KNOWLEDGE FROM MULTIPLE EXPERIENCES

Abstract

This paper models knowledge in cases where an agent has multiple experiences over time. Using this model, we introduce a series of observations that undermine the pretheoretic idea that the evidential significance of experience depends on the extent to which that experience matches the world. On the basis of these observations, we model knowledge in terms of what is likely given the agent's experience. An agent knows p when p is implied by her epistemic possibilities. A world is epistemically possible when its probability given the agent's experiences is not significantly lower than the probability of the actual world given that experience.

1 Introduction

This paper models knowledge in cases where an agent has multiple experiences over time. To sharpen the discussion, we focus on a specific kind of case where an agent inquires about some particular quantity of an object, such as its temperature, volume, color, or whatever. We assume this quantity does not change its value over time, and the agent knows this. The agent performs a series of experiments, and receives information about the object. In each experiment, the object appears to have a certain amount of the relevant quantity.[1] Our question is how an agent's knowledge about the real value of a quantity is affected by multiple appearances of the quantity's value.

Our paper contributes to a growing family of papers that offer fairly precise models of perceptual knowledge, including Williamson 2013a, Weatherson 2013, Goodman 2013, Cohen and Comesaña 2013, Williamson 2014, Carter 2019, Dutant and Rosenkranz 2019, Carter and Goldstein 2021, and Goldstein 2021. For closely related models of statistical knowledge, see Dorr et al. 2014 and Goodman and Salow 2018. Like these papers, our models have a number of simplifying assumptions, which we flag throughout.[2] For important work generalizing these assumptions to a wider family of cases, see Beddor and Pavese 2019. Existing efforts in this literature to model perceptual knowledge focus on an agent who has a single experience. Our paper shows how the case of multiple experiences raises interesting issues that might otherwise be overlooked. Throughout, we focus on the case of perceptual knowledge. But our discussion can also apply to other kinds of knowledge, such as testimony and memory.

[1] Here, as is standard in this literature, we assume that experiences represent the object as having a precise value. See Williamson 2013a for discussion.

[2] As in economics, the hope is that the models will provide insight into the phenomena that it models; not that it will capture exceptionless generalizations true of the target phenomenon. As Williamson 2017 notes: 'The traditional philosopher's instinct is to provide counterexamples to refute the simplifications and idealizations built into a model, which rather misses the point of the exercise... what defeats a model is not a counterexample but a *better model*, one that retains its predecessor's successes while adding some more of its own.' (Williamson 2017, p. 9)

Most often, we consider a simple case where a creature observes the fixed temperature of the water in a bowl. At a series of times, the creature dips her hands in the bowl and has an experience of the temperature. At some times, the object appears colder; at others, warmer. Each time the temperature appears a certain way, the creature learns something about what the temperature is really like.

Our preferred model is inspired by several striking facts. First, DEFEAT: perceptual knowledge can be defeated by the appearances. Imagine the agent dips her hands in the bowl. The water's temperature may be cold, warm, or hot. Imagine the temperature is warm and appears to be warm. In this case, the agent may know a lot about the temperature, ruling out that it is hot. But now imagine the agent dips her hands again and the water appears hot. After having both experiences, the agent may know less than she did when she only experienced the temperature as warm. In particular, she may no longer be able to rule out that the water is hot.

Our second striking fact (MONOTONICITY) is that sometimes an agent learns more from a new appearance with the same content as an experience she's already had. Imagine the agent dips her hands three times in the bowl of water. Each time, the water appears warm. Suppose it is warm. After her first experience, the agent knows something about the temperature. After her second experience, the agent knows more. After the third experience, she knows the most. Repetitions of matching experiences can produce increases in knowledge.

Our third striking fact (DIVERGENCE) is that by contrast with the case of a single appearance, multiple appearances which diverge from reality can produce more knowledge than multiple appearances which match reality.A naive conception of evidential support ties the strength of evidence to the match between perceptual content and belief. When our perceptual experiences match reality, we know more than when they do not. But now suppose that the water in the bowl is either moderately cold, warm, or moderately hot. Suppose that when the water is moderately cold, it is quite likely to appear moderately cold, almost as likely to appear warm, and very unlikely to appear moderately hot. When the water is warm, it is quite likely to appear warm, and almost as likely to appear moderately cold or moderately hot. When the water is moderately hot, it is quite likely to appear moderately hot, almost as likely to appear warm, and very unlikely to appear moderately cold.[3] Now imagine the agent dips her hands in the bowl twice, and the water appears first moderately cold and then moderately hot. In this case, the agent can come to know that the water is warm. Apprised of this probabilistic structure, she could reason as follows: if it were warm, then it could easily happen that the water appeared moderately cold and then moderately warm. But if the bowl were moderately cold, it would have been very unlikely for it to appear moderately hot; and vice versa. By contrast, now imagine the agent dips her hands in a bowl of warm water twice, and the water appears warm both times. Now the agent doesn't come to know anything.

[3]Here, we don't assume that this setup correctly describes the actual facts about human ability to perceive water temperature. We are simply imagining a possible creature who perceives water temperature in this way.

For all she knows, the water could be moderately cold, warm, or moderately hot. The upshot is that what we learn from experience is not a matter of matching up the content of experience to the world. Experiences that diverge from the world can produce more knowledge than experiences that match the world.

