaken.

¹¹ Switching ‘think’ with ‘believe’ might slightly improve (10), but is not without oddity. It seems to me that ‘think’ and ‘believe’ have slightly different interactions with intonation that may explain some differences. This is one of the reasons, as I note below, the evidence for (C) is suggestive rather than conclusive.

¹² It seems to me that (13-a) can be partially rescued by strong intonation, but the contrast with (13-b) suggests that this is not just normal contrastive focus. Strong intonations allow access to metalinguistic readings as well as, possibly, ad hoc meanings for words.

- (14) **A:** Can you pay for this?
B: Sorry, but I don't have any money left in my account.
A: ✓ Are you sure/certain you don't ?
 ? Do you believe/think you don't?

What is odd here is that the variant of **A**'s answer using 'believe' or 'think' is infelicitous. If 'believe' really did have a strong reading one would expect that reading to be accessed in this context and the contribution to be felicitous.

Before closing this section, I should comment on the relative strength of the different linguistic evidence. I consider the evidence for (A) to be strong: 'believe' does not systematically have a higher evidential requirement than 'think'. If anything, the reverse is true: 'believe' seems to have a use that suggests belief despite knowing the odds are against you.¹³ I consider the evidence for (B) to be definitive: it is completely clear that 'believe' and 'think' have easily available weak readings. The case for (C) is more tentative. There is suggestive evidence that strong readings are not available. This evidence might best be considered as a challenge to those who think that the strong readings should be the most easily accessible and prominent ones, rather than as constituting a solid case against the existence of those readings.

4 Homogeneous, strong belief

Despite these arguments, there is a rival view that can be teased out of the recent semantics literature that 'believe' is strong. Križ (2015) gives a general theory of homogeneity (the way certain terms require all-or-nothing readings) and applies it to attitude verbs such as 'believe'.¹⁴ In this section I will present a version of his homogeneous semantics for attitude verbs that seems to make 'believe' strong in an empirically plausible way.

Let me first outline the essential features of this homogeneous view of belief. We begin with a strong, Hintikkian semantics of 'believe': for x to believe p is for every world doxastically accessible to x to be a p -world.¹⁵ However, we classify 'believe' as a *homogeneous* expression in the sense developed by Križ. We then appeal to special properties of homogeneous expressions to reply to the arguments of HRS.

I should note that it is not clear to me that this view is Križ's. For Križ does not discuss what maximal readings of belief amount to, i.e. whether believing p is tantamount to being sure that p , or ruling out all doxastic possibilities that not- p in

¹³ Una Stojnic pressed this point to me (p.c.). For example, it seems as though you can know that your candidate is unlikely to win, but still believe he can win. Although I don't account for this use here, it doesn't affect any of my main points below.

¹⁴ In this last respect he follows Gajewski (2005), who applies a presupposition of homogeneity to belief ascriptions and other instances of neg-raising. Križ's overall account of homogeneity synthesizes and advances an earlier literature, starting at least with Schwarzschild (1993).

¹⁵ For a probabilistic variant of this view we could make the domain of universal quantification the probability measure, or something along these lines.

the usual sense. The only commitment of the homogeneity-view is that ‘believe’ acts as a universal quantifier over belief-worlds (however exactly belief-worlds are understood). The strength of such a universal semantics for ‘believe’ will depend on what the worlds quantifying over are: if these are a subset of the worlds that are doxastically possible in the normal sense then the view might be weak. Thus, we need to add to Križ’s view the commitment to ‘believe’ being strong. This commitment yields a position that is similar to the position by Moss (forthcoming). Moss argues that believing p is tantamount to being certain that p , and that loose talk accounts for the way in which we often ascribe belief in cases that fall short of certainty.¹⁶ Križ (ch. 3) also uses the the idea of loose talk in his treatment of homogeneity, arguing that the non-maximal reading (explained below) should be accounted for in terms of it. So both views share a strong basic semantics and a pragmatic weakening mechanism involving loose talk.¹⁷

In the rest of this section, I’ll outline some of the main characteristics of homogeneous expressions, and explain why we might think that ‘believe’ satisfies them. The four features I will focus on are STRENGTH, STRONG NEGATION, NON-MAXIMALITY, and NON-PRESUPPOSITION. Expressions that exhibit these properties, according to Križ, include plural definites, conditionals, questions, and bare plurals as well as neg-raising attitude verbs such as ‘believe’. I’ll use plural definites to illustrate these properties, as they are the canonical example of a homogeneous expression-type and the focus of much of Križ’s work.

