

Knowledge, Credence, and the Strength of Belief

(version of 31 March, 2020, to appear in Amy Flowerree and Baron Reed (eds.), *Expansive Epistemology: Norms, Action, and the Social World*, London: Routledge)

Timothy Williamson

1. *Outright belief: an initial sketch*

When philosophers discuss norms of belief, they often gloss what they mean by ‘belief’ as *outright* belief, in contrast to a mere *degree* of belief. The standard claim that knowledge entails belief is also typically understood as the claim that knowledge entails *outright* belief. What exactly is outright belief?

The contrast with mere degrees of belief is not altogether helpful, because the phrase ‘degree of belief’ is itself unclear. It can mean more than one thing. However, on the most salient account, degrees of belief are *credences*, which in an ideally rational agent are supposed to obey standard mathematical axioms for probability. Traditionally, credences were defined operationally, in terms of betting behaviour, following Frank Ramsey’s seminal essay of 1926 (Ramsey 1978). Like most operational definitions, this one must be taken with a pinch of salt. Someone who refuses to bet on religious grounds may still have credences in the intended sense, for example, a high credence that betting is wicked. Still, one might take betting behaviour as at least a moderately good guide to credence, for more or less rational agents who are willing to bet.

If the operational definition is even roughly on the right lines, there is a large normative difference between outright belief and any credence. Consider this example:

TEN THOUSAND TICKETS (OPERATIONAL)

Lottie knows that there are 10,000 tickets in a fair lottery and only one will win. She cautiously refrains from forming a belief either way as to whether her ticket will lose. Nevertheless, she knows and believes that its chance of losing is 0.9999; she makes bets on that basis. Thus, by an operational standard, her credence that her ticket will lose is 0.9999. In fact, her ticket wins.

Cautious Lottie is not wrong about anything, or mistaken in any way. Her belief that her ticket’s chance of losing is 0.9999 is true; indeed, it constitutes knowledge. It meets any reasonable normative standard, internal or external. More generally, the situation described does not amount to one in which Lottie is mistaken—about anything. Therefore, her belief

that her ticket's chance of losing is 0.9999 does not amount to a belief that her ticket *will* lose, for that belief would be mistaken: it would be false. Thus Lottie lacks the belief that her ticket will lose.

The relevant sense of 'chance' is objective. Lottie's credences are guided by her knowledge of objective chance in conformity with David Lewis's Principal Principle (1980). Although objective chances other than 0 and 1 are often thought to require indeterminism, that will not be assumed here; some kinds of objective probability in physics are independent of determinism (Loewer 2001, Maudlin 2007, Ismael 2009). But if you doubt that chancy statements about the future have truth-values, you can suppose that the draw has already taken place, although the results have not yet been announced; that will not change Lottie's credences. Alternatively, the example can be reworked to concern Lottie's beliefs about probabilities on her evidence (Williamson 2000); although such probabilities do not coincide with the agent's credences by definition, they may coincide in Lottie's case. The argument is robust with respect to such details.

The moral is *not* that outright belief should be equated with credence 1. The crucial difference is not between credence 1 and credence less than 1. Consider this example:

INFINITELY MANY TOSSES (OPERATIONAL)

Indira knows that there will be an ω -sequence (ordered like the natural numbers) of independent tosses of a fair coin. She cautiously refrains from forming a belief either way as to whether tails will come up at least once. Nevertheless, she knows and believes that the chance of tails coming up at least once is 1; she makes bets on that basis. Thus, by an operational standard, her credence that tails will come up at least once is 1. In fact, heads comes up every time.

Even more cautious Indira is not wrong about anything, or mistaken in any way. Her belief that the chance of tails coming up at least once is 1 is true; indeed, it constitutes knowledge. It meets any reasonable normative standard, internal or external. More generally, the situation described does not amount to one in which Indira is mistaken—about anything. Therefore, her belief that the chance of tails coming up at least once is 1 does not amount to a belief that tails *will* come up at least once, for that belief would be mistaken: it would be false. Thus Indira lacks the belief that tails will come up at least once.

Similar comments about the nature of the probabilities apply to Indira's case as to Lottie's. Even if probabilities are allowed to be non-standard real numbers, which can differ infinitesimally from 1, there are still reasons why the probability that tails will come up at least once must be exactly 1; probability 1 does not amount to certainty (Williamson 2007).

Not even the proposition that the event has chance 1 of occurring entails that it *will* occur. The two propositions can differ in both truth-value and epistemic status. Consequently, not even credence 1 (by operational standards) that the coin will come up tails at least once amounts to outright belief that it will come up tails at least once.

Of course, an infinite sequence of coin tosses is a very distant possibility. Nevertheless, it is useful to consider that possibility because it forces us to acknowledge key structural distinctions, between certainty and probability 1, and between impossibility and probability 0. Those distinctions arise for finite as well as infinite probability spaces, since the probability axioms are consistent with the assignment of probability 0 to a proper subspace. However, in the finite case one can stipulate that the space is *regular*, in the sense that each world in the space has nonzero probability (which can be constant across all worlds). The point of INFINITELY MANY TOSSES is that in the infinite case no such fix is feasible. But regularity is not preserved under updating by standard Bayesian conditionalization; when a regular probability distribution is so updated, the result is either no change or a non-regular probability distribution. In that sense, the stipulation of regularity is unstable. Thus, even in the finite case, the distinction between the empty event and a non-empty event of probability 0 is hard to escape.

Lottie and Indira are not just theoretical possibilities. There are sometimes good practical reasons to follow their example. For instance, a member of a jury may update their probability that the accused is guilty on the incoming strong evidence, while holding off the urge to form an outright belief in his guilt, because she wants to keep an open mind until she has heard all the evidence and the lawyers' closing arguments.

2. *Interlude on probability talk*

The arguments of §1 require that we can have beliefs *about* the probability of an event that do not reduce to degrees of belief as to whether the event occurs, and that we can use such beliefs to guide our behaviour. It does not really require that we can *express* such propositions about probability in a natural language using standard probability operators with their standard meanings (on which see Yalcin 2010, 2012 and Moss 2018). Although the descriptions of the two examples used ordinary English to express propositions about objective or evidential probabilities, expressing them in an artificial notation, such as mathematicians use, would have sufficed instead. But since §1 ascribed beliefs about probability in English, some remarks about probability talk in natural language may be appropriate.

The word 'probably' is not always used to express propositions about probabilities. It has a parenthetical use to modify the force rather than the content of an utterance. Like 'perhaps', it reduces the level of commitment, but less so:

- (1) Probably she is in Spain.
- (2) Perhaps she is in Spain.

One sign that ‘probably’ and ‘perhaps’ are not propositional operators in (1) and (2) is that applying a sentential negation to them is infelicitous:

(3)? It is not the case that probably she is in Spain.

(4)# It is not the case that perhaps she is in Spain.

Of course, we can move ‘probably’ and ‘perhaps’ outside ‘that’ and insert suitable pauses:

(3a) It is not the case, probably, that she is in Spain.

(4a) It is not the case, perhaps, that she is in Spain.

But (3a) and (4a) sound better than (3) and (4) respectively because we hear ‘perhaps’ and ‘probably’ in (3a) and (4a) as qualifying the whole statement, rather than the negated content.

Admittedly, (4) sounds even worse than (3) does, for (4) lacks an easy repair corresponding to the easy repair of (3) as (5) or (6):

(5) It is not probable that she is in Spain.

(6) It is not the case that it is probable that she is in Spain.

The contrast indicates that ‘it is probable that’ is not parenthetical in (7), of which (5) and (6) are felicitous negations (although (6) is of course unnecessarily long-winded):

(7) It is probable that she is in Spain.

Whatever (7) communicates, it does so by making a non-tentative statement about the probability of her being in Spain.