If evidential support isn't about match, what is it about? We replace match with probability. What we learn from our experiences doesn't depend on whether the real values match our experience. Instead, what we learn depends on whether the real values are probabilified by our experience. Perceptual knowledge must respect what is probable given the appearances. If one possibility is consistent with an agent's knowledge and another is not, then the former possibility is at least as likely given one's appearances as the latter.

Our final striking fact (NON-CONVEXITY) concerns this principle. The observation is that our knowledge of a given quantity is not always convex. That is, perceptual evidence may put us in a situation where we know the temperature is in one of two regions, even though we have ruled out the temperature being in an intermediate region. Again, this is a phenomenon that arises once we allow an agent to have multiple appearances instead of a single one. Imagine yet again that an agent dips her hands in a bowl of water. The water's temperature may now be very cold, warm, or very hot. Suppose when the water is very cold, it is extremely likely to appear very cold, slightly likely to appear warm, and slightly less likely to appear very hot. When the water is warm, it is extremely likely to appear warm, slightly likely to appear very cold, and slightly likely to appear very hot. When the water is very hot, it is extremely likely to appear very hot, slightly likely to appear warm, and slightly less likely to appear very cold. Now imagine the agent dips her hands in a bowl of very cold water twice, and the water appears first very cold and then very hot. In this case, the agent can come to know that the water is either very cold or very hot, but that it is not warm. Apprised of this probabilistic structure, she can reason as follows: if it were warm, then there were two freak accidents in which it appeared to be very different than it was. But if it were very cold or very hot, there was only one freak accident. Below, we'll see that several theories of perceptual knowledge go wrong here, predicting that perceptual knowledge is convex. In doing so, these theories disrespect the probabilities.[4]

Summarizing, we have outlined four facts about knowledge and appearance. Defeat: new appearances can defeat the knowledge gained from appearance. Monotonicity: multiple matches between appearance and reality can provide more knowledge than a single match. Divergence: appearances that diverge from reality can provide more knowledge than appearances which match reality. Non-convexity: appearances can provide knowledge of a real value that is not convex.

[4]Divergence and non-convexity can trivially obtain in situations where the agent possesses unusual background information. Imagine for example that the agent discovers she is wearing vision distorting goggles, or that the agent learns that a particular real value inside of a range cannot be actual. Crucially, though, we are modelling cases where the relevant effects arise even without unusual background knowledge. Moreover, in our examples and in the theory we develop, these features only emerge when the agent receives multiple pieces of evidence.

On the basis of these facts, we model perceptual knowledge in terms of what is likely given the appearances. An agent knows p when p is implied by her epistemic possibilities. A world is epistemically possible when its probability given the appearances is not significantly lower than the probability of the actual world on the appearances. (See Goodman and Salow 2021 and Hong In Preparation for other recent work that models knowledge in terms of probability.)

Here's our model of cases like the above. The agent receives a series of appearances. These appearances count strongly in favor of real values that are similar to the apparent values. The appearances count much less strongly in favor of more distant real values. In the cases we're most interested in, the probabilities are proportionate to the likelihood of having these experiences, given the real value.

As with all the competitor models we consider, an agent's knowledge depends on what the real value actually is. But we interpret this dependence probabilistically. In good cases, the real value is strongly supported by the appearances, and the agent knows quite a lot. In that case, most real values are probabilified by the appearances much less than the actual real value. So the agent can rule out most real values. By contrast, in bad cases the agent's experience does not probabilify the actual real value. It may even decrease the probability of the real value. In this case, the agent knows very little, because most real values are better confirmed by the appearances than the actual real value, and so most real values are epistemically possible.

§2 introduces an elegant model of knowledge from Williamson 2013a. This model makes predictions about what an agent knows when things appear a certain way. But this model does not cover cases of multiple appearance. §3 introduces a simple but flawed model of multiple appearance where an agent's knowledge grows intersectively. §4 introduces another model which allows knowledge defeat and explains how new appearances with the same content as previous experiences can still provide new evidence. But this model also goes wrong, because it models evidential support in terms of match instead of probability. §5 develops our own theory. §6 clarifies how our model relates to other notions in the literature. Crucially, in our own theory what an agent knows need not have a probability of 1 on her evidence.

Before we begin, a word of clarification. Throughout, our task is to model what an agent knows after they have a series of experiences. Our discussion appeals to facts about various probabilities of hypotheses about reality, conditional on facts about the appearances. We rely on these facts to model what an agent is in a position to know on the basis of her evidence, without any kind of assumption that the agent explicitly reasons with these facts as premises. Indeed, we are not committed to viewing this process as any kind of inferential one, in which an agent believes that the world appears a certain way, and deduces facts about reality on this basis.

On the other hand, we do not want to take a stand on the vexed question of exactly where perception ends and inference begins. Some may respond to our motivating observations above that while possible, these observations involve abductive rather than perceptual knowledge. Ultimately (as we alluded

to earlier), we think the positive theory we develop below is promising as a model not only of what an agent can know on the basis of how things appear, but also of what an agent can know on the basis of statistical samples, and evidence more generally. Whether perceptual or not, some key features of our account are that the relevant agents can gain knowledge without engaging in conscious reasoning, and without having prior beliefs or propositional evidence about how reality depends on appearance.