Let’s go through the four properties.

- I. STRENGTH: an unembedded assertion of a plural definite entails (at least as a default) something like a universal reading:

(15) The boys passed.
 \rightsquigarrow All the boys passed.

Similarly the strong theorist of belief thinks that normal belief attributions have strong entailments: e.g. ‘believes p ’ entails ‘is sure that p ’.

- II. Combined with STRENGTH we also have STRONG NEGATION: under negation a plural definite acts as if it is weak, leading to a stronger reading than expected:

(16) I doubt the boys passed.
 \rightsquigarrow I think none of the boys passed.

¹⁶ Lasersohn (1999) gives the canonical discussion of loose talk that influences both Moss and Križ’s discussions.

¹⁷ I thought of this particular variant of Križ’s view only after an earlier version of Moss’s paper. I won’t directly discuss Moss’s views here, but I (weakly) believe that the view I sketch here is the most promising way of developing her position.

This contrasts with:

(17) I doubt all the boys passed.

Another way of seeing this phenomenon is to note that negation acts as though it is scopeless with respect to ‘the boys’, but negation has a scope ambiguity with respect to ‘all the boys’.

(18) a. The boys don’t like me.

b. All the boys don’t like me.

The neg-raising property of ‘believe’ can now be explained as being the result of STRONG NEGATION.

(19) John doesn’t believe it’s raining

↔ John believes it isn’t raining.

This explanation of neg-raising is surprising since HRS used neg-raising as an argument for a semantics for weak belief. We pointed out that in general neg-raising verbs are weaker; for example, ‘love’ is not neg-raising but ‘like’ is. However, in the context of Gajewski and Križ’s observations this argument for belief being weak clearly does not work.¹⁸

III. NON-MAXIMALITY is the property that plural definites can tolerate some exceptions to their universality:¹⁹

(20) The boys came although little Sam stayed at home as he was ill.

versus

(21) ? All the boys came although little Sam stayed at home as he was ill.

As we shall see in the next section NON-MAXIMALITY is one of the most subtle properties of homogeneous expressions. While we tolerate exceptions, we cannot straightforwardly list counterexamples to the universal claims of plural definites:

(22) ? Athena stayed on Mount Olympus but the children of Zeus went down to Hades.

Similarly NON-MAXIMALITY leads to an explanation of why we can have weak readings of ‘believe’ but not ‘sure’, as shown in this contrast:

¹⁸ Indeed, I think neg-raising is just suggestive of the weakness of belief, it is not itself a definitive piece of evidence.

¹⁹ Brisson (1998) is an early and canonical discussion of this phenomenon.

- (23) a. John believes it's raining, but he has doubts.
 b. ? John is sure it's raining, but he has doubts.

Note that we can make an explicit analogy here: 'the boys' : 'all the boys' as 'thinking that' : 'being certain/sure that'. In each case, the two expressions have similar meanings but only the expression to the left of the colon is homogeneous.

- IV. Generally when an expression and its negation both lead to an entailment we say that the entailment is *presupposed*.²⁰ There is a common content to both (24) and (25):

- (24) The boys passed.
 (25) The boys didn't pass.

The common content is that all the boys passed or all the boys didn't. However, this content does not behave like a presupposition.

- (26) If the boys passed, then I'll be happy.
 $\not\rightarrow$ Either the boys all passed or they all didn't pass.

