SIMPLE BELIEF[1]

John Collins
Columbia University

For the *Synthese* special issue on "The Legacy of David Lewis"
edited by Marianna Antonutti and Pierluigi Graziani.
Final version of February 9th, 2018.

1. Two Kinds of Unlikelihood

Much recent philosophy of mind and language has focused on the problem of content; my primary concern has always been with attitude problems. That is to say, my main interest is not in what makes a particular mental state a belief *about X*, but in what makes it a *belief* about *X*, rather than some other attitude with that content. According to a commonsense functionalism, mental states are those inner states that occupy certain constitutive roles in the pattern of causal connection that obtains between a thinkers sensory input, her behavioral output, and her other mental states. I think that such an account of the mind must be correct, at least in broad outline.

Functionalists have tended to focus on belief-desire psychology. The attitudes of belief and desire are distinguished by their different directions of fit. Belief is the mental state that aims at truth (I try to adjust my beliefs so that they fit the way the world is); the aim of a desire is satisfaction (I try to adjust the world to fit my desires). Stalnaker has put it neatly in the following words:

> Belief and desire ... are correlative dispositional states of a potentially rational agent. To desire that P is to be disposed to act in ways that would tend to bring it about that P in a world in which one's beliefs, whatever they are, were true. To believe that P is to be disposed to act in ways that would tend to satisfy one's desires, whatever they are, in

[1]Early versions of his material were presented as the Jerrold Katz Memorial Lecture at the CUNY Graduate Center in February 7, 2007, at the Synthese Annual Conference "Between Logic and Intuition: David Lewis and the Future of Formal Methods in Philosophy", the Carlsberg Academy, Copenhagen, October 3, 2007, at "Another World is Possible: a Conference on the Work of David Lewis" at the University of Urbino, June 16-18, 2011, and at a Columbia Philosophy Work-in-Progess Seminar on March 6, 2014 . My thanks to the audiences on those occasions and to two anonymous referees for this journal for insightful comments and criticism.

> a world in which P (together with one's other beliefs) were true. ([19] p.15)

Now alongside this, which I must admit seems perfectly good to me as it stands, we have the idealization of commonsense belief-desire psychology provided by Bayesian decision theory. Bayesian decision theory is, as David Lewis reminds us, "no esoteric science", but simply "the core of our commonsense theory of persons, dissected out and elegantly systematised" ([11] pp. 337–338). But the elegance comes at a price. In several respects decision theory provides an unrealistic account of our ordinary thought and talk about rationalizations of intentional behavior. While the notion of a degree of belief is unquestionably of great epistemological importance, I would argue that in many contexts there is nothing incomplete about, and nothing at all wrong with, intentional explanations that appeal only to the sort of attitude that need not be construed as admitting of degree.

Call these the attitudes of simple belief and simple desire. Let me give an example:

> She is going to the fridge because she wants some milk for her coffee and she believes that that's where the milk is.

It seems to me that this simple kind of statement can serve as a perfectly good and quite complete intentional account. It should not be seen as shorthand for some more complete and accurate story involving subjective probabilities and utilities.

There's a long and interesting discussion of these issues in Richard Foley's monograph *Working Without a Net*. I'm very sympathetic to the view that Foley expresses when he says:

> There are deep reasons for wanting an epistemology of beliefs, reasons that epistemologies of degrees of belief by their very nature cannot possibly accommodate. ([4] p.170)

So we have these two ways of thinking about belief: simple belief, and belief that admits of degree. let me briefly note a few points about these two conceptions.

First of all, there is no straightforward way of translating back and forth between the two pictures. In particular, there is no straightforward reduction of simple belief talk to subjective probability talk. Simple belief is not the same thing as what is sometimes called "full belief". Simple belief that *p* doesn't require one to assign probability 1 to *p*. She believes that the milk is in the fridge. True. But that's not to say that she assigns probability 0 to a milkless fridge. Simple belief does not require certainty.

Secondly, we should also reject the "threshold" view, also known as the "Lockean Thesis", according to which one believes that *p* in the simple sense just in case one's degree of belief in the proposition that *p* exceeds a certain numerical threshold. More about this below.

Thirdly, it is tempting to think that the difference between the concept of belief that admits of degree and the concept of simple belief is just the difference between a formal and an informal account of belief. We should resist that temptation. There are those for whom "formal epistemology" simply means Bayesian epistemology, but I think it would be a mistake to think that the only possible formal development of the notion of belief is in terms of subjective probabilities.

Fourthly, neither of the two accounts of belief is particularly good descriptively. Our actual epistemic lives are messy and not easily regimented. I see both conceptions as normative, rather than descriptive, in that they set out ideals to which actual epistemic agents ought to aspire, even if in practice we fall short.

But, fifthly, they should, or need not, be seen as being in competition with one another. Perhaps, for example, the notion of simple belief might be thought of as figuring appropriately in some account of what it is to be minimally rational, or rational in some bounded sense.

Now in fact I am going to go on and suggest a somewhat different diagnosis of the situation. I think that each conception captures an important thread in our perhaps confused pre-theoretic thinking about epistemology. Here's roughly how the story will go. At bottom, the differences between the two accounts of belief stem from two different conceptions of what, for lack of a more neutral available term, I'll call *unlikelihood*.

On the one hand there is unlikelihood in the sense of *improbability*, on the other hand there is unlikelihood in the sense of *far-fetchedness*. When I consider something to be far-fetched, I am thinking of it as being not at all similar to the way things actually are. This is exactly the sense of distance from actuality familiar from the Stalnaker-Lewis semantics for subjunctive conditionals. We are entitled to ignore the more far-fetched ways that an antecedent might possibly have been true when evaluating such conditionals.

Corresponding to these two senses of 'unlikely' we have the two concepts of belief. In both cases believing is a matter of shunning those possibilities that are taken to be unlikely. If we take unlikely to mean improbable, then we get the Bayesian account of degrees of belief as subjective probabilities. I will assume that the reader is familiar with that story, and just take it for granted. But I'll need to say more below about how the notion of simple belief might be given a corresponding formal development.