2 Appearance and reality

We start by reviewing a simple model of perceptual knowledge from Williamson 2013a.[5] Here an agent has a single experience, and learns something about reality on this basis. A possible world is a pair of two quantities r and a, measuring the real and apparent value of a quantity. In our case, that quantity is degrees Fahrenheit. In the 'bad case' $(50, 60)$, the temperature is 50 degrees but appears 60 degrees. In the 'good case' $(50, 50)$ reality matches appearance. An agent knows p at (r, a) just in case p is true at every possibility (r', a') that is epistemically accessible from (r, a). E represents epistemic accessibility.

For the purpose of offering a simple model of perceptual knowledge, we assume throughout that appearances are luminous. When the temperature appears to be a, the agent knows this. So $a = a'$ whenever $(r, a)E(r', a')$. Epistemic accessibility is then determined by the distance between reality and appearance. A real value is accessible when its distance from appearance does not significantly exceed the actual distance between reality and appearance. The constant m represents the degree to which the distance between reality and appearance may exceed actuality. $|r - a|$ measures the distance between reality and appearance. A value r' is accessible from (r, a) just when the distance between r' and a is less than the sum of $|r - a|$ and m.

(1) $$E(r, a) = \{(r', a) : |r' - a| \leq |r - a| + m\}$$

In Figure 1, arrows depict the lower and upper bounds of accessibility from various real values, when the margin for error is 5. In the good case $(50, 50)$ where reality matches appearance, the agent knows a lot about the real value of the temperature: that it is within $[45, 55]$. But in the bad case $(56, 50)$ where reality does not match appearance, the agent knows much less: only that it is within $[39, 61]$.

Our model generalizes (1) to the case of multiple appearances. But before this, we consider two motivations for the model above. Each motivation is

[5] Williamson 2013a is primarily concerned with providing an internalist model of Gettier cases. But the models there have a variety applications outside of this internalist setting. For more discussion of these models and their varied significance, see the citations above.

As a simplifying assumption, Williamson 2013a's models also assume that the agent knows everything that they are in a position to know. Obviously the real world isn't like this; often, for example, we don't even believe that which we are in a position to know. Readers that dislike this aspect of the idealization can treat the model as a model of what we are in a position to know rather than as a model of knowledge.

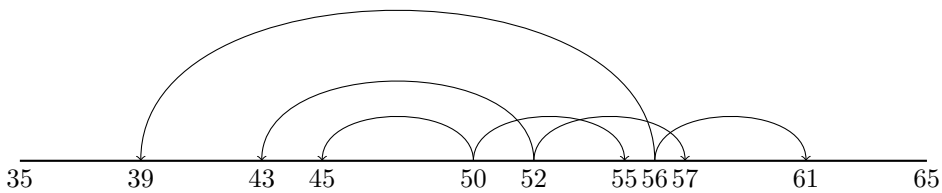

Figure 1: $a = 50$, $m = 5$

relevant to the case of multiple appearances.

First, the model above can be understood in terms of relative normality. On this proposal, an agent knows p at (r, a) when p is true at any world that is almost as normal as (r, a). $(r', a') < (r, a)$ represents that (r', a') is less normal than (r, a). $(r', a') \ll (r, a)$ represents that (r', a') is significantly less normal than (r, a).[6] (r', a') is almost as normal as (r, a) just in case $(r', a') \not\ll (r, a)$. Then $E(r, a)$ is the set of worlds (r', a) that are almost as normal as (r, a).

(2) $E(r, a) = \{(r', a) : (r', a) \not\ll (r, a)\}$

This theory agrees with the one above, provided that we connect normality to the distance between reality and appearance. In particular, say one world is significantly less normal than another just in case the distance between reality and appearance in the first world is m greater than the corresponding distance in the second world.

(3) $(r', a) \ll (r, a)$ iff $|r' - a| > |r - a| + m$

As Beddor and Pavese 2019 observes, standards for normality are naturally interpreted relative to various tasks. This is why the normality ordering can supervene on the distance between reality and appearance. Other aspects of the world, such as whether there is a flying spaghetti monster, may contribute to the general normality of that world. But they aren't relevant to the task relative notion of normality we are modeling, which is forming a belief about a certain quantity on the basis of its appearance.[7][8]

[6] See Goodman and Salow 2018 (although note that their notation for the normality ordering is reversed).

[7] Even here we are simplifying. Of course some strange features of the environment might be relevant to the normality of forming a belief about a certain quantity on the basis of its appearance. For example, it would be abnormal for the light from an object to take a detour through Beijing on its way to our eyes. Like Williamson 2013a, we abstract from this issue in our model.

[8] Arguably a task relative conception of normality also provides some motivation for Appearance Luminosity. The task that is relevant to our model is forming a belief about a certain quantity on the basis of its appearance. For the purposes of evaluating this task, we only consider worlds where the quantity has that perceptual appearance.

On the other hand, we may for certain purposes want a more fine-grained conception of methods. Even holding fixed the appearance, there might be various methods used to form

Below, we consider how to extend this conception of normality to the case of multiple appearances. Ultimately, our own theory understands normality in terms of probability rather than distance.

Second, the theory above can be understood in terms of principles governing the epistemology of reality and appearance. First, Williamson suggests that knowledge of reality is subject to margin for error. Whenever the real value is r, the agent can know at most that the real value is within m of r. More precisely, let $\mathsf{Real}(r, a)$ represent the set of accessible real values at (r, a), or $\{r' : (r', a) \in E(r, a)\}$.