‘Improbably’ and ‘improbable’ pattern in related ways. Thus (8) is fine; it too makes a non-tentative statement about the probability of her being in Spain; it is contrary to (7) and pretty much equivalent to (9):

(8) It is improbable that she is in Spain.

(9) It is probable that she is not in Spain.

By contrast, the relationship between (10) and (11) is quite unlike that between (8) and (9):

(10) Improbably she is in Spain.

(11) Probably she is not in Spain.

The latter two are far from equivalent. The natural use of (10) is to say non-tentatively that she is in Spain, while commenting that in advance one would have expected otherwise; (10) is inconsistent with (11). 'Improbably' is parenthetical in (10), like 'probably' in (1) and (11), but since it would be self-defeating to state something (when not merely guessing) while deprecating its present probability, the comment must be reinterpreted in terms of past probabilities.

Speakers' interpretation of (10) in terms of past probabilities is *ad hoc*. By contrast, the verb in the sentential operator 'it is probable that' is tensed in the normal way, unlike 'probably'. Its past tense is unproblematic, as in (12):

(12) It was probable yesterday that she would be in Spain today.

That clearly makes a statement about yesterday's probabilities; the past tense probability operator is being used to describe probabilities, not to reduce the speaker's level of commitment. If such a descriptive compositional reading is available for the past tense operator 'it was probable that', a corresponding descriptive compositional reading is also available for the present tense operator 'it is probable that', whether or not a non-descriptive reading is available too.

'Probably' can also occur within a predicate, as in (13), more fully integrated into the syntax than when it prefixes a sentence:

(13) She is probably in Spain.

It is not obvious whether to read (13) like (1), with 'probably' parenthetical, or like (7), with 'probably' a propositional operator. Embedding a sentence in a negative context, such as the antecedent of a conditional, typically requires the operator reading, as in (14):

(14) If she is probably in Spain, she is probably on vacation.

Thus (14) is read like (15), because the parenthetical reading of 'probably' in the antecedent of (14) makes little sense:

(15) If it is probable that she is in Spain, she is probably on vacation.

By contrast, it is much less clear whether to give 'probably' in the consequent of (14) and (15) a parenthetical or operator reading. Both make sense. Even (16) has non-equivalent readings:

(16) If she is probably in Spain, she is probably in Spain.

When both occurrences of 'probably' have the operator reading, (16) is a trivial tautology (given no gratuitous semantic variation in 'she', 'is', and so on). But when only the first occurrence has the operator reading, while the second has the parenthetical reading, the speaker is making a tentative risky statement: if it is probable but false that she is in Spain, the statement tentatively made by (16) is then itself false.

We can make the antecedent of a conditional the main clause by using 'only if' instead of plain 'if', which makes it easier to hear 'probably' as parenthetical, but simply because we can hear it as qualifying the whole statement:

(14a) She is probably in Spain only if she is on vacation.

Does 'probably' also have a *third* reading, as some philosophers have thought, more fully integrated into the compositional semantics than the first reading, but in contrast to the second reading more expressive than descriptive of current probabilities? It would need to be shown that the appearance of such a reading is anything more than an artefact of indiscriminately lumping together data involving parenthetical uses with data involving operator uses. This is not the place to pursue that issue further. For present purposes, suffice it to stipulate that English operators such as 'It is probable that' will henceforth be used to express propositions about objective or evidential probability.

3. *Outright belief and knowledge*

Although operational credence 1 is insufficient for outright belief, it does not follow that outright belief requires a perfectionist standard of certainty. A more natural explanation of the difference is that outright belief in a proposition p involves a disposition to *rely* on p , more specifically, to act on p by using p as a premise in practical reasoning.

Dispositions to rely on p come in degrees. How willing one is to rely on p may depend inversely on how much is at stake: the greater the stakes, the less the willingness. There may be few propositions, if any, on which one is willing to rely no matter how high the stakes. But the degree to which one is disposed to rely on p does *not* correspond to one's credence in p , as normally understood. For example, Lottie treats the proposition that her ticket will lose as very highly probable, but she may be fully disposed *not* to rely on that proposition.

Thus Lottie fails to believe outright that her ticket will lose because she is not disposed to use the proposition that it will lose as a premise in her practical reasoning. Rather, the premise she uses in her practical reasoning is that its chance of losing is 0.9999. That is what she relies on, and believes outright. In so doing, she does not take the proposition that her ticket will lose to meet some unrealistically high epistemic standard.

Clearly, as premises in practical reasoning, the proposition that the ticket will lose and the proposition that the chance of its losing is 0.9999 have quite distinct logical powers.

Similarly, Indira fails to believe outright that tails will come up at least once because she is not disposed to use the proposition that tails will come up at least once as a premise in her practical reasoning. Rather, the premise she uses in her practical reasoning is that the chance of tails coming up at least once is 1. That is what she relies on, and believes outright. In so doing, she does not take the proposition that tails will come up at least once to meet some perfectionist epistemic standard of absolute certainty. As premises in practical reasoning, the proposition that tails will come up at least once and the proposition that the chance of tails coming up at least once is 1 have quite distinct logical powers.

This account fits the knowledge-first view that to believe p is to be disposed to treat p as if one knew p —that is, to be disposed to treat p as agents treat propositions they know (Williamson 2000: 46-47, 2017; see Wimmer 2019 for complications). Acting on what one knows is the paradigm of intelligent life; that one is also liable to act on what one merely believes is a corollary of that paradigm. In acting on p , one so far treats p as if one knew p . Thus reliance is a good candidate for the main way of treating a proposition in common between knowing and merely believing. In particular, this view suggests, knowledge entails belief: if one knows p , one is disposed to treat p as if one knew p : the ‘as if’ does not imply that one lacks knowledge, still less that one has a mental reservation to that effect (see section 7 below for more on the connection between knowledge and belief).

Lottie may indeed refrain from acting on the proposition that her ticket will lose precisely because it is clear to her that she does not *know* that it will lose (compare Hawthorne and Stanley 2008). She is willing to act on the proposition that its chance of losing was 0.9999 because it is clear to her that she knows that its chance of losing was 0.9999, or at least it is not clear to her that she does not know (depending on the details of the example). Similarly, Indira may refrain from acting on the proposition that tails will come up at least once precisely because it is clear to her that she does not know that tails will come up at least once. She is willing to act on the proposition that the chance of tails coming up at least once is 1 because it is clear to her that she does know that the chance of tails coming up at least once is 1, or at least it is not clear to her that she does not know. Such a connection between knowledge and belief imposes no perfectionist standard of certainty, unless one has already imposed such a standard on knowledge, which was not done here.

This talk of practical reasoning should not be taken to imply an over-intellectualized, over-sophisticated conception of knowledge and belief. The reasoning need not be deductive, nor need it be articulated in words. Even languageless animals and very small children can integrate information, and misinformation, from various sources to determine what to do. They can act on what they know, and they can act on what they treat as if they knew it when in fact they merely believe it.

Of course, we sometimes treat a proposition as a working assumption, and integrate it into our practical reasoning, without believing it. We may even know that it is false. For

example, engineers rely on Newtonian rather than relativistic mechanics in their calculations, because it is much simpler and a good enough approximation for their purposes. They do not thereby believe Newtonian mechanics; they know that it is strictly speaking false. We might describe them as believing Newtonian mechanics ‘for practical purposes’, but that qualification is needed. They do not fully treat Newtonian mechanics as if they knew it, because their practical reliance on it is mediated by their belief that it is a good enough approximation. After all, if you relied on an assumption p for practical purposes because your evidence showed p to be a good enough approximation, although you held p to be strictly speaking false, and then p turned out to be true, you could not claim to have known p all along. Your arms’-length treatment of p is not the embrace of knowledge.

More generally, you may rely on p on a particular occasion—for example, when very little is at stake—without being sufficiently *disposed* to rely on p to count as believing p , at least by the standards of a given context.