2. Can Simple Belief be Reduced to Degrees of Belief?

Let me now say some more about the third of the points made above. Can we reduce talk about simple belief to talk about degrees of belief? That is, given a subjective probability function P that represents an agent's degrees of belief, can we arrive at that agent's *simple beliefs* in a way that doesn't reduce simple belief to full belief, i.e. to certainty?

One naive thought is encapsulated by

The Lockean Thesis: For some $0 \leq r < 1$: X is believed if and only if $P(X) > r$.

Here r is a probabilistic threshold level for belief, The parameter r may be thought of as measuring the agent's *cautiousness* as a believer. The higher the value of r the more cautious the agent is in accepting propositions as true.

Of course the Lockean Thesis runs afoul of Kyburg's Lottery Problem. No matter how high a value of r is chosen (< 1) we can consider a fair lottery with more than $1/(1-r)$ tickets. Then the Lockean Thesis entails that the agent believes of each ticket that it will not win, while yet believing that some ticket will win.

There has been a lot of recent formal work on this problem, some notable examples of which are by Hannes Leitgeb, in [8],[9], and [10], by Hanti Lin and Kevin Kelly [14], and by Kenny Easwaran [3].

Each of these various proposals has its costs. A full discussion of these matters is beyond the scope of the present paper, but let me try to give you the flavor of this in the case of Leitgeb's account, without going too deeply into the technical details.

Leitgeb's main idea is to think of an agent's simple belief state as a set of worlds which not only receives a sufficiently high subjective probability but for which the probability remains *resiliently* or *stably* high when one conditionalizes on other propositions that are consistent with what the agent believes. The scheme is relativized to a particular partition of the possibility space into "worlds". Leitgeb's central technical notion is of what he calls *P-stability*:

Proposition A is said to be *P-stable* if and only if for every proposition B consistent with A, and for which $P(B) > 0$, the conditional probability $P(A/B) > \frac{1}{2}$. ([10]

p.120.)

The *P*-stable propositions turn out to be nested, and so we may define an epistemic ordering of worlds given by the order in which they first appear in a *P*-stable proposition, when those nested propositions are ordered from strongest to weakest.

Then, as Leitgeb suggested in [8], we might take the agent's belief state to be the strongest *P*-stable proposition. Or, as Leitgeb allows in his more recent [9] and [10], the belief state might also be given by one of the other, weaker, *P*-stable propositions, depending on the agent's degree of caution. In either case a proposition is believed (in the simple sense) if and only if it is entailed by the agent's belief state.

So where's the difficulty here? It comes in the form of what I call the Refinement Problem. The finer we make an agent's partition of possible states, for example by giving the agent a fair coin which might be tossed arbitrarily many times, the higher we can make the threshold probability required for a proposition to be believed. By sufficient refinement of the partition of possible states, we can collapse the stability proposal into the claim that simple belief requires certainty. Leitgeb acknowledges this problem, see e.g. [10] p.137, but I think it is more troublesome than he allows. The worry is a serious one, since most standard stories about the elicitation of degrees of belief from preference (e.g. Savage, Bolker-Jeffrey etc.) involve some sort of "fine partition" assumption to ensure that we end up with a real valued representation of degree of belief.

3. A Problem for the Reconcilation of AGM Simple Belief with Credence

The recent literature on the reconcilation of simple belief and probability has concentrated on Kyburg's lottery problem. In this section I'd like to sketch a new and rather different problem facing such proposals. I am indebted in what follows in this section to a helpful conversation with Anubav Vasudevan.

Suppose that we think of simple belief on the AGM model of qualitative belief and its revision formulated by Alchourrón, Gärdenfors, and Makinson. The details of the AGM revision axioms need not detain us here (but see [5]). It will suffice to say that the AGM account presupposes an ordering of those possibilities inconsistent with what the agent currently believes, and that when faced with the requirement to revise one's belief state so as to accept a disjunction of such possibilities, one of which is more far-fetched than the other according to that ordering, the AGM scheme prescribes coming to believe the disjunct that is the less remote of the two.

Now where the relationship between AGM belief revision and probabilistic talk of degrees of belief has been discussed in the literature, authors have almost invariably followed Gärdenfors's lead in Chapter 5 of [5] and taken qualitative belief on the AGM scheme to model full belief in the probabilistic sense.

It seems they have had good reason to do so. For suppose we resist the idea that AGM models *full belief* and want to allow that one might believe, in the AGM sense, propositions in which one has credence less than 1. And suppose, secondly, that the AGM ordering of those possibilities inconsistent with what one currently believes is indeed an ordering with respect to what we have been calling 'far-fetchedness', rather than, for example, something that might be cashed out in probabilistic terms.

Then we face a new problem. Consider some very far-fetched possibility Y which, however, is uneliminated by your evidence and to which you assign some very small positive probability ε.

For example let Y be the counterfactual possibility that you are a brain in a vat, connected to a supercomputer simulation of a virtual world which has been providing you with experiences which are qualitatively a perfect match for the entire course of your life experience in the actual world. If you are like me then (i) you believe that Y is false—I believe that I'm not a brain in a vat, and I assume that you do too, it's a completely crazy far-fetched hypothesis—but (ii) if pushed, you will admit that you cannot be absolutely certain that Y is false. If the stakes are moderate ones, then I will bet that Y is false. But if you raise the stakes on me high enough, then at a certain point I will waver. Would you accept a bet that won you a one-billionth chance of a penny if Y is false, but which condemned you to eternal damnation and hellfire if Y turned out to be true? I wouldn't. And that's simply to say that, for me, $P(Y) = \varepsilon$, for some very small $\varepsilon > 0$

So Y is a far-fetched and very improbable hypothesis which I believe to be false, because, while it is uneliminated by my evidence, I properly choose to ignore it, i.e. I don't take it to be a serious possibility even though I cannot be certain of its falsehood.

Now, however small the positive number ε is, one may construct a fair lottery with sufficiently many tickets that the probability of any single ticket winning is less than ε. Suppose you hold a single ticket in this lottery, and let X be the proposition that your ticket is the winner.