(4) **Margin for error.** $\forall r, a : [r - m, r + m] \subseteq \mathsf{Real}(r, a)$

Second, Williamson suggests that knowledge from appearance is symmetric around appearance. When the temperature appears a certain value a, all you know is that the temperature is within some distance x of that value.

(5) **Appearance centering.** $\forall r, a \exists x : \mathsf{Real}(r, a) = [a - x, a + x]$

One might interpret this as an evidential constraint on knowledge, saying that our evidence from an appearance never favors values on one side of that appearance over values on the other side.

Let Appearance Luminosity be the principle that if $(r, a)E(r', a')$, then $a = a'$. Then the accessibility relation in (1) is the smallest relation that satisfies Margin for Error, Appearance Centering, and Appearance Luminosity.[9] On this theory, every agent knows as much as possible, consistent with these three principles. When we move to the setting of multiple appearances, we consider which of these principles to accept, and how exactly to formulate them. We'll argue that Appearance Centering and Margin for Error are both challenged in cases of multiple appearances. Our failures of convexity from the Introduction produce violations of Appearance Centering. The problem for Margin for Error concerns another case from the Introduction, where a series of perfect matches provides more information than a single perfect match between appearance and reality.

3 Intersectivism

One of our main goals is to give a model of how an agent's knowledge changes over time in response to new appearances. To explain this, we enrich the framework above to allow the world to appear in multiple different ways. A world is not just a pair of a real value and an apparent value (r, a). Rather, we let a world be a pair of a real value r and a sequence of apparent values $A = \langle a_1, \ldots, a_n \rangle$, where the value of n can differ between worlds. We're primarily interested in cases where an agent has a series of experiences over time. In those cases, we let

beliefs based on the appearance, and one might want one's conception of knowledge to be somehow sensitive to which method is in play. The literature that we are contributing to abstracts away from questions of fine-grained methods, and this is one of the ways in which our model like theirs' makes simplifying assumptions.

[9] See Goodman 2013.

1 through n be a series of times, and when $t = i$ the agent has appearance a_i. But our framework could also make sense of multiple pieces of evidence at the same time.

We introduce a time-relative accessibility relation E^t, which describes what an agent knows after receiving appearances up to time t. So $E^2(50, \langle 40, 60, 70\rangle)$ describes what the agent knows if the real value is 50 and the agent has experienced apparent temperatures of 40 and 60 degrees. $E^3(50, \langle 40, 60, 70\rangle)$ describes what the agent knows if the real value is 50 and the agent has experienced apparent temperatures of 40, 60, and then 70 degrees.

The simplest way to extend the theory above to multiple appearances is to let the agent's knowledge grows conjunctively with each new appearance. On this proposal, we find what an agent would know if she had each individual experience, and then combine all of these claims. Where $E(r, a)$ is as in (1), $A = \langle a_1, \ldots, a_n\rangle$, and $E^t(r, a) = \{(r', A') : (r', a'_t) \in E(r, a)\}$:

(6) **Intersectivism.** $E^t(r, A) = \bigcap\{E^i(r, a_i) : i \leq t\}$

Equivalently, (r', A) is epistemically accessible from (r, A) at t just in case (1) predicts that (r', a_1) is accessible from (r, a_1), ..., and (r', a_t) is accessible from (r, a_t).

Intersectivism fails to predict any of our four observations from the Introduction. We focus on the first two: that new appearances can defeat the knowledge gained from appearance; and that multiple matches between appearance and reality can provide more knowledge than a single match.

Let's start with knowledge defeat. To see the problem, consider Figure 2. Imagine $m = 5$, $r = 50$, and the temperature appears first 45 and then 55 degrees. Intersectivism predicts that at t_2 the agent knows that the temperature is within 45 and 55 degrees. If the temperature had only appeared 45 degrees the agent would have known it was between 35 and 55 degrees; and if it had only appeared 55 degrees the agent would have known it was between 45 and 65 degrees. Now suppose instead that the second appearance had been very unreliable. Bizarrely, the agent would still learn exactly as much: that the temperature is between 45 and 55 degrees.

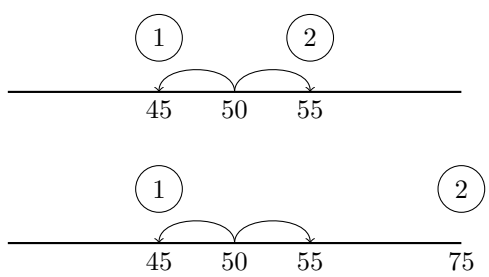

Figure 2: $r = 50, A = \langle 45, 55\rangle$, $A' = \langle 45, 75\rangle$, $m = 5$

Intersectivism goes wrong because it predicts that defeat is impossible. In

the second example, the agent is initially able to rule out values higher than 55. But once she has an unreliable appearance of 75, this knowledge should be defeated.

Intersectivism rules out knowledge defeat. It says that an agent's knowledge can only grow as they have more experiences. More precisely:

(7) **No Defeat.** If $t \leq t'$, then $E^t(r, A) \supseteq E^{t'}(r, A)$.

We now show that No Defeat is absurd, because it is Kierkegaardian, rewarding epistemic leaps of faith.[10]

To see the problem, first suppose that evidence is commutative. What you know at a time doesn't depend on the order in which you received your evidence up to that time. More precisely, say that A and A' are commutative at t ($A \sim_t A'$) just in case A and A' contain exactly the same appearances, in possibly distinct order, up until time t. For example, $\langle 80, 70, 60\rangle$ and $\langle 60, 70, 80\rangle$ are commutative at t_3 because they contain the same appearances in different order. Then:

(8) **Commutativity.** If $A \sim_t A'$, then $E^t(r, A) = E^t(r, A')$.