4. *Non-operational credences*

What happens when credences are understood *non-operationally*? Could a high enough credence constitute belief? The proposal might be that credences are psychologically real, natural propositional attitudes with unique contents, so no credence in a proposition about the probability of p would constitute a credence in p , though it might constitute a *belief* in the proposition about the probability of p .

The earlier argument can be generalized to non-operationalized credences. Although an agent’s dispositions to betting behaviour and the like may not *constitute* her credences, in favourable circumstances nothing need prevent her from *conforming* her credences in the relevant propositions to her knowledge of their objective or evidential probabilities, after the manner of Lewis’s Principal Principle. Thus her credences will be probabilistically coherent, and determine the same behaviour as her knowledge of the probabilities. We can refine the two examples accordingly:

TEN THOUSAND TICKETS (NON-OPERATIONAL)

Everything is as in TEN THOUSAND TICKETS (OPERATIONAL), but in addition Lottie conforms her credences to her knowledge of the chances. Thus, by a non-operational standard, her credence that her ticket will lose is 0.9999.

Thus Lottie still has credence 0.9999 that her ticket will lose without believing outright that it will lose. By normal standards, her position may be cautious but is not irrational.

INFINITELY MANY TOSSES (NON-OPERATIONAL)

Everything is as in INFINITELY MANY TOSSES (OPERATIONAL), but in addition Indira conforms her credences to her knowledge of the chances. Thus, by a non-operational standard, her credence that tails will come up at least once is 1.

Thus Indira still has credence 1 that tails will come up at least once without believing outright that tails will come up at least once. By normal standards, her position may be cautious but is not irrational.

Could it be argued that understanding credences non-operationally wipes out the distinction between outright belief and high credence, or at least the possibility of credence 1 without outright belief? The thought would be that to assign an outcome credence 0 just *is* to exclude it doxastically. But that does not work in INFINITELY MANY TOSSES, since Indira has credence 1, but not outright belief, in the proposition that tails will come up at least once. There is no clear reason to dismiss the case as psychologically impossible. Whether or not Indira is rational, if the case is possible, credence 1 is psychologically insufficient for outright belief.

Indira's position reflects a fundamental distinction in mathematical probability theory between probability 0 and impossibility. A probability space is based on a set of *outcomes*, conceived as mutually exclusive and jointly exhaustive possibilities; they are the analogues of possible worlds in a model for modal logic. Probabilities are assigned to *events*, identified with sets of outcomes, which are analogous to propositions, identified with sets of worlds. Each probability space has a natural modality: the only necessary event is the set of all outcomes; the only impossible event is the empty set; every other event is contingent.¹ Typically, in an infinite probability space, many possible events (nonempty sets of outcomes) have probability 0, for combinatorial reasons and perhaps others. No way of understanding credences can wipe out that structural distinction. It would be hopeless to treat the all heads event as impossible (that is, empty) on the grounds that it has probability 0: the argument for assigning it probability 0 generalizes to any other maximally specific event, any other ω -sequence of heads and tails, so by parity of reasoning one would have to treat *every* ω -sequence of heads and tails as impossible (empty), and so the original scenario as impossible, which it is not.

A similar distinction applies to practical reasoning, for example about what bets to take. Think of it as concerning a range of *live possibilities*. Any proposition incompatible with a premise of the reasoning is thereby excluded from the range of live possibilities (even if it is in fact true). If Indira were to exclude every proposition with probability 0 from the range of possible outcomes, she would exclude each outcome of the sequence of tosses, and so to exclude the sequence of tosses itself as impossible. That would be absurd, for she knows that one or other of them will occur. Thus she needs a nonempty category of live possibilities each with probability 0. To avoid assigning invidious privileges to some specific outcomes of the coin-tossing over others, we may assume, she assigns all of them to that category. Indira treats whatever she outright believes as eligible to figure as a premise in her practical reasoning. Thus, for any specific possible outcome of the coin-tossing, she avoids

outright belief that it will not happen, since that would mean excluding that outcome as a live possibility. Consequently, she avoids outright belief in some propositions to which she assigns probability 1.

Whether we understand credences operationally or non-operationally, the examples suggest that they cannot add up to outright belief.

Sometimes, a purely normative hypothesis is proposed, the so-called Lockean thesis that outright belief in p is rational if and only if high enough credence in p is rational (Foley 2009, Leitgeb 2017). That does not automatically require outright belief in p to *be* high enough credence in p . The case of Indira indicates that outright belief and the highest degree of credence can come apart even in a rational agent. One might try to reconcile that point with the Lockean thesis by adopting a permissive view of the rationality of outright belief, on which, if credence 1 in p is rational, so is outright belief in p , but refraining from outright belief in p can be rational too, as with Indira. Indeed, rational agents willing to use their outright beliefs as premises for reasoning in the way explained will get into trouble with the coin-tossing case if they insist on outright believing everything to which they assign probability 1.

For a different take on such cases, consider a knowledge norm for outright belief, corresponding to the imperative 'Believe p only if you know p '. Given that norm, it would arguably be irrational for someone in Indira's position to believe outright that tails will come up at least once, since she knows that she does not know that tails will come up at least once, even though it is rational for her to have credence 1 that tails will come up at least once. On that view, the Lockean thesis fails. Of course, the knowledge norm for outright belief is highly controversial; we will return to it in section 8.

5. *Sceptical interlude on credences*

Scepticism about credences initially seems unreasonable, since we can readily make sense of questions about which of two propositions we are more confident of. However, we should not assume without argument that one's degree of confidence in p just is one's credence in p , in the technical sense of 'credence'. The normal use of the word 'confident' does not fit the ideology of credences. For example, imagine that a race is about to start. I ask you:

(17) Who are you most confident will win?

The question has a tendentious presupposition. You might reject it by saying:

(18) I think that number 7 has the best chance, but I have no confidence that he will win.

'No confidence' is quite different from 'no chance', even on an epistemic reading of 'chance'. One cannot properly explain 'credence' in terms of 'confidence' if one has first projected the ideology of credences onto one's use of 'confidence'. One's confidence in p might instead correspond to the degree to which one is disposed to rely on p , to use it as a premise in practical reasoning (Williamson 2000: 99). There might also be negative degrees of confidence in p , corresponding to the degree to which one is disposed to rely on its negation, to use $\neg p$ as a premise in practical reasoning. For a large body of intermediate propositions, of varying degrees of specificity, one may have *no* disposition to rely on them in practical reasoning, and *no* disposition to rely on their negations.

A good question is whether anything like the philosopher's credences is psychologically real and natural for humans. That has been doubted (Holton 2008, 2014; Weisberg 2020).

The friend of credences may respond that postulating psychologically real and natural credences is the best way to understand rational behaviour. Even if they are not *defined* operationally, they still *manifest* themselves in betting and other behaviour, in ways not plausibly explained by beliefs about probabilities. Moreover, on pain of irrationality, these credences must conform to the mathematical laws of probability, otherwise the agent will be subject to a Dutch book. But it would be simple-minded to assume that the agent's pattern of responses to a long series of artificial tasks, such as deciding which of numerous weird bets to accept, must reveal structures present all along, as opposed to constructing artefacts. At the very least, one would want evidence that the credences so 'revealed' were approximately independent of the order in which the bets were presented. Nor should one assume that there are strong evolutionary pressures to have a coherent pattern of credences over large numbers of propositions. Nature and culture are full of dangers, but clever bookies trying to turn us into money pumps are not usually the problem. Bayesian thinking makes heavy demands on our limited resources of memory and computation. In ecologically realistic conditions, its benefits may be too slight to be worth the additional costs in time and energy, when compared with fast and frugal heuristics (see for example Martignon and Laskey 1999).