Compare to how the presupposition of 'stop' behaves in the antecedent of a conditional:

- (27) If Wisconsin stops electing Republican governors, UW will prosper.
 \rightsquigarrow Wisconsin has been electing Republican governors.

I call this property of having some common content that does not behave like a presupposition the NON-PRESUPPOSITION of homogeneous expressions. Belief attributions clearly share this property, as this example shows:

- (28) If Lilly thinks it's raining, she'll bring her umbrella.
 $\not\rightarrow$ Either Lilly is confident it's raining or she's confident it's not raining.

Thus, there seems to be good evidence that 'believe' behaves like a homogeneous expression. To have a strong, homogeneous view of 'believe', we might want to adopt the semantics and pragmatics that Križ (2015) uses in his dissertation for all homogeneous expressions. We need not go through the details here, but the basic idea is to use a trivalent semantics that makes ' x believes p ' true iff all of x 's belief-worlds are ones in which p is true, false iff all of x 's belief-worlds are ones in which p is false, and neither true nor false in other cases.²¹ The trivalent logic naturally explains neg-raising, since the conditions for the negation of a sentence to be true require more than just the sentence failing to be true. Moreover Križ (2015, ch. 3) provides a compelling account of NON-MAXIMALITY as loose speech, which seems to

²⁰ Though there are exceptions, such as the case of conventional implicatures.

²¹ This is explored in Križ (2015, ch. 2). For those allergic to non-classical logics there may be bivalent or supervaluationist variants of this view.

have the potential to explain why, although belief is semantically strong, we can attribute belief in cases where people only have high credences.

5 Against homogeneous, strong belief

I do not think that the strong, homogeneous view of ‘believe’ is correct. I am attracted in general to Križ’s account of homogeneity and its potential to unify an array of diverse and puzzling semantic and pragmatic phenomena. And I am prepared to concede that belief attributions might even be homogeneous. However, I do not think that belief attributions—homogeneous or not—are semantically strong (i.e., that literally believing p requires ruling out the possibility that p is false). I have two arguments against the homogeneous strong belief hypothesis.

- First, the observed weakness of belief is not adequately explained by NON-MAXIMALITY. Note, first, that with plural definites it is not, in general, possible to assert both the plural definite and a negative witness in one breath.²²

(29) ? The students passed but one of them didn’t.

By contrast it is possible, as HRS observe, to assert belief while expressing doubts (in one breath):

(30) [Context: open question that it’s raining] Tilly thinks it’s raining, but she’s not sure.

(31) [Context: common ground that it’s raining] Tilly thinks it’s raining, but she doesn’t know it.

As I said, there are some cases in which one can explicitly discuss exceptions with plural definites, but these are rather special cases such as the following examples from Križ (2015, p. 72):

- (32) a. The professors are smiling. Of course, not Smith, but you know, he never smiles; it doesn’t mean anything.
 b. The townspeople are asleep. Of course, the gatekeeper is probably still up, but we know that he’s always there anyway.

NON-MAXIMALITY is fragile and seems to be order-dependent with plural definites. We can contrast (32) with (33).²³

(33) Smith isn’t smiling, but you know, he never smiles; it doesn’t mean anything. Still, the professors are smiling.

In (33), unlike in (32), one gets a strong contextual inference that Smith is not a professor. No such pattern can be seen with exceptions to belief statements, where either order seems fine.

²² Križ (2015, ch. 3).

²³ Connections between plural definites and Sobel sequences have long been noted in this literature, see, for example, Schlenker (2004).

(34) Tilly isn't sure Steve passed, but she still believes/thinks he passed.

- The second reason I do not accept the homogeneous view is that belief attributions simply do not behave as if they are as semantically strong as attributions of sureness/certainty. I have already given arguments for this claim in section 3, particularly in the context of examples (11) and (14). For more evidence in this vein, specifically distinguishing belief-attributions from plural definites, consider this contrast:²⁴

(35) a. ? Bill and Lilly's children came to the event, in fact, all of them did.
b. I believe it's raining, in fact, I am certain/sure it's raining.