Second, suppose that our knowledge of the real value now does not depend on our future experiences. More precisely, say that A and A' are alike up to t ($A \approx_t A'$) just in case $\langle a_1, \ldots, a_t\rangle = \langle a'_1, \ldots, a'_t\rangle$. Let $\mathsf{Real}^t(r, A)$ represent the set of possible real values at t (the set $\{r' : \exists A'.(r', A') \in E^t(r, A)\}$). Then we say that any two worlds that share the same real value and are alike up to t have the same knowledge about that real value at t.[11]

(9) **No Future Dependence.** If $A \approx_t A'$, then $\mathsf{Real}^t(r, A) = \mathsf{Real}^t(r, A')$.

The combination of Commutativity, No Future Dependence, and No Defeat is absurd, because it requires epistemic leaps of faith. Imagine an agent receives a bunch of misleading evidence, and then receives a small bit of good evidence. An epistemic leap of faith is a case where the agent trusts the final small bit of good evidence, believing that the real value is concentrated around that evidence. A theory is Kierkegaardian when it rewards an epistemic leap of faith, by saying

[10] The line of thought that follows was suggested independently in conversation by both Jeremy Goodman and Yoaav Isaacs.

[11] There's a generalization of No Future Dependence that is obviously hopeless. According to the generalization, what we know now is independent of our future evidence. One way to see that this is obviously wrong is to consider propositions about our future evidence. I may now know that I am not about to have an experience as of a polka-dotted elephant dancing the polka. But whether I know depends on whether I will in fact have the experience in the future.

On the other hand, we admit that No Future Dependence is not completely sacrosanct even when restricted to knowledge of current real values. As a test case, suppose one has an experience of a barn. One might think that whether one knows now that it is a barn depends on whether in the near future there will be fake barns constructed that produce misleading experiences. On the other hand, its not clear how convincing this way of pushing back really is. Suppose someone looks at a working watch, and says they know it is 7 o'clock. Can I falsify their assertion two minutes later by showing them a stopped clock that says it is 7:10. Can one really retroactively prevent knowledge through such performances?

that epistemic leaps of faith can produce knowledge.

Consider the case $(20, \langle 20 \rangle)$, where the agent receives good evidence that the real value is 20, and never receives any misleading evidence. The agent should get to know a lot about the real value. Otherwise, skepticism threatens.

Now consider the world $(20, \langle 20, 80, 80, 80 \rangle)$ where the agent initially receives the same good evidence, but then receives lots of misleading evidence. No Future Dependence implies that at t_1, after receiving the good evidence, the agent knows the same thing about the real value as she knows at $(20, \langle 20 \rangle)$. Further, No Defeat implies that she knows at least as much at t_4 as she does at t_1.

Finally, consider the Kierkegaardian world $(20, \langle 80, 80, 80, 20 \rangle)$, where the agent receives the evidence in opposite order, first receiving huge amounts of misleading evidence about the real value, and then receiving some good evidence. Commutativity implies that she knows the same thing at t_4 in $(20, \langle 80, 80, 80, 20 \rangle)$ and $(20, \langle 20, 80, 80, 80 \rangle)$. The result is that if we hold fixed our initial assumptions, we must either be skeptics about the agent's knowledge of the real value at t_1 in $(20, \langle 20 \rangle)$, or allow that the agent has lots of knowledge of the real value at t_4 in the leap of faith world t_4 in $(20, \langle 80, 80, 80, 20 \rangle)$. A natural response to this argument is to deny No Defeat. Since Intersectivism validates No Defeat, we need a new model.

Intersectivism sometimes predicts you know too much after receiving new evidence. There are other cases where Intersectivism predicts you know too little. Return to our second fact from §1, that two perfect matches with reality should provide more knowledge than one. Consider Figure 3. In this case, the agent observes the temperature twice, and both times receives a perfectly accurate appearance.

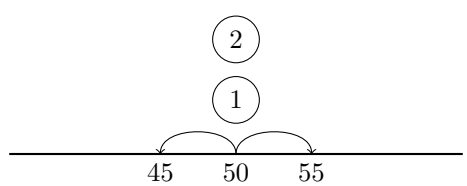

Figure 3: $r = 50, A = \langle 50, 50 \rangle$, $m = 5$

In this example, the agent knows more at t_2 than at t_1. But Intersectivism denies this. Each appearance individually produces the same amount of knowledge: that the real value is between 45 and 55 degrees. When we conjoin this knowledge, we produce the same range of ignorance.

This concludes our main objections to Intersectivism. But to introduce our next theory, it is worth considering a few structural features of Intersectivism. First, Intersectivism is incompatible with the normality theory of knowledge in (2), at least supposing that any two worlds are comparable with respect to normality. Compare $low = (20, \langle 20, 80 \rangle)$ and $high = (80, \langle 20, 80 \rangle)$. Intersectivism predicts that in *low* the agent knows the real value is low (within m of 20), and in *high* the agent knows the real value is high (within m of 80). But now assume

that either *high* is at least as normal as *low*, or *high* is less normal. The normality theory predicts that if *high* is at least as normal as *low*, then in *low* the agent doesn't know the real value is low, since for all she knows she is in *high* where the real value is high. Similarly, if *high* is less normal than *low* then in *high* the agent doesn't know the real value is high, since for all she knows she is in *low* where the real value is low. So Intersectivism is incompatible with the normality theory. In the rest of this paper, we hold fixed the normality theory and offer various conceptions of normality that produce various theories of knowledge.