This scepticism about credences is quite consistent with a key role for unconscious probabilistic Bayesian processes governing various aspects of perception and motor behaviour in humans and other animals. For a module that deals only with a restricted range of possibilities, such processes may be highly efficient. But that efficiency cannot be assumed to generalize to conscious beliefs, especially to all those which can be verbalized with the full expressive power of a natural language. Consider an elementary step of reasoning, for example by disjunctive syllogism: if one acquires the belief 'A or B' from one source, and the belief 'Not A' from another, one goes on to form the belief 'B'. There is no fully analogous step for credences, since it would require getting from the probabilities of 'A or B' and 'Not A' to that of 'B': but the probability of 'B' is not a function of the probabilities of 'A or B' and 'Not A'. If the probabilities of 'A or B' and 'Not A' are both 80%, the

probability of 'B' can be anywhere between 60% and 80%. It is far from clear what reasoning with credences in place of outright beliefs is supposed to be.

Obviously, the psychological reality or unreality of credences is a large issue, which cannot be settled here. Instead, we turn to a different sort of challenge to outright belief.

6. *A challenge from natural language semantics*

So far, the English word 'believe' has been treated as expressing outright belief. That has produced no obvious dissonance. Of course, semantic questions about the meaning of the English verb 'believe', and the even more common verb 'think', are not equivalent to psychological and philosophical questions about the reality and naturalness of the sort of state philosophers intend by 'outright belief'. Nevertheless, if ordinary speakers continually talk about outright belief, that boosts its credentials, at least by rebutting the suspicion that it is just a philosophers' invention.

In a recent article, John Hawthorne, Daniel Rothschild, and Levi Spectre (2016), henceforth HRS, present evidence against the assumption that the role of 'believe' (and 'think') is to express the sort of state philosophers intend by 'outright belief' (see also Holguín 2020). Rather, they argue, it expresses a much weaker sort of doxastic attitude. For example, they point out, one can quite naturally utter (1), without making either Mary or oneself sound at all irrational:²

(19) Mary believes it's raining, but she's not sure it's raining.

In this case, 'be sure' seems much closer than 'believe' to expressing what was intended by 'outright belief'.

HRS also note that 'believe' and 'think' are subject to neg-raising, where an occurrence of negation syntactically operating on the verb is interpreted as though it operated on its complement. For instance, (20a) is naturally interpreted like (20b), and (21a) like (21b):

(20a) Mary does not believe it will rain.

(20b) Mary believes it will not rain.

(21a) Mary does not think it will rain.

(21b) Mary thinks it will not rain.

If Mary is agnostic about whether it will rain, one should not use (20a) or (21a). By contrast, as (19) already suggests, stronger doxastic terms do not exhibit neg-raising. For example, (22a) is not naturally interpreted like (22b):

(22a) Mary is not sure it will rain.

(22b) Mary is sure it will not rain.

Neg-raising is characteristic of weak verbs. For instance, ‘want’ exhibits neg-raising, while ‘need’ does not. Thus (23a) is naturally interpreted like (23b), while (24a) is not naturally interpreted like (24b):

(23a) John does not want to exercise.

(23b) John wants to not exercise.

(24a) John does not need to exercise.

(24b) John needs to not exercise.

One might initially suppose that neg-raising holds for a weak operator O because, if O is weak enough, it commutes with negation in the sense that $O\neg p$ is logically or necessarily equivalent to $\neg Op$. However, as HRS are aware, that explanation fails. For, with respect to any of the standard weak attitudes, one can be simply *indifferent* between two contradictory propositions, so that Op holds if and only if $O\neg p$ holds. But, if O commutes with \neg , that means that Op holds if and only if $\neg Op$ holds. That is a contradiction, at least in classical logic, which HRS are not challenging. Since such states of indifference or agnosticism are quite common, natural, and salient, it seems unlikely that the semantics of natural language has somehow just overlooked their possibility. Thus the point about neg-raising is only circumstantial evidence, as it were, for the weakness of belief.³

In any case, one might expect believing p in the weak sense to be a matter of treating p as more likely than not, giving p a credence over 50%, but HRS argue that it is even weaker than that. For example, at the beginning of the season, if asked ‘Which team do you believe will win the league?’, I may answer by naming the team I take to be most likely to win, even if I think that the combined probabilities of the other teams winning add up to more than 50%. Thus I take the team I named to be less likely than not to win. 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.⁴

Can we interpret the answer as merely claiming that the specified team has the best chance of winning? That move is *ad hoc*. It involves either a non-literal interpretation of the question, or a failure to answer the question. Why think that anything so indirect is involved in such a smooth and commonplace conversational exchange? If the hearer interpreted the

question non-literally, what is supposed to happen when it is interpreted literally? If the hearer interpreted the question literally, why did they answer a different question, and why did they not sound evasive in doing so?

HRS argue against the fallback position that ‘believe’ is *ambiguous* between strong and weak senses, outright belief and the watered-down version. For if it were, one should be able to hear (25) as correctly attributing weak belief while denying strong belief (with the cumbersome negation ‘it’s not that’ to resist neg-raising):

(25)? He thinks Spurs will win, but it’s not that he believes they will.

No such reading is straightforwardly available; (25) stubbornly sounds like a contradiction. By putting intense emphasis on ‘believes’ in (25), a speaker might be able to get the hearer to concoct a reading on which (25) is correct, but that seems more like constructing a sense for ‘believes’ on the spot to fit the speaker’s importunate intonation than just selecting the appropriate lexical entry from a standing menu. Summing up, HRS say that outright belief ‘is not a disambiguation of what we ordinarily mean by “belief”; rather it seems a theoretical posit. [...] The everyday notion of belief is a weak one’ (2016: 1402).

An alternative to ambiguity is *context-dependence*. Given that only one lexical entry for ‘believe’ is at issue, it does not follow that ‘believe’ expresses the same relation in all contexts of utterance. The adjective ‘tall’ expresses different properties as the contextually relevant comparison class varies, with no corresponding multiplicity of lexical entries. Indeed, HRS’s ‘rough preliminary account’ involves two context-dependent elements: ‘to believe something might require (a) it be significantly more likely than the salient alternatives, and (b) it be above some contextually determined threshold of likeliness’ (2016: 1400; presumably, ‘likely’ in (a) and (b) is to be evaluated with respect to the putative believer’s subjective probabilities, her credences). Which alternatives are salient will also depend on the conversational context, in particular on what questions are at issue. Thus both (a) and (b) are context-dependent conditions.

If we provisionally treat the conjunction of (a) and (b) as necessary and sufficient for belief, then Indira counts as believing that tails will come up at least once. In this context, the only salient alternative is its negation, in effect that heads will come up every time, which is less likely than that tails will come up at least once ($0 < 1$), so (a) is satisfied. Moreover, since the possibility that tails will come up at least once has subjective probability 1 for Indira, it is above any threshold of likeliness except for one which would absurdly count her as not believing trivial tautologies, so (b) is satisfied too. Thus the analysis counts Indira as believing that tails will come up at least once. Or perhaps we should say instead that it counts the description of INFINITELY MANY TOSSES as incoherent, for it included the stipulation that Indira ‘cautiously refrains from forming a belief either way as to whether tails will come up at least once’.

That consequence of the analysis is implausible, for the state of mind attributed to Indira is readily intelligible. She assigns the all-heads outcome credence 0, but she does not exclude it doxastically. It is not doxastically impossible for her. She treats any other maximally specific outcome the same way. She was described as cautiously refraining from forming a belief either way (with no modifier ‘outright’); that description seems apt for the

envisaged state of mind. Thus INFINITELY MANY TOSSES is arguably a counterexample to the conjunction of (a) and (b) as an analysis of belief.