The infelicity of (35-a) is evidence that 'Bill and Lilly's children' is almost the same in meaning as 'All of Bill and Lilly's children, however, the felicity of (35-b) suggests that 'I believe *p*' does not bear such a close relationship to 'I am certain/sure that *p*'.

I conclude that we should not treat 'believe' as a strong, homogeneous expression in the sense I explored in the previous section.

6 What it takes

What type of semantics can we give 'believe'/'think' that would explain the kind of behavior I outlined above? What do the semantic considerations above tell us, if anything, about the nature of the mental states we commonly call beliefs? I will go through a few proposals.

6.1 Believing as being sure

Some philosophers want to equate belief with certainty or probability one.²⁵ While it may be that certainty plays an essential role in our mental lives, my arguments here suggest that 'be certain' is not even a contextually possible meaning of 'believe'. For this reason, I think it is a mistake to use the English verb 'believe' as a philosophical term for the state of being sure or certain.

6.2 Lockeanism

Here is Foley (1992) explicating a classic version of the Lockean thesis:

To say that we believe a proposition is just to say that we are sufficiently confident of its truth for our attitude to be one of belief. Then it is epistemically rational for us to believe a proposition just in case it is

²⁴ Thanks to Diego Feinmann for this example.

²⁵ Recent examples include Clarke (2013), Greco (2015), and Moss (2017, forthcoming)

epistemically rational for us to have sufficiently high degree of confidence in it, sufficiently high to make our attitude towards it one of belief.

This view equates belief with having sufficient confidence in a proposition. HRS argue that the behavior of the word ‘believe’ would seem to support such a Lockean view as long as the threshold is sufficiently low to account for the data. In particular, while the threshold may exhibit some context sensitivity, some of the cases above show that in many cases we cannot access a reading of ‘believe’ with a high threshold, even where it is strongly favored by context.

There are a number of unresolved problems facing the Lockean view, however, perhaps the most pressing being the failure of closure (Kyburg 1961). It seems, after all, a basic datum of natural language inference that belief is closed under conjunction:²⁶

- (36) ? John believes it will rain today and he believes it will rain tomorrow, but he doesn’t believe it will rain today and tomorrow.

Another worry about the Lockean view comes from considering epistemic contradictions involving belief. Both (37) and (38-b) sound as though they attribute contradictory beliefs to John.²⁷

- (37) ? John thinks [it’s raining but it might not be raining].
 (38) a. John thinks/believes it’s raining but he knows/realizes/acknowledges it might not be.
 b. ? John thinks it’s raining but he thinks it might not be raining.

The Lockean needs to explain why these examples sound contradictory: To believe might not-*p* seems generally just to require having not-*p* being compatible with one’s belief.²⁸ Then if believing *p* only requires believing *p* is likely, it should be possible to both believe *p* and believe that *p* might not be true. Luckily Beddor and Goldstein (2017) have already done the work of explaining why (37) is bad despite belief being weak. However, what explains the contrast between (38-a) and (38-b) remains an open question and a pressing one for the Lockean.²⁹

6.3 Knowledge first (belief last)

Here is Williamson’s description of his knowledge-first account of belief.

²⁶ Leitgeb (2014) defends a Lockean view that accepts closure, but only at the cost of putting rather strong constraints on which probability distributions agents can hold.

²⁷ Beddor and Goldstein (2017) discuss (38-b) and argue that it is not as infelicitous as standard epistemic contradictions such as (i).

²⁸ See Yalcin (2007) for a defense of this view.

²⁹ As Yalcin (p.c.) notes, this is one case where ‘believe’ and ‘believe likely’ do seem to come apart. For (i) seems substantially better than (38-b):

- (i) John thinks it’s likely it’s raining but he thinks it might not be raining.