The proposition X is significantly unlike proposition Y, because although it is also wildly improbable (more improbable in fact than Y is) it is not at all unlikely in the other sense. That your ticket wins the lottery is not at all a far-fetched hypothesis. Whatever happens, there will be a possible world that is very close indeed to the actual world and at which your ticket does win. And most of us are of the opinion that the possibility X is *not* properly ignored, so that whether or not you believe that the ticket will lose, it would be incorrect to say that you *know* that it will.

Now here's the problem. Suppose that you receive completely unimpeachable evidence that the disjunction X or Y is true. Suppose that God calls you on the telephone and tells you: "Either your ticket wins the lottery, or you are a brain in a vat!"

How do you update your beliefs in response to this new information? Well, if you are Bayesian, the answer is easy: you conditionalize your current subjective probabilities on the proposition X or Y and after you have done this you will have a degree of belief somewhat greater than one half in the proposition that you are a brain in a vat, and a degree of belief corresonding less than one half that your ticket will win the lottery. Since almost everyone in the business of reducing simple belief to degrees of belief agrees that you cannot be said to believe X when your subjective probability for X is less than one half, almost everyone is going to say that your posterior Bayesian belief state is one in which you do not believe that you will win the lottery. But if your revision goes according to the AGM picture of simple belief, then you will update completely differently. You will shift to a belief state that includes only what you take to be the closest X or Y worlds to your prior belief state, and clearly this will be a posterior state in which you *do* believe that you will win the lottery!

Note that the problem can be made even more stark by lowering the prior probability of winning the lottery. If, for example, the number of tickets is so great that your prior probability of winning is only one ninth the probability that you initially assign to being a brain in a vat, then, after conditionalizing on the disjunctive information, your credence in the proposition that you win the lottery will be only 0.1 and yet that will be what AGM revision prescribes that you believe.

This simple point seems to me to be a fairly devastating objection to a project which (i) distinguishes simple belief from full belief, (ii) takes simple belief to be a matter of shunning possibilities that are unlikely in the sense of being far-fetched, (iii) uses AGM revision to model the dynamics of simple belief, and (iv) attempts to relate simple belief to credence. I would like to explore options (i) and (ii) further here at the expense of (iv) and perhaps (iii) as well. I will have nothing more to say about AGM revision in the remaining sections of the paper. [2]

4. Simple Belief: Towards a Free-Standing Account

I propose instead a free-standing formal account of simple belief that does not attempt to reduce or relate it to talk of degrees of belief and which in fact avoids probabilistic talk completely. What are the normative constraints on belief in the simple sense? I think we should start from the assumption that belief aims not just at truth but at attaining the status of knowledge. Conditions that we think knowledge should satisfy are conditions that should guide the formation of our beliefs.

So I shall start by considering what I take to be a (rather weak) necessary condition on knowledge. It is a condition that I introduced and defended in Collins [1] as the:

Close Shave Principle: S knows that X entails that there is no possibility that is very close to actuality at which X is false and to which S assigns non-zero probability.

As just stated, the principle mentions the agent's subjective probabilities, i.e. her degrees of belief. Since we are working towards a free-standing, non-reductive account of belief, the principle will eventually be reformulated below in a form that doesn't mention probability. But let me first illustrate and motivate the principle in its original form.

Note that the Close Shave Principle (henceforth 'the CSP') is a *modal* condition on knowledge. So it might be helpful to distinguish it carefully from two more familar modal principles that have been offered as necessary conditions on knowing and which have been widely discussed in the epistemology literature. The first of these is what has become known as:

[2] Leitgeb is a notable exception to the rule mentioned above. He does not assume that belief in the AGM sense requires probability 1 and in [10] develops a correspondence between AGM revision and Bayesian credence in which the former is allowed to model simple belief. He holds on to (i), (iii), and (iv) here. However Leitgeb violates assumption (ii): that the AGM ordering should be one of comparative far-fetchedness rather than something that captures likelihood in probabilistic terms. This can be seen most clearly from his "Outclassing Condition" p.108 which states that for a proposition A to be P-stable, the least probable way for A to be true must be more probable than the negation of A. Since the nested P-stable propositions are precisely the ordered AGM fallback positions, this yields a probabilistic notion of comparative likelihood.

Sensitivity: If X had been false, S wouldn't have believed that X.

This condition was introduced by Robert Nozick [15] as part of his account of knowledge as *truth tracking*. Nozick's idea was that the Sensitivity condition could simultaneously serve to diagnose what had gone wrong in the Gettier examples, and also block the standard argument for radical skepticism about knowledge by invalidating the closure principle to which that skeptical argument appeals. If knowledge requires Sensitivity of belief, knowledge will not be closed under known entailment.

For example: my belief that I have hands is sensitive to the fact that I have hands, since the closest worlds to the actual world at which I don't have hands are worlds in which perhaps, my mother while pregnant took the drug thalidomide, or worlds in which I suffered an unfortunate accident with a chainsaw. In such situations I would clearly be aware of my handlessness.

Now the proposition that I have hands entails that the skeptical hypothesis that I am a handless brain in a vat is false. Yet my belief that the skeptical hypothesis is false does *not* satisfy Sensitivity. That skeptical hypothesis has been carefully constructed so that, were it true, which it's not, my beliefs would be exactly the same as they actually are, for the reason that the entire course of my life experience would be qualitatively indistinguishable from the way it actually has been.

That was Nozick's plan. Few people today defend the Sensitivity condition. The demolition job undertaken by Saul Kripke in the mid 1980s on Nozick's truth-tracking, and finally published as [7], made it clear that the counterexamples to Closure could not be contained or restricted to far-fetched cases like the skeptical hypothesis, but in fact popped up almost everywhere in forms that not even fanatical opponents of Closure could countenance.

Kripke's demolition of the truth tracking theory has sometimes struck me as similar to what the Roman general Scipio did to the ancient city of Carthage. Yet the idea of imposing modal conditions on knowledge is such a natural and attractive one that it's hard to resist the temptation to sow seeds in that salted earth.

The second more familiar modal condition which I would like to distinguish from CSP is Ernest Sosa's [18] 'Safety' condition. This is the contrapositive of Sensitivity, and sometimes expressed this way:

If S were to believe that X, then X would be true.