Additionally, Intersectivism doesn't seem to validate Appearance Centering. Once we allow that the world can appear multiple different ways to the agent, we need a more expansive notion of Appearance Centering. At first glance, there's no version of Appearance Centering that holds in Figure 2. In this case, the apparent values are at 45 and 75, but the agent knows the real value is between 45 and 55.[12]

4 Centrism

We now turn to a model that explains our first two striking facts, regarding knowledge defeat and multiple perfect matches, but struggles with our third and fourth facts.

To develop this model, we return to the normality theory of knowledge, where a world is possible just in case it is not significantly less normal than actuality.

(10) $E(r, A) = \{(r', A) : \ (r', A) \not\ll (r, A)\}$

We develop a new model of knowledge by offering a new theory of normality. As before, we understand normality in terms of the total amount of error at a world, measured by the distance between reality and appearance.

(11) $(r, A) \leq (r', A)$ iff $\mathsf{Error}(r, A) \geq \mathsf{Error}(r', A)$

But in the setting of multiple appearances, we measure distance by appealing to the intuitive notion of the center of a sequence of appearances. So if $A = \langle 20, 80 \rangle$, let $\mathsf{Center}(A) = 50$. Generalizing, let the center of a sequence of appearances be its mean. Then the error at a world is the distance between reality and the center of appearances, multiplied by the number of appearances (this multiplication helps model the epistemic effect of multiple matching appearances). Then an agent's epistemic possibilities are those where the error is no more than m greater than the actual error. Let A^t be the initial segment of A up until t.

We now consider a model of knowledge we call 'Centrism'. Centrism says the epistemically possible real values are those whose error does not exceed the sum of the actual error and the fixed margin m.

(12) **Centrism.**

[12] On the other hand, Intersectivism does validate Margin for Error. Suppose the real value is 50. Then no individual appearance will allow the agent to learn that he is within m of 50. It follows that the sequence of appearances together does not allow the agent to learn this.

a. $\mathsf{Center}^t(A) = \frac{\sum_{i=1}^{t} a_i}{t}$
b. $\mathsf{Error}^t(r, A) = t \times |r - \mathsf{Center}^t(A)|$
c. $E^t(r, A) = \{(r', A^t) : \mathsf{Error}^t(r', A) \leq \mathsf{Error}^t(r, A) + m\}$

When degrees of normality are identified with amounts of error, and m is the threshold for a significantly lower amount of normality, this definition agrees with (2). Equivalently, this model implies that $E^t(r, A) = \{(r', A) : |r' - \mathsf{Center}^t(A)| \leq |r - \mathsf{Center}^t(A)| + \frac{m}{t}\}$.

As will become clear, we don't think Centrism is a very promising theory (and would likely be anathema to statisticians). However, a surprising number of philosophers we've spoken to were initially tempted towards it, and moreover it turns out to be an excellent foil for introducing some of the key concepts and constraints that guide our own theory.[13]

To illustrate, consider a case where the appearances have a clear center matching reality. Centrism reasonably predicts that such a body of evidence produces a lot of knowledge.

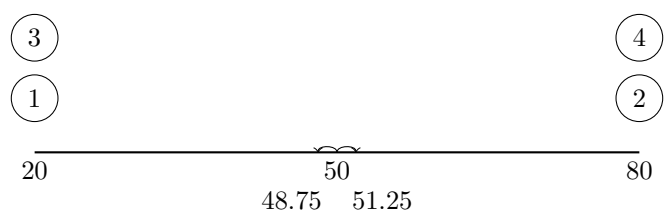

Figure 4: $r = 50, A = \langle 20, 80, 20, 80\rangle$, $m = 5$

The center of $A = \langle 20, 80, 20, 80\rangle$ is 50. Reality matches this center, and so the error is 0. The margin with four appearances is 1.25. So at t_4 in $(50, \langle 20, 80, 20, 80\rangle)$ the agent knows that the real value is in [48.75, 51.25]. This is even more knowledge than at $(50, \langle 50\rangle)$, where the agent knows only $[45, 55]$. By contrast, a theory that simply summed up the individual distance between reality and each appearance would predict that $(50, \langle 20, 80, 20, 80\rangle)$ has a great deal of error, and so would predict that the agent would know very little.

We now apply Centrism to our four initial facts (defeat, monotonicity, divergence, and non-convexity). First, Centrism explains knowledge defeat. For a simple case, consider Figure 5. The real value is 50 and the apparent values are first 50 and then 75, with a margin of 5. Centrism predicts that at t_1 the agent knows the real value is between 45 and 55. By contrast, at t_2 the agent knows

[13] A more promising version of Centrism would perhaps replace our crude measure of distance above with squared distance (cf. the superiority of the Brier score over cruder measures of accuracy in the case of credences). On this proposal, the error at a world is the sum of the squared distance between the real value and each appearance. At each world, the agent knows that the error is not significantly less than actual. Throughout, we restrict attention to the simpler definition of error in the main text because our two main challenges to this theory (regarding DIVERGENCE and NON-CONVEXITY) also affect the proposal that relies on squared distance.

that the real value is between 47.5 and 82.5. So at t_1 the agent knows that the real value is not 80, and at t_2 the agent loses this knowledge.