...to believe p is to treat p as if one knew p —that is, to treat p in ways similar to the ways in which subjects treat propositions which they know. In particular, a factive propositional attitude to a proposition is characteristically associated with reliance on it as a premise in practical reasoning, for good functional reasons; such reliance is crucial to belief. A creature which lacks a concept of knowing can still treat a proposition in ways in which it treats propositions which it knows. The primitive creature does not treat the proposition that food is present like that when merely desiring that food is present; it does not use the proposition as a premise in practical reasoning. By contrast, the person who genuinely believes that there is a god by a leap of faith does rely on that premise in such reasoning. The unconfident examinee who tentatively gives p as an answer is little disposed to rely on p as a premise, and for that reason does not clearly believe p , but for the same reason does not clearly know p . Although a full-blown exact conceptual analysis of *believes* in terms of *knows* is too much to expect, we can still postulate a looser connection along these lines. (Williamson 2000, pp. 46–47)

One feature of this view which HRS picked out to criticize is Williamson's claim that knowledge is the 'norm' of 'believe' (i.e. believe p only if you know that p). We pointed out that this view seemed improbable since it does not suggest any form of error or irrationality to both attribute a belief to an individual and deny that they know it.

(39) John doesn't know Bill is his friend, but he believes he is.

Rather (39) simply seems to say, on one prominent reading, that John lacks the kind of evidence or confidence necessary for knowledge. In this example, then, whatever state we are attributing to John is not one of which knowledge could be the norm of. Thus at least some attributions of belief appear to be attributing a different kind of state from the one Williamson describes.

Williamson (2018) has responded to many of the arguments in HRS, conceding in effect that there is a use of 'believe p ' in which it expresses a relatively weak form of commitment (presumably rationally compatible with taking oneself not to know p). However, Williamson insists that 'believe' can and often does express the stronger 'full belief' or 'outright belief' of his knowledge-first account. While Williamson does not think full belief in p requires certainty in p in the usual sense, it would seem to be incompatible with taking oneself not to know p . There does not seem to me to be good evidence that 'believe' gets to have such a strong reading.

To support his view, Williamson suggests that our locution 'fully believe' directly expresses the paradigmatic state of outright belief that comes out of his knowledge-first account. However, consideration of actual usage of 'fully believe', I think, tells against this suggestion. Consider (40):³⁰

(40) Amanda doesn't know Peter will come home, but she fully believes he will.

As far as I understand it, Williamson takes full belief in p to be a state in which one accepts p to the same extent that one would accept p if one knew p . However, as I

³⁰ Thanks to Levi Spectre and Matt Mandelkern for making this point to me.

read (40) it is attributing to Amanda a belief which self-consciously lacks the normal support enjoyed by knowledge.³¹ Similarly a first person case is no admission of irrationality:

(41) I don't know Peter will come, but I fully believe it.

More generally, I worry that Williamson's account has too little to say about the semantics of 'believe' to give a satisfying explanation of its behavior. For instance, if 'believing p ' expresses as its primary meaning a commitment to p similar to the commitment we would have if we knew p , it is mysterious why it should be so easy to use 'believing p ' to express a relatively weak state such as merely thinking p likely. Williamson (2018) argues these weaker states are *similar* to knowledge states, but he gives no particular account of why 'believe' should express the particular range of states that it does.^{32,33} Thus, even if believing does involve being in knowledge-like states, this connection does little to explain how we talk about belief. The Lockean view, by contrast, directly explains why belief attributions are true of those propositions that we think more likely than not: as on the Lockean view that is just what 'believe' means.³⁴

7 Hintikkian mysterianism

The Lockean view is often taken to be opposed to the classic view attributed to Hintikka (1962) according to which 'believe' can be understood as expressing universal quantification over those worlds compatible with one's belief. This view has both been dominant in epistemic logic and in natural language semantics. It is natural to oppose Hintikka's universal quantification view to the Lockean view since the former has it that belief rules out doubt while the latter has it that belief is compatible with considerable doubt.³⁵

³¹ For me the natural reading of (40) is not one in which Amanda is either unjustified or is in a Gettier type case.