Note that this is not logically equivalent to Sensitivity because contraposition is not a valid inference scheme for the subjunctive conditional. Note also that the above formulation is a particularly infelicitous way of expressing the idea in English. For Safety is being offered as a necessary condition for knowledge, yet if S knows that X then clearly X is true and S believes that X, so what sense does it make to assert the counterfactual sounding claim that 'If S were to believe that X, then X would be true'?

The only way I can think of to express the thought in relatively natural sounding English is to say:

S wouldn't believe that X unless X were true.

which has the disadvantage, however, of being couched as a fairly opaque 'unless' construction which only serves to disguise the logical relation of this claim to Sensitivity. Perhaps the clearest way to put it is directly in terms of closeness. Thus:

Safety: In all the closest worlds in which S still believes that X, X is true.

I won't say any more about Safety here, other than to draw your attention to one important feature that distinguishes between Sensitivity and Safety on the one hand, and the CSP on the other. Both of the former conditions, Sensitivity and Safety, are concerned with what your belief state would be in close counterfactual situations. They have to do with counterfactual states of belief. The CSP, by contrast, concerns just your *actual* belief state and what that belief state has to say about close counterfactual situations in which X is false, namely: that you are certain that none of them is actual. So the distinction is one between counterfactual states of belief, and belief about counterfactual situations.

The CSP is a lot weaker than those more familiar conditions. It is important to recognize that the Principle is only being offered as a very weak *necessary* condition on knowing. In general, satisfaction of the CSP falls far short of what is sufficient for knowledge. Take the case of "Paranoid Joe" for example.[3] Joe is given a lottery ticket as a present. But Joe has the paranoid belief that the lottery is rigged and so he assigns zero probability to the proposition that his ticket wins. Clearly Joe doesn't know that his ticket will lose, but note that the CSP doesn't explain why, in this case, that is so. The CSP is satisfied. Presumably we would want to say that the reason Joe doesn't know his ticket will lose is because he is unjustified in assigning probability zero to the proposition that it wins.

[3]This example was suggested by an anonymous referee.

Weak as it is, the CSP does capture an important feature of knowing. It was originally formulated in an attempt to capture a minimal necessary condition on knowledge that would allow one to explain the *Lottery Observation* (not to be confused with Kyburg's Lottery Problem mentioned earlier): When I hold a single ticket in a fair lottery with many tickets, then, no matter how probable it is that my ticket will lose, I cannot correctly be said to know that it will lose.

Why? Well: because even when my ticket actually loses there is a very close possible world in which it wins. That world gets non-zero probability, and so according to the CSP, I don't know that my ticket will lose.

My idea in isolating this modal principle and using it to explain the Lottery Observation was to go on and argue that lottery cases don't generalize in the way that many have claimed they do and hence that the Lottery Observation poses no particular skeptical challenge.

Jonathan Vogel [20] was an early proponent of that supposed challenge. He illustrates the point with the following example. Suppose I parked my car a couple of hours ago on a city side street. I remember where I parked it. But now, two hours later, sitting in a cafe across town, do I know where my car is? If knowledge is closed under known entailment, then my knowing that my car is still there on the side street, entails that I know it has not been stolen and driven away. But Vogel thinks that lottery considerations show that I cannot know the latter. He says:

> In effect, when you park your car in an area with an appreciable rate of auto theft, you enter a lottery in which cars are picked, essentially at random, to be stolen and driven away. Having your car stolen is the unfortunate counterpart to winning the lottery. And, just as one doesn't know that one will not have ones number come up in the lottery, it seems one doesn't know that one's number won't come up, so to speak, for car theft. ([20], p.16)

Hawthorne [6] argues from similar examples that this kind of skeptical threat is a widespread phenomenon. Here's a variant on one of his cases. I may think I know that I'll be visiting Venice later this summer. I've booked plane tickets and accommodation. But do I really know I'll be in Venice? Do I know, for example, that I won't suffer a fatal pulmonary embolism in the next couple of weeks? Is medical self-examination quite so easy?

But if the Lottery Observation is explained by violations of the CSP, then this skeptical argument depends on there being a genuine lottery mechanism involved, i.e. it depends, in this case, on there being a very close possible world to the actual world in which I do die in the next two weeks as the result of an embolism. But there will only be a close such alternative to actuality if I actually have some condition that puts me at risk, for example: an untreated thrombosis in my leg. And I may know that I have no such clot in my leg, not by statistical reasoning, but because e.g. I'm taking anti-coagulant medication. The mechanisms by which pulmonary embolisms, and car thefts, occur are not remotely like lotteries. The disanalogy between real lotteries and the supposed generalizations can only be ignored if we confuse the two distinct kinds of unlikelihood, if we take the fact that a possibility is merely improbable to be an indication that it is also outlandish or far-fetched.

Mention might well be made here of a parallel between my use of the CSP in [1] and the notion of *normic support* developed by Martin Smith in, for example, [16] and [17]. In Smith's terminology:

> ... a body of evidence *E normically supports* a proposition *P* just in case the circumstance in which *E* is true and *P* is false requires more explanation than the the circumstance in which *E* and *P* are both true. ([17] p.40, see also [16] p.16)

If we suppose that worlds can be ranked with respect to their normalcy, then we can follow Smith in reformulating the concept of normic support in terms of variably restricted quantification over worlds:

> [A] body of evidence *E* normically supports a proposition *P* just in case *P* is true in all the most normal worlds in which *E* is true. ([17] p.42)

Now Smith goes on to suggest that "in order for one to have justification for believing a proposition *P*, it is necessary that ones' body of evidence *E* normically supports *P*." ([17] p.42). Since justification for believing a proposition is a necessary condition for knowing that proposition, Smith has now given us a necessary condition on knowing that is at least close enough in spirit to the CSP as to allow him to provide a parallel diagnosis of lottery skepticism to that which I gave in [1]. For example, Smith says of Hawthorne's heart attack case (of which the pulmonary embolism example above was a variant):

> If a young, fit, apparently healthy friend were to suffer a fatal heart attack, this is something that would immediately prompt us to seek

> possible *explanantions*: was he suffering from a virus? Did he have a congenital heart abnormality? Was he poisoned? We may even develop a psychological *need* to identify some explanation of this sort. This is not a reaction to something we regard as a 'pure chance' event. If a friend had won the lottery, rather than suffering a fatal heart attack, then, while this may be equally suprising, it would not prompt us to look for explanations. ([17] p. 58).