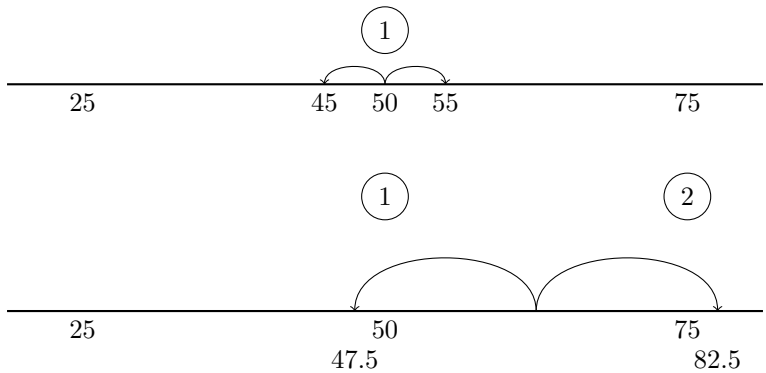

Figure 5: $r = 3$, $A = \langle 3, 5\rangle$, $m = 2$

Second, Centrism predicts that multiple matches between appearance and reality can provide more knowledge than a single match. To see why, note that Centrism validates a version of Margin for Error where the margin decreases with the number of appearances:

(13) **Shrinking Margin for Error** $\forall(r, A) : [r - \frac{m}{t}, r + \frac{m}{t}] \subseteq \mathsf{Real}^t(r, A)$

In a simple good case $(50, \langle 50\rangle)$ the agent knows less at t_1 than in a richer good case $(50, \langle 50, 50\rangle)$ at t_2 where even more appearances match reality. In the former case, the agent knows only that the real value is in $[45, 55]$; in the latter case the agent learns that it is within $[47.5, 52.5]$. This difference is motivated by normality and error. In the second case, a real value of 47.5 produces just as much error as a real value of 45 in the first case.

Centrism not only validates a version of Margin for Error. It also validates Appearance Centering. Centrism implies that epistemic possibility is centered around the center of appearances:

(14) **Mean Appearance Centering** $\forall(r, A)\exists x : \mathsf{Real}^t(r, A) = [\mathsf{Center}(A) - x, \mathsf{Center}(A) + x]$

Centrism is the smallest relation satisfying Appearance Luminosity, Shrinking Margin for Error, and Mean Appearance Centering.[14] In this way, Centrism naturally generalizes the theory in (1).

[14]Suppose there is a smaller relation R satisfying these principles. If R is smaller than E, then there is some (r, A) where $R(r, A) \subset E(r, A)$. Since R satisfies Mean Appearance Centering, we know that $R(r, A)$ is centered on the center of A, just like $E(r, A)$. But since $R(r, A) \subset E(r, A)$, we know that Shrinking Margin for Error fails. Suppose for simplicity that r is above or at the center of A. Then the possible real values extend above r by $\frac{m}{t}$. Since $R(r, A) \subset E(r, A)$, we know that the R-accessible real values extend above r by less than $\frac{m}{t}$. So R violates Shrinking Margin for Error.

Unfortunately, Centrism faces serious problems. First, there is a basic problem with the concept of a center. Not every quantity supplies a series of values that have an intuitive center. For example, consider Williamson's case of the unmarked clock. Now imagine that one receives the series of appearances $A = \langle 12, 3, 6, 9 \rangle$. The agent's evidence completely circles the clock, and there is no natural notion of a center from which to calculate error.

Centrism faces more serious problems. Unfortunately, Centrism does not explain our last two striking facts (DIVERGENCE and NON-CONVEXITY). Our third fact was that appearances that diverge from reality can provide more knowledge than appearances which match reality. Our fourth fact was that appearances can provide knowledge of a real value that is not convex. Let's consider each in turn.

Fundamentally, Centrism models evidential support in terms of the match between the content of one's experiences and the world. Holding fixed the number of appearances, Centrism implies that whenever your appearances perfectly match reality, you know as least as much as in any other case.[15]

(15) **Match.** If $\forall i \leq t : r = a_i$, then $E^t(r, A) \subseteq E^t(r, A')$

In this way, Centrism consider the best evidence to be evidence whose content perfectly matches reality.

We now object to Match, because it ignores the probabilities of real values conditional on the appearances. From a probabilistic perspective, multiple appearances that diverge from reality may provide more evidence than multiple matching appearances.[16] To see the problem, return to our case from §1. Again imagine an agent dipping her hands in water. The water is either moderately cold, warm, or moderately hot. When the water is moderately cold, it is quite likely to appear moderately cold, almost as likely to appear warm, and very unlikely to appear moderately hot. When the water is warm, it is quite likely to appear warm, and almost as likely to appear moderately cold or moderately hot. When the water is moderately hot, it is quite likely to appear moderately hot, almost as likely to appear warm, and very unlikely to appear moderately cold. Now imagine that the agent dips her hands in a bowl of warm water on two different occasions, and the water appears first moderately cold and then moderately hot. In this case, we think the agent can come to know that the

[15] We can distinguish Match from another principle, Strong Match, which says that an agent knows strictly more in cases of perfect match than in any other case. Centrism validates Match, but invalidates Strong Match. Below, we offer a counterexample to Match and hence also to Strong Match.