Of course, HRS presented (a) and (b) as only a 'rough preliminary account', so the existence of counterexamples need not surprise them. But they are more confident of their general view: 'we have seen that thinking or believing *p* is likely is sufficient for thinking *p*' (2016: 1401; they treat thinking *p* and believing *p* as equivalent). INFINITELY MANY TOSSES is also a threat to any other way of sharpening that general view, for on any reasonable sharpening Indira believes it likely that tails will come up at least once.

On their side, HRS can cite examples such as (26) and (27):

(26) John thinks it's likely Levi is in Sweden but he doesn't go so far as to think Levi is in Sweden.

(27) John believes it's likely Levi is in Sweden but he doesn't go so far as to think Levi is in Sweden.

With flat intonation, or with emphasis on 'Levi' or 'Sweden', (26) and (27) certainly sound odd. But so does (28):

(28) Levi went to Sweden but he didn't fly to Sweden.

A speaker who uses two similar expressions to draw a contrast is normally expected to use emphasis to draw the hearer's attention to the point of contrast. Thus the normal way to pronounce (28) would be something like (28+):

(28+) Levi *went* to Sweden but he didn't *fly* to Sweden.

Not to use such emphasis is uncooperative, perhaps even socio-linguistically incompetent. The flat intonation sounds robotic. Exactly the same applies to (26) and (27). The normal way to pronounce them would be something like (26+) and (27+):

(26+) John thinks it's *likely* Levi is in Sweden but he doesn't go so far as to think Levi *is* in Sweden.

(27+) John believes it's *likely* Levi is in Sweden but he doesn't go so far as to think Levi *is* in Sweden.

To my ear, these sound much better than (26) and (27). Of course, emphasis sometimes distorts examples, but to refuse emphasis when it is independently expected is equally distorting.⁵

None of this implies that 'believe' always expresses outright belief. It does not undermine HRS's evidence that in many contexts it expresses something much weaker. Rather, it suggests that in *some* contexts, perhaps in many, it expresses a more categorical state of outright belief.

That conclusion is hardly surprising. Imagine a preacher telling his congregation:

(29) You've got to believe that Jesus can save you.

He is not just telling them that they've got to believe it likely that Jesus can save them. He is telling them to exclude the possibility that Jesus can't save them. A quick Google search on the words 'I believe that' turns up numerous examples in which the insertion of 'it likely' between 'believe' and 'that' has a comic effect, undermining a solemn creed with an implication of uncertainty.

Might the preacher be saying literally that his congregation have got to believe in the weak sense that Jesus can save them, but be naturally interpreted in context as communicating to them the proposition that they have got to be confident in a much stronger sense that Jesus can save them? The difficulty is to understand where such a strengthened content would come from in this case. After all, if the contrast between 'believe' and 'be sure' is operative, then one would expect the preacher's choice of the word 'believe' over 'be sure' to be communicatively significant, to generate the conversational implicature that he is not in a position to assert the equally relevant and supposedly stronger (30):

(30) You've got to be sure that Jesus can save you.

But no such implicature is generated. The preacher is not hinting that his congregation should not be too sure that Jesus can save them. That is evidence that the contrast between 'believe' and 'be sure' is *not* operative in that context.

The history of analytic epistemology since 1963 suggests a similar conclusion. Gettier put the 'B' in the JTB analysis of knowledge. His second conjunct for the analysis was 'S believes that P', whereas of the two sources he cites for it, Ayer (1956: 34) had 'is sure' and Chisholm (1957: 16) had 'accepts'. After Gettier, 'believes' was the canonical term in that component of attempted analyses of knowledge. Of course, he was far from the first to use 'believes' as such a foil for 'knows'; Russell (1912, chapter 13) discussed the relation between 'knowledge' and 'true belief'. But 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.⁶

Can one understand the epistemological literature by supposing that, although the sentences used literally express only propositions about belief in the weak sense, they are naturally interpreted in context as communicating propositions about confidence in a much stronger sense? As in the case of the preacher, the difficulty is to understand where such a

strengthened content would come from. After all, when the contrast between ‘believe’ and ‘be sure’ is operative, then one would expect the choice of ‘believe’ over ‘be sure’ to be communicatively significant, to generate conversational implicatures of the unassertibility of the equally relevant and supposedly stronger statements with ‘be sure’ in place of ‘believe’. The fact that no such implicature was generated is evidence that the contrast between ‘believe’ and ‘be sure’ was *not* operative in those contexts.

A more natural explanation for epistemologists’ comfort with ‘believe’ is that it is easy to have a context in which ‘believe’ expresses something like the state philosophers have taken it to express, outright belief. HRS’s suggestion that outright belief is a mere philosophers’ ‘theoretical posit’ is unconvincing.

One might wonder: why is it so hard to hear sentences like (25) as correctly attributing weak belief (expressed by ‘think’) while correctly denying outright belief (expressed by ‘believe’)? A similar question arises for HRS themselves: why is it so hard to hear those sentences as correctly attributing belief relative to a very low standard of likelihood (expressed by ‘think’) while correctly denying belief relative to a middling standard of likelihood (expressed by ‘believe’)? The answer is the same. Without special stage-setting, the relevant standard is set by general features of the conversational context, not by features specific to one word rather than the other. This is a common feature of pairs of context-dependent terms with similar meanings. For instance, ‘high’ and ‘tall’ are more or less interchangeable as applied to buildings, although the contextually relevant standard may vary, for example depending on whether towers in San Gimignano or towers in New York form the comparison class. But, without elaborate stage-setting, a sentence such as (31) is baffling:

(31) That is a tall tower but not a high tower.

It is very hard to hear (31) as both correctly attributing height relative to a low standard (expressed by ‘tall’) and correctly denying height relative to a high standard (expressed by ‘high’). Again, even though the relevant standard for ‘like’ and ‘similar’ varies with conversational context, it is very hard to hear (32) as both correctly attributing similarity relative to a low standard (expressed by ‘like’) and correctly denying similarity relative to a high standard (expressed by ‘similar’):

(32) They are like each other but not similar to each other.

The problem with (25) can be explained along the same lines.

By contrast with ‘think’ and ‘believe’, terms such as ‘be sure’ and ‘be certain’ are restricted to the high end of the spectrum. Thus there is no problem in accommodating to examples where ‘believes’ is contrasted with ‘is sure’, such as HRS’s (19). Since the threshold for ‘believe’ can go much lower than that for ‘sure’, but cannot go higher, the two terms form a naturally contrasting pair. The utterance of (19) typically creates a low contextual standard for ‘believe’, as required to verify (19). One might compare (19) to unproblematic examples such as (33):

(33) That is large but not enormous.

One may further conjecture that ‘think’ and ‘believe’ are normally mandatorily maximally inclusive, in the sense that no other term can stand to them as they stand to ‘be sure’ and ‘be certain’. For example, the attempt to cast ‘suspect’ in that role fails, since (34c) sounds so much worse than (34a) or (34b) (Rothschild 2019):

(34a) Mary *thinks* it will rain, but John is *sure* it will.

(34b) Mary *suspects* it will rain, but John is *sure* it will.

(34c)? Mary *suspects* it will rain, but John *thinks* it will.

Although ‘grasps the thought’ has a reading on which it is more inclusive than ‘thinks’, it is so only by including cases where the thinker’s attitude to the content is neutral or negative, whereas the inclusiveness which ‘think’ and ‘believe’ may maximize is only amongst terms—such as ‘suspect’—which imply that the thinker’s attitude to the content is to some degree positive. Mandatory maximal inclusiveness is only a comparative feature; it just forbids contexts in which some other term from the relevant range is more inclusive than ‘think’ or ‘believe’. That is quite consistent with ‘think’ and ‘believe’ setting a moderately strong standard when (as often) they are not competing with other such terms.