One major difference beween the current proposal and Smith's normic support account is that the latter appeals to pragmatic elements, such as the need for explanation, in establishing a normalcy ranking for worlds, whereas the CSP appeals to an objective ordering of worlds with respect to their closeness to (or far-fetchedness from) actuality. To someone skeptical of the metaphysics, this might be a count in favor of Martin Smith's approach. I will return briefly to discuss pragmatic issues in the concluding section of the paper.

But I would now like to recast the CSP so as to remove the reference to degrees of belief. One way of doing this is to take the qualitative notion of *certainty* to correspond to the quantitative notion of assignment of subjective probability 1. We shall assume that neither belief nor knowledge entails certainty.

CSP*: If S knows that p then for any possible world w that is very close to actuality and at which p is false, S is certain that w is not the actual world. [4]

Suppose S knows that p but is not certain that p. Then the reformulated principle tells us that whatever $\neg p$ possibilities S is not certain are false cannot be very close to the actual world.

If we say, following Lewis, that a possibility w is uneliminated by S's evidence iff the entire course of S's perceptual experience in w exactly matches her actual perceptual experience, and if we identify what is certain for S with what is entailed by S's perceptual experience, then the necessary condition on knowing that I have in mind can be expressed as follows:

CSP:** If S knows that p, then there is no possibility uneliminated by S's perceptual experience that is very close to actuality and at which p is false.

Note that CSP** is somewhat stronger than our original principle. This can be seen, e.g., from the fact that in the earlier Paranoid Joe example, the claim that

[4] I am indebted here to an anonymous referee who pointed out an error in an earlier version of this reformulation of the Principle.

Joe knows his ticket will lose *is* contradicted by CSP**, though not by the CSP as originally formulated.

CSP** is very similar, of course, to a combination of what Lewis calls the "Rule of Actuality" and the "Rule of Resemblance". The latter states:

> Suppose one possibility saliently resembles another. Then if one of them may not be properly ignored then neither may the other. ([12], p.429)

When Lewis refers to a possibility as being properly ignored, he means that one can know it to be false despite the fact that it is uneliminated by one's evidence. The Rule of Actuality states that actuality may not properly be ignored; it follows from that and from the Rule of Resemblance that if S knows that p, then there is no uneliminated possibility that saliently resembles actuality and at which p is false.

Now this is just CSP** with 'saliently resembles' replacing 'is very close to'. So if 'closeness' is then cashed out as similarity in some salient respect, the fit between these two conditions is pretty tight.

But in fact there are two key respects in which my proposal differs from Lewis's. First of all, a relatively minor point, as becomes clear when he turns to the discussion of particular examples, Lewis's standard of what is required for resemblance is not as strict as what I have in mind; roughly speaking, where he means close, I mean very close. I'm talking about tweaking the positions of a few particles. So, e.g., CSP** is not well suited to handling Gettier cases, as Lewis ambitiously hoped for his Rule of Resemblance.

Secondly, and more importantly, when I speak of "closeness" in the various formulations of CSP, what I intend is what I earlier referred to as closeness in the objective sense. On the other hand, it becomes clear from Lewis's discussion that he intends 'resembles' in a fairly liberal sense and worries that there will be no principled way to prevent this from also including resemblance with respect to the way things seem to the subject to be. This would lead straight to skepticism. As Lewis writes:

> We must apply the Rule of Resemblance with care. Actuality is a possibility uneliminated by the subjects evidence. Any other possibility W that is likewise uneliminated by the subjects evidence thereby resembles actuality in one salient respect; namely, in respect of the subjects evidence. That will be so even if W is in other respects very

> dissimilar to actuality— even if, for instance, it is a possibility in which the subject is radically deceived by a demon ... that would be capitulation to scepticism. ([12], p.430)

In fact, this running together of objective and subjective notions of closeness in the formulation of a modal condition on knowing is something that goes right back to Nozicks truth-tracking theory.

5. A Problem for Inductive Reasoning?

We often acquire knowledge of law-like generalizations by inductive reasoning from limited samples. The various forms of CSP we are considering here face a prima facie problem, however, in cases where it turns out that the law-like generalization is only statistical in nature. Consider, for example, the general claim that: a cube of ice placed in a large container of lukewarm water will melt. Suppose that earlier today I put an ice cube in a thermos of lukewarm water and then closed the lid. Surely, now, several hours later, I know, even without removing the lid of the thermos, that the ice cube has melted. Yet physics tells us that the generalization on which that knowledge claim is based is only a statistical one. There are highly unusual, yet physically possible, states of the system consisting of the water plus the ice cube in which the initial trajectories of all of the particles are coordinated in just such a way that the ice will remain frozen and floating in the warm water. Of course these configuations, though physically possible, are extremely improbable. Yet if I assign even a miniscule positive probablity to such a possible state, doesn't the CSP then entail that I cannot know the ice has melted? Since I obviously can know that the ice has melted, that would count as a refutation of the CSP.

But recall the way the CSP is framed. It doesn't say that my knowing that *p* is inconsistent with my assigning non-zero probability to a possibility in which *p* is false. Rather it says that in order to know that *p* I may only assign non-zero probability to such possibilities if, in fact, they are not close to the actual world.

In the case of the ice cube, that's a safe bet. Not only is it highly improbable that the actual initial state of the system is one of those unusually coordinated ones in which the cube remains frozen; it is also highly improbable that the actual world is very close to being in such a state. That is, it is highly probable that the actual world is not one that could be turned into an unusually coordinated one by the tweaking of the positions and velocities of a couple of particles, but one in which, rather, that would require wholescale adjustment of many, many particles.

So the CSP does not force the conclusion that I don't know the ice has melted.