³² Williamson in particular argues that 'believe' should inherit some of the context sensitivity of 'similar'—since it describes the propositions that are treated in a similar way to knowledge. While much of the behavior of 'believe' is compatible with this claim, still given the range of ways in which things can be treated similarly, very little about 'believe' actually follows from this claim.

³³ Moreover, these states of weak belief, which Williamson acknowledges are attributed by belief attributions, do not seem to be governed by the knowledge norm for belief.

³⁴ Though as Williamson notes, there are interesting cases that may be hard to explain if 'believe' is weak. He gives the example of a preacher telling a congregation:

- (i) You need to believe that Jesus can save you.

I think here the religious connotations of 'believe' play a strong role. In addition, as I argued in Sect. 2, even a weak belief attribution can give rise to strong readings.

³⁵ Dorst (2019), for example, explicitly treats the views as opposed.

However, merely taking ‘believe’ to express universal quantification over some worlds does not necessarily make belief strong.³⁶ We could have a view, thus, that fails to identify the belief worlds—the worlds quantified over in belief attributions—with doxastic possibilities in the usual sense. Rather the belief worlds would be a subset of the doxastic possibilities, though presumably a subset with a reasonably high probability.

One thing this view could do is provide a satisfying explanation of the contrast in (38) repeated here:

- (38) a. John thinks/believes it’s raining but he knows/realizes/acknowledges it might not be.
 b. ? John thinks it’s raining but he thinks it might not be raining.

The key idea is that ‘might,’ when embedded in this belief report, requires that there be a belief-world in which it doesn’t rain, leading to the contradictory sense of (38-b). Whereas when ‘might’ is embedded under other attitude verbs it does not force there be such belief-worlds (since its force is not derivative on its embedding inside belief).³⁷

Of course, such a view would also explain the closure of attitudes under conjunction, since this follows immediately from the universal quantification posited for belief attributions. Moreover, this view is compatible with a homogeneity-type treatment of belief-attributions, and thus a natural explanation of neg-raising.

However, I label this view a mysterian one, as I have little idea how we determine what the belief worlds are if they are not simply those worlds that we consider possible in the usual, epistemic sense.³⁸ In so far as we are just interested in semantics, this may be as far as we need to go, however.³⁹ It is worth noting, though, that this view is not strictly compatible with Lockeanism: merely having a high enough probability in a proposition is compatible with believing, it but it is not a sufficient condition.⁴⁰

³⁶ I thank Diego Feinmann, Manuel Kríž and Matt Mandelkern for pressing this point. Beddor and Goldstein (2017) draw a similar moral in a dynamic context.

³⁷ The embedding behavior of ‘might’ under attitudes is a complex topic, but Mandelkern (forthcoming) provides a framework compatible with the kind of account I am proposing.

³⁸ Some kind of ‘normal’ or ‘ideal’ worlds might be appealed to as in many treatments of modality (e.g., Kratzer 1981).

³⁹ There are significant compositional problems for this account, however, that need to be addressed in an explicit semantics. In particular it’s hard to see how embeddings of operators like ‘probably’ function on this account.

⁴⁰ A natural comparison is Stalnaker’s (1968) analysis of conditionals, which relies on an unanalyzed selection function.

8 Semantics and belief

I hope the data and arguments I've presented here have been illustrative of both the power and the limits of the type of thinking we find in natural language semantics. At the least, I hope I've shown that those theorists who want to argue for a particularly strong notion of belief, must distance themselves from how we speak about believing. I am, however, not at all certain how philosophers should think about belief.