There is much more to be said about the semantics of ‘think’ and ‘believe’ (as emerges from Rothschild 2020). One clue that the present approach is on the right lines is the common form ‘fully believe’. Whereas ‘believe outright’ is mainly philosophers’ English, ‘fully believe’ is quite standard, as a Google search confirms. By contrast with the failed contrast between ‘think’ and ‘believe’ in (25), the insertion of ‘fully’ restores felicity:

(35) He thinks Spurs will win, but it’s not that he fully believes they will win.

Significantly, ‘fully believe’ is not subject to neg-raising: there is no inclination to treat (36) as equivalent to (37):

(36) Mary does not fully believe that it will rain.

(37) Mary fully believes that it will not rain.

Instead of (35), we might also say:

(38) He sort of believes Spurs will win, but he doesn’t fully believe they will.

The semantic effect of ‘fully’ in ‘fully believe’ is similar to its effect with other verbs, as in ‘fully participate’, ‘fully enjoy’, ‘fully understand’, ‘fully accept’, ‘fully comply’, and ‘fully execute’ (a contract), which suggests that ‘fully believe’ has a compositionally determined meaning and is no mere idiom.

By contrast with ‘fully believe’, ‘fully think’ is infelicitous, though a Google search shows it to have been used by Donald Trump and a few philosophers. This suggests that ‘think’ and ‘believe’ are not quite as similar in meaning as they are often taken to be. However, ‘totally believe’ and ‘totally think’ are both quite common, though colloquial and recent.

Examples such as (38) indicate that we do associate ‘believe’ with a rough scale. The top end does not demand an outrageously Cartesian degree of certainty, for sentences such as (38) can be felicitously uttered in everyday contexts:

(39) He fully believes that he passed.

Perhaps (39) can be paraphrased by (40) in a suitable context:

(40) He is sure that he passed.

For reasons explained in §5, we should avoid assuming that the bottom of the scale corresponds to credence 0: (41) seems false of someone who believes that the chance that she passed is one in a million:

(41) She believes to some degree that she passed.

A reasonable hypothesis is that ‘believe’ stands for full belief in some contexts, such as those reviewed above, in which the contrast between ‘believe’ and ‘be sure’ is not operative.

Might ‘believe’ stand for full belief in *all* contexts, with appearances to the contrary being explained away as loose speech (Lasersohn 1999), as Sarah Moss and others have in effect proposed? That view is challenged by felicitous examples such as (42):

(42) She believes that she passed, but she’s not sure that she did.

The proposed view makes (42) similar in meaning to (43), with the antecedent understood as loose talk:

(43) She believes that she passed, but she doesn’t exactly believe that she passed.

However, (43) is infelicitous, just like (44), again with the antecedent understood as loose talk:

(44) It’s three o’clock, but it isn’t exactly 3 o’clock.

A more plausible view of (42) is that once the contrast between ‘believe’ and ‘be sure’ is activated, the contextually relevant threshold for ‘believe’ may be lowered. That would make it more like (45), where once the contrast between ‘good’ and ‘great’ is activated, the contextually relevant threshold for ‘good’ may be lowered.

(45) It's good, but it's not great.

In any case, if philosophers' believing outright is ordinary speakers' believing fully, and 'believe' sometimes stands for full belief, that is vindication enough of philosophers' way of using 'believe'.

We may be able to support the tentative argument of this section by reflecting on the cognitive value of the words 'think' and 'believe'. The next section broaches such a strategy.

7. Cognitive considerations

'Think' is one of the most commonly used cognitive words in the language. In particular, it is more commonly used than 'sure' and 'certain'. A reasonable working assumption is therefore that it stands for one of the cognitive phenomena most worth talking about—not just in philosophy or linguistics, but in ordinary life. A disturbing feature of views which make 'think' mandatorily weak is that they seem to violate that assumption.

Regarding p as at least slightly more probable than the relevant alternatives often has few practical consequences. For example, suppose that $\neg p$ is the relevant alternative to p , and I change from regarding p as 51% probable and $\neg p$ as 49% probable to regarding p as 49% probable and $\neg p$ as 51% probable. That is to change from thinking and believing p to thinking and believing $\neg p$, on a weak view of 'think' and 'believe'. But such a slight change of probabilities often has very few practical consequences. When I plan my actions, I have to take almost as much account of a possibility which I regard as 49% probable as I do of a possibility which I regard as 51% probable. Normally, I have to plan for both contingencies; I cannot afford to ignore the less probable one. The point applies even more strongly when more than two possibilities are relevant. If I regard p as 2% probable, and each relevant alternative to p as only 1% probable, I will not simply plan what to do on the assumption of p . Yet, in that context, I count as thinking and believing p , on a weak view of 'think' and 'believe'. Thus such views tend to minimize the very connections between thinking and believing on one hand and acting on the other which make thinking and believing *matter*, and give us reason to talk about them in ordinary life.

If I think that it will be dry, but in fact it is wet, I am mistaken, and my mistake may explain why things go badly for me, why I do not get to where I wanted to go. Mistakes have practical consequences. By contrast, if I just regard it as more likely to be dry than wet, but in fact it is wet, that does not yet mean that I am mistaken, and we have much less basis for explaining why things go badly for me. Of course, overestimating or underestimating probabilities can have practical consequences too, but one can have 51% credence in a proposition which turns out false without overestimating or underestimating anything.

These functional considerations cast some doubt on semantic accounts of 'think' and 'believe' as mandatorily weak, but by themselves they do not explain the linguistic evidence in favour of such accounts. Still, we can make a start. Perhaps *prototype theory* has something to teach us about belief ascription (for a *locus classicus* see Rosch 1978). It is

helpful in highlighting some important aspects of the cognitive psychology of categorization, even though much remains to be understood.

Consider a toy model of colour ascription. In judging whether an object *O* is red, we compare *O*'s colour profile to a *prototypical* shade associated with the word 'red' (focal red), as best we can. If the similarity is close enough, we describe *O* as 'red', otherwise not. We have a vague default standard of closeness, which is non-competitive, in the sense that it does not involve competing comparisons. It would be cognitively inefficient to bother making several comparisons when the first one already gives us a good match. However, sometimes the context of judgment motivates a competitive standard. For example, when asked 'Which of these apples is red?', we compare the colour profiles of the apples with focal red, and describe as 'red' those with the best match. In general, the result will depend on which comparison class is salient in the context. A reddish-yellow apple may be described as 'red' when the other apples are focal green, but not when they are focal red.

We can construct an analogous toy model of belief/thought ascription. In judging whether an agent *S* believes a proposition *p*, we compare *S*'s (dispositional) treatment of *p* to a prototypical case associated with the word 'believe' or 'think', as best we can. If the similarity is close enough, we describe *S* as 'believing' or 'thinking' *p*, otherwise not. We have a vague default standard of closeness, which is non-competitive, in the sense that it does not involve competing comparisons. It would be cognitively inefficient to bother making several comparisons when the first one already gives us a good match. However, sometimes the context of judgment motivates a competitive standard. For example, when we think of *S* as answering a question, we compare *S*'s treatment of the various possible answers to the question with the prototypical case associated with 'think' or 'believe', and describe *S* as 'thinking' or 'believing' the answer with the best match. In general, the result will depend on which comparison class is salient in the context. Suppose that *S* gives Liverpool a higher probability of winning the league than any other single team, but still less than 50%. When the question is 'Which team does *S* think will win the league?', we imagine *S* as addressing the ground-level question 'Which team will win the league?', whose possible answers are propositions of the form *team T will win the league*; since *S*'s treatment of *Liverpool will win the league* is closer to the prototype than is *S*'s treatment of any other possible answer, we describe *S* as 'thinking' that Liverpool will win the league. But when the question is 'Does *S* think Liverpool will win the league?', we imagine *S* as addressing the ground-level question 'Will Liverpool win the league?', whose possible answers are just *Liverpool will win the league* and *Liverpool will not win the league*; since *S*'s treatment of the latter is closer than *S*'s treatment of the former to the prototype, we now describe *S* as 'thinking' that Liverpool will *not* win the league, and not as 'thinking' that Liverpool will win the league.