Similar considerations apply to recent discussions of skepticism about counterfactuals. Alan Hájek has argued that most counterfactuals are false. Hájek starts with examples in which we know that there is a non-zero probability of some quantum event that might occur were the antecedent true. But others, see e.g. Keith DeRose [2] and Karen Lewis [13], have suggested that quite ordinary possibilities of low probability might suffice to make the skeptical point. So, for example:

> If Sophie had gone to the parade, she would have seen Pedro dance.

is inconsistent with the might conditional

> If Sophie had gone to the parade, she might have been stuck behind someone tall (and so not seen Pedro dance).

so if the latter is true, the former is false.

But in this case, whether or not the might conditional is true depends on whether there really is a close world to the actual world in which Sophie gets stuck behind someone tall at the parade. And that depends, for example, on where the tall people at the parade are actually standing, and whether some of them might be friendly enough to step aside to give a shorter person a view, etc, etc. Who knows what is actually the case here? My main point here is simply this: even if (before I learned she didn't go) I assigned non-zero credence to Sophie's getting stuck behind someone tall, it doesn't follow that I am also committed later to the conditional 'If she had gone, she might have been stuck behind someone tall'. That's to confuse matters of credence and epistemic possibility, subjective notions, with objective issues about which possibilities are close to the way things actually are. It's another version of the conflation of our two distinct notions of likelihood. What's the argument from the epistemic possibility of Sophie's view being blocked to the conclusion that there must be possible worlds that diverge from the actual world via a small miracle which lead to Sophie's view being blocked? I don't think this follows.

The situation is even clearer with the statistical mechanical ice cube example. At noon I contemplate placing an ice cube in a thermos of warm water but then think better of it. It is true that there are physically possible initial microstates of the system consisting of the ice cube and thermos that are (a) macroscopically indistinguishable from the actual microstate of the system, and which (b) evolve according to the laws of physics into a state in which the ice now remains frozen in the warm water. Does it follow that at 2 p.m. I am committed to saying: If I had placed the ice cube in the thermos at noon, it might now still be frozen?

No, it doesn't. It's not just that the possibility that the ice remains frozen is overwhelmingly improbable. In addition it's also overwhelmingly improbable that any such possibility is even remotely like the way things actually were in the freezer at noon yesterday. Almost certainly there is no small adjustment to the actual state of the system at noon yesterday that would transform it into one in which the ice fails to be melted now. That would require a quite extraordinary coordination of the trajectories of the individual particles in the system. Such extraordinary coordination is physically possible, yes, but that possibility differs from the way things actually were at noon yesterday with respect to the positions and velocities of a vast number of molecules. Any possibility that differs from actuality with respect to a vast number of independent matters of particular fact is not a possibility that is even remotely close to the way things actually are. So almost certainly you can't get from the actual world to an ice-still-frozen-now world via a small miracle at noon yesterday. So almost certainly the might counterfactual is false.

6. Towards a Non-Probabilistic Model of the Statistical Inductive Cases

The discussion of the previous section falls short of providing anything like a formal account of the way in which statistical inductive reasoning could be handled on the simple conception of belief. Can we do better? Here's a quick sketch of an idea.

If the subject's aim is not just to believe only what is true, but also to know those truths, then I think we can usefully see the CSP as providing at least one central normative principle guiding the formation of one's simple beliefs. What would the content of such a normative guiding principle be?

The CSP** tells us that my belief that *p* fails to count as knowledge if there is a way for *p* to be false that is uneliminated by my evidence and very close to the way things actually are. So if my aim is not just true belief, but knowledge that *p*, then I should strive to avoid being in that situation. That is, if I'm going to omit from my belief state any possibilities that are not directly eliminated by experience (and I will have to do so as a practical matter if I'm to avoid ending up in the motivational coma of a radical skeptic) then I should try to avoid omitting any possibilities that are very close to the way things actually are. My doxastic state at any time should comprise actuality along with those uneliminated possibilities that are, in some sense, the best candidates for being close alternatives to actuality.

By framing things in this decision-theoretic way I might be thought to be taking a stand in favor of some voluntarist notion of belief. But in fact I'm not sure where I stand on that score. So while I find the decision-theoretic framework is a useful way here of presenting these ideas, I'd like to remain officially neutral as to whether the framework is to taken completely literally as providing an explicit canon for belief formation, or rather as some kind of heuristic rational reconstruction.

And in fact I have no such general account to offer anyway. Lacking a general account I propose to proceed here by discussing one particular simple example, a toy example in some respects, but one which I hope possesses some of the key features of a wide range of real life cases. The example I'm going to discuss is intended to be an idealization of a case discussed by Brian Weatherson. Here is Weatherson's example:

The Mighty Cats: Mark is watching his favorite Australian Rules football team, Geelong, lose another game. Geelong are down by eight goals with fifteen minutes to go. His housemates are leaving to go see a movie, and want to know if Mark wants to come along. He says that he's watching the end of the game because Geelong might come back. One of his housemates replies sarcastically, "I guess it's possible they'll win. Like it's possible they'll call you up next week to see if you want to join the team and play for them!" Mark replies, "Yeah, you're right, this one's over. What's the movie?" Mark doesn't just give in to his housemates, he forms the belief that Geelong will lose. Later that night, when asked what the result of the game was, he says that he didn't see the final score, but that Geelong lost by a fair bit. The upshot is that what Mark does can count as belief formation, even if his credence that Geelong will lose does not rise. ([21], pp.552–553.)

What interests me most about this example is Weatherson's description of what goes on here as a case of belief formation without any change in the underlying probabilities. Weatherson goes on to say that while it is tempting to interpret the housemate's mocking retort as offering Mark a reason of some sort for believing that Geelong will lose the game, a reason that, presumably, impacts his doxastic state by causing him to assign a lower probability to the possibility of his team's making a dramatic last minute come-back, we should resist this temptation, and see this change for what it really is: an exercising of a capacity to cease to take an epistemic possibility seriously.

This voluntarist gloss aside, let's see if we can make sense of the change in terms of the goal of admitting as doxastic possibilities only those uneliminated possibilities that are, in some sense, the best candidates for being close alternatives to actuality.