Acknowledgements I am grateful to Julien Dutant, Mark Johnston, Manuel Križ, Harvey Lederman, Sarah Moss, Ian Phillips, Hanna Pickard, and Seth Yalcin as well as audiences at MIT, Princeton, and UCL for discussion. Particular thanks to Diego Feinmann, Angelika Kratzer, Matt Mandelkern, Levi Spectre, and Tim Williamson for extensive comments on earlier drafts. Special thanks to an anonymous reviewer for this journal.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Beddor, B., & Goldstein, S. (2017). Believing epistemic contradictions. *The Review of Symbolic Logic*, *11*, 1–28.
- Brisson, C. (1998). *Distributivity, maximality, and floating quantifiers*. Ph.D. thesis, Rutgers University.
- Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Oxford University Press.
- Clarke, R. (2013). Belief is credence one (in context). *Philosopher's Imprint*, *13*(11), 1–18.
- Dorst, K. (2019). Lockean maximize expected accuracy. *Mind*, *128*, 175–211.
- Foley, R. (1992). The epistemology of belief and the epistemology of degrees of belief. *American Philosophical Quarterly* *29*(2), 111–124. <http://www.jstor.org/stable/20014406>.
- Gajewski, J. (2005). *Neg-raising: Polarity and presupposition*. Ph.D. thesis, MIT.
- Greco, D. (2015). How i learned to stop worrying and love probability 1. *Philosophical Perspectives*, *29*(1), 179–201. <https://doi.org/10.1111/phpe.12059>.
- Hawthorne, J., Rothschild, D., & Spectre, L. (2016). Belief is weak. *Philosophical Studies*, *173*(5), 1393–1404.
- Hintikka, J. (1962). *Knowledge and Belief*. Ithaca: Cornell.
- Horn, L. R., & Bayer, S. (1984). Short-circuited implicature: A negative contribution. *Linguistics and Philosophy* *7*(4), 397–414. <http://www.jstor.org/stable/25001177>.
- Horn, L. (1978). Remarks on neg-raising. In P. Cole (Ed.), *Syntax and semantics 9: Pragmatics*. New York: Academic.
- Horn, L. (1989). *A natural history of negation*. Stanford: CSLI.
- Kratzer, A. (1981). The notional category of modality. In H.-J. Eikmeyer & H. Reiser (Eds.), *Words, worlds, and contexts* (pp. 38–74). Berlin: Walter de Gruyter.
- Križ, M. (2015). *Aspects of homogeneity in the semantics of natural language*. Ph.D. thesis, University of Vienna.
- Kyburg, H. (1961). *Probability and the logic of rational belief*. Middletown: Wesleyan University Press.
- Lasnik, P. (1999). Pragmatic halos. *Language* *75*(3), 522–551. <http://www.jstor.org/stable/417059>.
- Leitgeb, H. (2014). The stability theory of belief. *Philosophical Review*, *123*(2), 131–171.
- Mandelkern, M. (2019). Bounded modality. *Philosophical Review*, *128*(1), 1–61.
- Moss, S. (forthcoming). Full belief and loose speech. *Philosophical Review*. <http://www-personal.umich.edu/~ssmoss/Moss%20-%20Full%20Belief%20and%20Loose%20Speech.pdf>.

- Moss, S. (2017). *Probabilistic knowledge*. Oxford: Oxford University Press.
- Recanati, F. (2013). Pragmatic enrichment. In D. G. Fara & G. Russell (Eds.), *Routledge companion to philosophy of language*. London: Routledge.
- Romoli, J. (2013). A scalar implicature-based approach to neg-raising. *Linguistics and Philosophy*, 36(4), 291–353. <https://doi.org/10.1007/s10988-013-9136-2>.
- Rothschild, D. (2015). Lexical pragmatics and the weakness of belief. Lecture notes. <http://danielrothschild.com/wbp.pdf>.
- Schlenker, P. (2004). Conditionals as definite descriptions. *Research on Language and Computation*, 2, 417–462.
- Schwarzschild, R. (1993). Plurals, presuppositions and the sources of distributivity. *Natural Language Semantics*, 2(3), 201–248. <https://doi.org/10.1007/BF01256743>.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. New York: Blackwell.
- Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98–112). Oxford: Oxford University Press.
- Williamson, T. (2018). The strength of belief, manuscript. <http://media.philosophy.ox.ac.uk/docs/people/williamson/Strengthofbelief.pdf>.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
- Yalcin, S. (2007). Epistemic modals. *Mind*, 116, 983–1026.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.