The binary case may also help explain the appeal of neg-raising. For when we are deciding which of two competitor descriptions to apply to a given case, by comparison with their prototypes, a tie in similarity suggests that it does not matter which we apply; they are equally good. One of the prototypes will always be at least as close as the other to the target case, so if the comparisons do not license us to apply one of the descriptions, they license us to apply the other. Either *S*'s treatment of *p* is at least as close as *S*'s treatment of

$\neg p$, or *vice versa*. From this perspective, either we can describe S as thinking p or we can describe S as thinking $\neg p$, which is tantamount to neg-raising.

However, one should not concentrate too much on competitive contexts. As already observed, it is cognitively efficient to have a non-competitive default. In this respect, the standard linguistic method of working with minimal pairs may be misleading, by making such competitions more salient than they tend to be 'in the wild'. After all, in perception, facts often just strike us, without being answers to any question we had previously raised.

A natural conjecture is that the default non-competitive standard for applying 'think' and 'believe' just is that of *outright belief*, as we have been calling it: a good enough initial offer from the prototype to make competitive tendering unnecessary. That would require something considerably stronger than mere 51% credence. On this view, philosophers may well have been using 'believe' literally for outright belief all along.

A further natural conjecture is that the prototypes for 'think' and 'believe' include cases of *knowing*. After all, the most straightforward and familiar way of coming to think or believe that it is raining is by *seeing* that it is raining, and thereby coming to *know* that it is raining. Of course, when that happens, to describe you as 'thinking' or 'believing' that it is raining is to under-describe the case; it would be more informative and usual to describe you as 'seeing' that it is raining. Nevertheless, it is clearly literally true that you 'think' and 'believe' that it is raining. Analogously, asked to pick a prototype for the noun 'building', one might pick a house. Of course, when faced with a house, to describe it as a 'building' is to under-describe the case; it would be more informative and usual to describe it as a 'house'. Nevertheless, it is clearly and literally true that it is a 'building'. If cases of knowing can serve as prototypes for 'think' and 'believe', it is hardly surprising that knowing should entail believing. But the reason is not that we understand knowledge in terms of belief; it is that we understand belief in terms of knowledge.

We must be careful, however, not to treat the process of using comparisons with prototypes to judge when to apply a word as if it were built into the word's *semantics*. The psychological process is highly fallible, and subject to idiosyncratic variation, in ways which have no place in the semantics. We do better to treat comparison to prototypes as a fast and frugal *heuristic* for categorization, a cognitively efficient way of judging, reliable enough for ordinary purposes but capable of leading us into error. Philosophy and semantics have arguably been led into over-reliance on pre-theoretic judgments of cases by a failure to recognize the role of fallible heuristics in the aetiology of many kinds of judgment (Williamson 2020). In particular, if prototype-based heuristics play a role in our pre-theoretic applications of the words 'think' and 'believe' to cases, we cannot simply read the semantics of those words off those judgments. Nevertheless, we may expect the semantics to be a decent though imperfect fit with the judgments; the point of the heuristics is to achieve as much reliability as we easily can.

When we investigate the fit between the heuristics and the semantics, we need to ask questions like these: Does contextual variation in which prototypes are activated force contextual variation in the reference of 'think' and 'believe'? How responsive is the semantics to the psychological kinds underlying ordinary use of 'think' and 'believe', if the heuristics operate at a more superficial level? Those questions will be left open here.

8. *The knowledge norm of belief*

HRS use their account of belief to criticize the putative knowledge norm for belief: ‘Believe p only if you know p ’. As they point out, knowledge is an inappropriately high standard for weak belief. The slogan ‘Believe p likely only if you know p ’ is quite wrong-headed. After all, one might *know* p to be likely without knowing p . But if one says ‘Believe p only if you know p ’ in a context where ‘believe’ refers to outright belief, there is no such problem.

It is instructive to consider Daniel Whiting’s objection to a knowledge norm for belief in this setting. He is explicit that he is ‘concerned only with outright belief, not degrees of belief or confidence’ (2013: 184). Here is his objection (2013: 186):

The knowledge view appears to conflict with the ways in which we criticize and evaluate beliefs. Suppose David asks, ‘Who do you believe will win the next election?’ Kelly might reply, ‘The Republicans’. It would be very odd for David to reply, ‘You don’t know that!’ And it would be entirely appropriate for Kelly to reject this challenge by saying, ‘I never said that I did—I was only telling you what I believe’. Note that David might be right that Kelly does not know this but, still, his remark seems out of order.

Whiting’s comments on the conversation itself are plausible. However, David’s question ‘Who do you believe will win the next election?’ creates exactly the sort of context in which, RHS argue, ‘believe’ is interpreted especially weakly. Indeed, HRS use almost exactly the same example to argue that ‘believe’ does not refer to outright belief (2016: 1401):

[A]sking, ‘Who do you think/believe will win the election?’ does not seem to make any strong presupposition about the addressee’s knowledge.

Such wh-questions present the respondent with a forced choice between answers of the form ‘I believe X will win the election’. Evens if there are only two parties, Republicans and Democrats, the contextual standard for ‘believe’ is driven down almost to 50%. Thus Whiting’s example is wholly inappropriate, given his professed focus on outright belief. It is irrelevant to the knowledge norm for outright belief, his supposed target.

The effect of wh-questions on the interpretation of ‘believe’ is similar to its effect on comparable verbs with weak readings, such as ‘want’ or ‘like’. For example, a lawyer about to bargain with the other side might ask her client ‘Which do you want, to pay them \$100,000 now or to risk losing in court?’ In this context, either ‘I want to pay them \$100,000 now’ or ‘I want to risk losing in court’ is true as uttered by the client. In a context with a higher standard for ‘want’, of course both are false. Likewise with ‘Would you like to go first or second?’, as said by one prisoner to another in the queue for the guillotine. As usual, the contrasting mandatorily strong terms are much more resistant to such effects. ‘Which do you need, to pay them \$100,000 now or to risk losing in court?’ has a false presupposition if

both are live options. 'Would you love to go first or second?' sounds ironic. Nor do such examples show outright wanting and outright liking to be philosophers' inventions.

Rothschild (2020) objects to the knowledge-first account that an utterance of (46) is 'no admission of irrationality':

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

In his envisaged scenario, (46) is both true and fully believed by the speaker. Since (46) is true, its second conjunct is true, so she fully believes that Peter will come. Since she fully believes (46), she fully believes its first conjunct, in other words, she fully believes that she doesn't know that Peter will come. Given that she is also fully rational, her full beliefs are presumably closed under conjunction. Thus she fully believes (47):

(47) Peter will come but I don't know that he will come.

By hypothesis, the speaker fully and rationally believes (47), so she should be able to say (47) to herself with no inner reservation or qualification. But (47) is a Moore paradox, and it is hard to understand someone who says (47) to herself with full belief as free of irrationality. That tells in favour of a knowledge norm for full belief, not against it.