But first let's simplify the example. Many of the processes whose outcomes we have an interest in might be thought of as consisting of a long sequence of individual chancy events that are to some greater or lesser extent independent of one another. In the case of the football game every time the ball is passed or kicked there's a chance that the pass or kick will be on target and a chance that it will go astray. Will the ball be fumbled or caught? If it is fumbled will possession be regained or turned over to the opposing team? And so on and so forth. So let's simplify things by thinking of a case that involves a straightforward random walk. Suppose that you and I each have twenty-four dollar bills and we decide to toss a coin twenty-four times. Every time it lands Heads I will give you a dollar, and every time it lands Tails you will pay a dollar to me.

Here are three possible ways this little game might go:

W_1 : H H H H H H H H H H H H H H H H H H H H H H H H
W_2 : H H H T H T H H H H H T T T T H H H T T H T T H
W_3 : T T T H T H T T T T T H H H H T T T H H T H H T

In W_1 you end up winning \$24 and I am left bankrupt, I don't even have money to catch the bus home and it's a long walk. In W_2 the coin lands Heads 14 times and tails 10 times, I lose only \$4. W_3 is the exact "opposite" of W_2 in the sense that the result of every single toss ends up differently. In this one I win \$4.

One nice thing about this example is that there is a simple and fairly obviously appropriate notion of what should count as the distance between any two possibilities in the space. It is what is known in information theory as the Hamming distance (or signal distance) between two strings of symbols of equal length. The Hamming distance is simply the number of positions at which the corresponding symbols in the two strings differ. It measures, in other words, the number of individual changes that would have to be made to the one string in order to transform it into the other. Thus the distance between W_1 and W_2 is 10, the distance between W_1 and W_3 is 14, while the distance between W_2 and W_3 is 24. W_2 and W_3 are as far apart as any two points in this little possibility space can be.

Now there are many more atomic possibilities besides these three; there are 2^{24} in all. And if were thinking about the example probabilistically then, since the coin is assumed to be fair, all 2^{24} of these atomic possibilities are equiprobable and that's pretty much all there is to be said about it. But what if we frame this problem non-quantitatively, in terms of what I have been calling simple belief? Intuitively, what we would like to see is the elimination of "extreme" possibilities like W_1. But it's hard to find a justification for that. We seem to be stymied by the symmetry of the situation. That is, we seem to have no choice but to retain all of the 2^{24} possible outcomes as serious doxastic possibilities, since to omit some while retaining others would appear to be completely arbitrary. In fact, it may seem that the relationship between our two ways of modeling this example is constrained by the following:

Equiprobability Principle: Whenever S assigns the same probability to two possible states of affairs (according to the degree-of-belief model) then according to the simple-belief model either both states are serious possibilities for S, or neither of them is.

But I think this would be too hasty a capitulation. It ignores some of the conceptual resources available to the purely qualitative model. In fact I think that the Equiprobability Principle should be rejected.

Now the conceptual resource that we have up to now been ignoring involves the agent's values, what he or she cares about. Lets suppose that in the coin tossing example we have been discussing, all I care about is how much money I end up winning or losing in the game, i.e. the total number of times that the coin lands heads, and that it matters to me not at all which particular sequence leads to that total number.

> A *partition* of a possibility space is a set $\mathcal{S}$ of propositions such that the elements of $\mathcal{S}$ are pairwise inconsistent, and the disjunction of all the elements of $\mathcal{S}$ is the whole space.

In our little example, the natural way for me to partition the space of possibilities will have only 25 cells, each corresponding to a certain number of heads (from 0 to 24). Furthermore, some of the possibilities in this equal-value partition will themselves be disjunctions of a vast number of subpossibilities that I don't care to distinguish. In a less toy-like example, like the ice cube case, it would not only be that I don't care to distinguish among the vast quantity of subpossibilities, but that I am actually unable to do so.

> A partition of the space of possibilities is called the *equal-value partition* for an agent when it is the coarsest partition such that for each

> element of the partition the agent is indifferent between all of its sub-possibilities.

Now our epistemic ordering of possibilities is such that X is a closer possibility than Y iff some way that X might be true is closer to actuality than any way that Y might be true. It follows from this that for any suitable measure of distance between possibilities that are themselves disjunctions of other sub-possibilities, i.e. if we want to extend the measure given by the Hamming distance from the atoms of the space to the full Boolean algebra of possibilities, then the way to do it is to say:

> The *distance* between X and Y is minimum distance between x and y, where x ranges over all the ways X might be true, and y over all the ways Y might be true.

And if the distance of a disjunctive possibility from actuality is the minimum of the distance from each of the disjuncts, then "big" disjunctions, measured by number of subpossibilities, will often be better candidates for actuality than the "smaller" ones (like, e.g., the cell of the partition that contains only the single sequence W_1 or the cell of the partition for the ice cube example in which the cube remains unmelted.

> Suppose that $\mathcal{S}$ is a partition of the space of possibilities and X is a possibility that is an element of $\mathcal{S}$. Then say that the *estimated distance from actuality* for X with respect to the partition $\mathcal{S}$ is the mean of the distances from X to each of the elements of $\mathcal{S}$.

The guiding idea now is that an account of simple belief should rank possibilities according to how close they are judged to be to the actual world, rather than by probability. The agent's belief state should include as doxastic possibilities those which are judged to be better canidates for actuality in the following sense:

> If X and Y are two elements of $\mathcal{S}$, say that X is a *better candidate for actuality* than Y iff X's estimated distance from actuality is less than Y's.

Of course this suggestion leaves open the question of how bold or cautious one should be in eliminating those possibilities that are judged not to be close to the way things actually are. That may depend, for example, on pragmatic considerations concerning how high the stakes are and how costly it might be to be wrong.

Here's an illustrative example. The table below shows a smaller version of our random walk example, in which the coin is to be tossed only six times. Again, the agent is assumed to be interested only in the total number of heads that results, and not in the order in which those outcomes occur. So the equi-value partition has seven elements, and the crosstable gives the distances between those possibilities. The number in each row of the final column is the average of the distances listed in that row, and hence gives the estimated distance from actuality of the possibility corresponding to that row.