The knowledge-first approach to belief, and in particular the knowledge norm for outright belief, require far more development and defence than they have been given here. Nevertheless, on current evidence, the knowledge-first account has the resources to deal effectively with the case for weak belief.⁷

Notes

- 1 A probability space typically has no natural analogue of the accessibility relation in a Kripke model for those modal logics which allow contingently necessary and contingently contingent propositions. Thus the modal logic naturally associated with a probability space is S5 (or a proper extension thereof, if there are only finitely many outcomes).
- 2 Several of HRS's examples involve 'think' or 'believe' in the first person. I have avoided such examples because the frequent use of 'I think' and 'I believe' as a parenthetical qualifying comment or discourse marker is a potential confound. See Dehé and Wichmann 2010 and Kaltenböck 2009, 2010 for discussion.
- 3 There is a possible analogy with theories of scalar implicature which postulate a default assumption of opinionatedness to explain why a speaker who says 'Lilly believes that Bill ate some of the cake' is typically understood as implicating 'Lilly believes that Bill didn't eat all of the cake', otherwise for Gricean reasons the speaker would have said 'Lilly believes that Bill ate all of the cake'. For if Lilly has no opinion as to whether Bill ate all of the cake, the latter statement is false; moreover, a speaker with no opinion as to whether Lilly has the stronger belief is in no position to make the stronger statement. Given the opinionatedness assumption, $\neg Bp$ implies $B\neg p$. However, the opinionatedness assumption is much easier to cancel than neg-raising. One can simply add 'She has no opinion as to whether he ate all of it', whereas in response to the question 'Does he think it will rain?', the answer 'He doesn't think so; he has no opinion' sounds inconsistent. Thus it is unclear how the easily cancelled opinionatedness default can explain the persistence of neg-raising.
- 4 One worry about such examples is that the second person question 'Which team do you believe will win the league?' invites answers using the first person 'I believe', which raise the problems noted in fn. 1. The third person question 'Which team does he believe will win the league?' is in more danger of making a false presupposition.
- 5 Variations on HRS's example (19) which draw the stress away from the 'believes'/'is sure' contrast sound much worse: consider 'Mary believes it's raining, but she's not sure it's raining' and 'Mary does believe it's raining, but she's not sure it's raining' (perhaps in response to a denial that she believes it's raining).

- 6 An exception is Keith DeRose's suggestion that 'S is certain that P' is a better candidate than 'S believes that P' for expressing the subjective aspect of 'S knows that P' (2009: 186).
- 7 Ideas in this paper were presented in the 2017 Whitehead Lectures at Harvard, talks at Yale and the universities of Leuven and York, a conference at Cardiff University, a summer school at the University of Urbino, and classes at Oxford and Yale universities; thanks to participants for helpful discussion, and especially to Keith DeRose, Daniel Greco, John Hawthorne, Sarah Moss, Jennifer Nagel, Baron Reed, Daniel Rothschild, and Simon Wimmer for detailed comments.

References

- Ayer, Alfred. 1956: *The Problem of Knowledge*. London: Penguin.
- Chisholm, Roderick. 1957: *Perceiving: A Philosophical Study*. Ithaca: Cornell University Press.
- Dehé, Nicole, and Wichmann, Anne. 2010: 'Sentence-initial *I think (that)* and *I believe (that)*: prosodic evidence for use as main clause, comment clause and discourse marker', *Studies in Language*, 34: 36-74.
- DeRose, Keith. 2009: *The Case for Contextualism: Knowledge, Skepticism and Context, Volume I*. Oxford: Oxford University Press.
- Foley, Richard. 2009: 'Beliefs, degrees of belief, and the Lockean thesis', in Franz Huber and Christoph Schmidt-Petri (eds.), *Degrees of Belief*, 37-47. Dordrecht: Springer.
- Gettier, Edmund. 1963: 'Is justified true belief knowledge?', 23: 121-123.
- Hawthorne, John, Rothschild, Daniel, and Spectre, Levi. 2016: 'Belief is weak'. *Philosophical Studies*, 173: 1393-1404.
- Hawthorne, John, and Stanley, Jason. 2008: 'Knowledge and action', *Journal of Philosophy*, 105: 571-590.
- Holguín, Ben. 2020: 'Thinking, guessing, and believing', https://www.benholguin.com/uploads/1/1/3/6/113613527/tgb_website.pdf
- Holton, Richard. 2008: 'Partial belief, partial intention', *Mind*, 117: 27-58.
- Holton, Richard. 2014: 'Intention as a model for belief', in Manuel Vargas and Gideon Yaffe (eds.), *Rational and Social Agency: Essays on the Philosophy of Michael Bratman*, 12-37. Oxford: Oxford University Press.
- Ismael, Jenann. 2009: 'Probability in deterministic physics', *Journal of Philosophy*, 106: 89-108.
- Kaltenböck, Gunther. 2009: 'Initial *I think*: main or comment clause', *Discourse and Interaction*, 2: 49-70.
- Kaltenböck, Gunther. 2010: 'Pragmatic functions of parenthetical *I think*', in Gunther Kaltenböck, Wiltrud Mihatsch and Stefan Schneider (eds.), *New Approaches to Hedging*, 243-272. Bingley: Emerald Group.
- Lasersohn, Peter. 1999: 'Pragmatic halos', *Language*, 75: 522-551.
- Leitgeb, Hannes. 2017: *The Stability of Belief: How Rational Belief Coheres with Probability*. Oxford: Oxford University Press.
- Lewis, David. 1980: 'A subjectivist's guide to objective chance', in Richard Jeffrey (ed.), *Studies in Inductive Logic and Probability*, vol. II, 263-293. Berkeley: University of California Press.
- Loewer, Barry 2001: 'Determinism and chance', *Studies in History and Philosophy of Modern Physics*, 32, pp. 609-620.
- Martignon, Laura, and Laskey, Kathryn Blackmond. 1999: 'Bayesian benchmarks for fast and frugal heuristics', in Gerd Gigerenzer, Peter M. Todd, and the ABC Research Group, *Simple Heuristics that Make Us Smart*, 169-188. Oxford: Oxford University Press.

- Maudlin, Tim 2007: 'What could be objective about probabilities?', *Studies in History and Philosophy of Modern Physics*, 38, pp. 275-291.
- Moss, Sarah. 2018: *Probabilistic Knowledge*. Oxford: Oxford University Press.
- Ramsey, Frank. 1978: 'Truth and probability', in his *Foundations: Essays in Philosophy, Logic, Mathematics, and Economics*, ed. Hugh Mellor, 58-100. London: Routledge & Kegan Paul.
- Rosch, Eleanor. 1978: 'Principles of categorization', in Eleanor Rosch and Barbara Lloyd (eds.), *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum.
- Rothschild, Daniel. 2020: 'What it takes to believe', *Philosophical Studies*, forthcoming.
- Russell, Bertrand. 1912: *The Problems of Philosophy*. London: Williams and Norgate.
- Shope, Robert. 1983: *The Analysis of Knowing: A Decade of Research*. Princeton: Princeton University Press.
- Weisberg, Jonathan. 2020: 'Belief in psyontology', *Philosophers' Imprint*, forthcoming.
- Whiting, Daniel. 2013: 'Nothing but the truth: on the norms and aims of belief', in Timothy Chan (ed.), *The Aim of Belief*, 184-203. Oxford: Oxford University Press.
- Williamson, Timothy. 2000: *Knowledge and its Limits*. Oxford: Oxford University Press.
- Williamson, Timothy. 2007: 'How probable is an infinite sequence of heads?', *Analysis*, 67: 173-180.
- Williamson, Timothy. 2017: 'Acting on knowledge', in Adam Carter, Emma Gordon, and Benjamin Jarvis (eds.), *Knowledge First: Approaches in Epistemology and Mind*, 163-181. Oxford: Oxford University Press.
- Williamson, Timothy. 2020: *Suppose and Tell: The Semantics and Heuristics of Conditionals*. Oxford: Oxford University Press.
- Williamson, Timothy. 202X: 'Epistemic ambivalence', to appear in Nicholas Hughes (ed.), *Epistemic Dilemmas*. Oxford: Oxford University Press.
- Wimmer, Simon. 2019: 'Knowledge-first believing the unknowable', *Synthese*, forthcoming.
- Yalcin, Seth. 2010: 'Probability operators', *Philosophy Compass*, 5: 916-937.
- Yalcin, Seth. 2012: 'Bayesian expressivism', *Proceedings of the Aristotelian Society*, 112: 123-160.