$\mathcal{S}$	0H	1H	2H	3H	4H	5H	6H	EDA
0H	0	1	2	3	4	5	6	3.00
1H	1	0	1	2	3	4	5	2.29
2H	2	1	0	1	2	3	4	1.86
3H	3	2	1	0	1	2	3	1.71
4H	4	3	2	1	0	1	2	1.86
5H	5	4	3	2	1	0	1	2.29
6H	6	5	4	3	2	1	0	3.00

We see that the best candidate for actuality is the possibility that the coin lands heads three times, and the worst candidates are the extreme ones in which either the coin always lands heads, or always lands tails. It follows from our guiding idea that among the rationally permitted simple belief states for an agent with this equal-value partition, is that which eliminates the set of possibilities {0H, 6H} but that it is not permitted, e.g., to eliminate 5H as a doxastic possibility without also ruling out 0H, 1H and 6H. Again, the main point to appreciate here is that none of this appeals to probabilities.

Of course this example is not completely general. It will not always be the case that the best candidates for actuality are the largest elements of the agent's partition. For example, suppose the partition includes an element which consists of only a sparse number of atoms, spread out in a isolated manner across the whole space. Such an element could be chosen so as to have a lower estimated distance from actuality than a larger possibility clumped in one region of the space. But we have good reason to be suspicious of such gerrymandered partitions, at least in real life examples where the total number of atomic possibilities is vast. It's reasonable to suppose that an agent can only care about differences she can distinguish, and if a possibility consists of a sparse collection of isolated microstates spread though the whole space of possibilities, such a possibility will not be one that an agent can discriminate as being true rather than false. The "clumping" of atoms into partition elements that are connected and convex would seem to be necessary for such possibilities to be graspable or discriminable by the agent in the first place. Just how such a necessary condition should be framed in order to prove the kind of formal result we want is beyond the scope of this paper, but it is my hope that something like the following will turn out to be true.

Conjecture: If $\mathcal{Z}$ is the equal-value partition of some possibility space [that satisfies some fairly general and natural looking conditions] and X and Y are any two elements of $\mathcal{Z}$, then X is a better candidate for actuality than Y iff X has more atoms than Y has.

7. Conclusion

If something like the story I have just sketched can be made to work, then I think we have the beginnings of an autonomous normative account of simple belief that will underwrite the claims I made earlier in discussing how the simple theory might approach the statistical examples.

It is worth drawing attention, in conclusion, to the features that are doing the work. The method sketched above yields an ordering of better and worse candidates for closeness, but doesnt tell us at which point along that spectrum to draw the line between the serious possibilities and the rest. Here I presume, pragmatic considerations should come in to play: How high are the stakes?

But pragmatic considerations are playing another more subtle and interesting role. I don't want to get things really wrong, but if I do get it really wrong, then I want to be wrong in ways that I dont care about. So by limiting attention to the equivalue partitions we have effectively weighted the estimated distances by the cost of being wrong. This is how the symmetry was broken and how we got something non-trivial.

If I right about this, epistemology may have another pragmatic aspect of a hitherto unnoticed kind.

References

[1] Collins, John, 2006. ‘Lotteries and the Close Shave Principle’. In Stephen Hetherington (ed.) *Aspects of Knowing: Epistemological Essays*, Elsevier 83–96.

[2] DeRose, Keith, 1999. ‘Can It Be That It Would Have Been Even Though It Might Not Have Been?’, *Noûs*, (Supplement: Philosophical Persepctives 13, Epistemology) 33: 385–413.

[3] Easwaran, Kenny, 2016. ‘Dr. Truthlove or: How I Learned to Stop Worrying and Love Bayesian Probabilities’, *Noûs* 50: 816–853.

[4] Foley, Richard, 1993. *Working Without a Net*, Oxford University Press.

[5] Gärdenfors, Peter, 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, Bradford Books, MIT Press.

[6] Hawthorne, John, 2004. *Knowledge and Lotteries*, Oxford, Clarendon Press.

[7] Kripke, Saul, 2011. ‘Nozick on Knowledge’, in *Philosophical Troubles: Collected Papers, Vol. 1*, Oxford University Press,162–224.

[8] Leitgeb, Hannes, 2013. ‘Reducing Belief Simpliciter to Degrees of Belief’. *Annals of Pure and Applied Logic* 164: 1338–1389.

[9] Leitgeb, Hannes, 2014. ‘The Stability Theory of Belief’. *Philosophical Review* 123: 131–171.

[10] Leitgeb, Hannes, 2017. *The Stability of Belief: How Rational Belief Coheres with Probability*, Oxford University Press.

[11] Lewis, David, 1974. ‘Radical Interpretation’, *Synthese*, 23: 331–344.

[12] Lewis, David, 1996. ‘Elusive Knowledge’, *Australasian Journal of Philosophy* 74:4, 549–567.

[13] Lewis, Karen, 2016. ‘Elusive Counterfactuals’, *Noûs* 50: 2, 286–313.

[14] Lin, Hanti, and Kevin T. Kelly, 2012. ‘A Geo-Logical Solution to the Lottery Paradox’. *Synthese* 186: 531–575.

[15] Nozick, Robert, 1981. *Philosophical Explanations*. Harvard University Press.

[16] Smith, Martin. 2010. 'What Else Justification Could Be'. *Noûs* 44: 10–31.

[17] Smith, Martin. 2016. *Between Probability and Certainty: What Justifies Belief*. Oxford University Press.

[18] Sosa, Ernest, 1999.'How to Defeat Opposition to Moore'. *Noûs*, (Supplement: Philosophical Perspectives, 13, Epistemology) 33: 141–153.

[19] Stalnaker, Robert, 1984. *Inquiry*. Bradford Books, MIT Press.

[20] Vogel, Jonathan, 1990. 'Are There Counterexamples to the Closure Principle?'. In M. Roth and G. Ross (eds), *Doubting: Contemporary Perspectives on Skepticism* 13–27) Kluwer.

[21] Weatherson, Brian, 2008. 'Deontology and Descartes's Demon'. *The Journal of Philosophy*, Special Issue on Epistemic Norms, Part 1. 105: 9, 540–569.