[16] There are other potential counterexamples to Match that we abstract from in this paper. Imagine that it is unclear whether the real values are reliably generating appearances. Imagine that there is instead a significant chance that the agent just experiences water as warm, no matter its temperature. In that case, having a series of warm experiences can provide less evidence than having a mixed series of experiences, even if the water is warm. The former series of appearances is consistent with the hypothesis that the water will feel warm no matter its temperature. The latter series is not. This case is similarly a counterexample to the principle that having many experiences that perfectly match reality always provides more knowledge than just one. Have many experiences that are exactly the same could actually provide evidence that one's appearances are unreliable because insensitive.

water is warm. Someone apprised of these probabilities could reason as follows: if it were warm, then it could easily happen that the water appeared moderately cold and then moderately warm. But if the bowl were moderately cold, it would have been very unlikely for it to appear moderately hot; and vice versa. By contrast, now imagine that the agent dips her hands in a bowl of warm water on two different occasions, and the water appears warm both times. In this case, we think that the agent doesn't come to know anything. For all she knows, the water could be moderately cold, warm, or moderately hot. The upshot is that what we learn from our experiences is not a matter of matching up the content of that experience to the world. Experiences that diverge from the world can provide more knowledge than experiences that match the world.

To make these points precisely, we consider an abstract version of this ordinary case. To make the calculations simple, we move to a setting with a finite number of worlds, where the agent learns about a quantity with 5 values. But we could make the same points in a continuous setting.[17] So consider Figure 6. The real value is 3. In the first case, the apparent values are 1 and 5; in the second case, they are 3 and 3, matching reality perfectly.

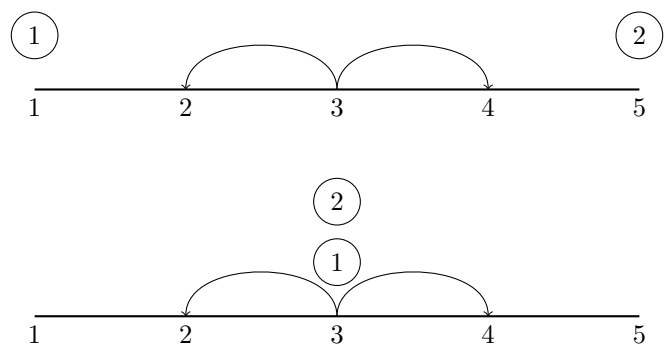

Figure 6: $r = 3, A = \langle 1, 5\rangle, A' = \langle 3, 3\rangle$, $m = 2$

Match and therefore Centrism imply that the latter case generates at least as much knowledge as the former. We reject this claim as a description of the ordinary case above. To see the problem, we need to supply a prior probability function which models the ordinary case above. Throughout, we make several simplifying assumptions about the priors to aid discussion. First, the priors are indifferent about the real values.[18] Second, the priors encode dependencies between real and apparent values. Many of our examples have the following

[17]The reader may worry that our examples require bounded quantities (there is no value below 1). But our observations throughout also apply for example to cases with a circular structure, such as an agent's knowledge of the time on the basis of looking at a clock. Note that even in the case of a single appearance, Appearance Centering would need some qualification with bounded domains, but can be imposed unrestrictedly when the values form a circle, as with the position of a clock hand.

[18]In our own theory, developed later, this assumption will imply that prior to any observations, the agent knows nothing about the real value.

structure: when the temperature is really n degrees, it is quite likely to appear n degrees. It is less likely but still possible for it to appear degrees nearby n. It is extremely unlikely but still possible for the temperature to appear to be degrees far from n. Third, the priors treat appearances at different times as independent conditional on a given real value. Given that the real value is 50, there is no connection between the probability of the first appearance being 50 and the second appearance being 50. We can think of the appearances as like a machine which tracks the real values, but has a certain chance of malfunctioning at any time.[19]

The ordinary case above had a particular pattern of dependence between real and apparent values. The prior we need predicts that when the real value is 1, it is pretty likely that the appearances could either 1, 2, or 3. But it is extremely unlikely that the appearances would be 4 or 5. Similarly for when the real value is 5. By contrast, when the real value is 3, any values in 1 through 5 are fairly likely. Here is one example.

a	r	P
1	1	.075
2	1	.065
3	1	.055
4	1	.003
5	1	.002
1	2	.050
2	2	.055
3	2	.050
4	2	.040
5	2	.005
1	3	.038
2	3	.040
3	3	.044
4	3	.040
5	3	.038

a	r	P
1	4	.005
2	4	.040
3	4	.050
4	4	.055
5	4	.050
1	5	.002
2	5	.003
3	5	.055
4	5	.065
5	5	.075

Given our assumption of indifference over real values, the probability of a real value given an apparent value is proportional to the likelihood of the apparent value given the real value. In addition, the likelihood of producing a sequence of appearances is just the product of the likelihoods of producing each appearance. This produces the following predictions about what is likely given the appearances in each of our two cases, where $P(\cdot \mid A = \langle a_1, \ldots, a_n \rangle)$

[19]This independence assumption is not essential to the description of our cases, or to the positive theory we develop later. We use it here as a simplifying assumption that is appropriate for some range of cases. For some discussion of cases where the assumption is inappropriate, see Garber 1980. Often, in cases where independence fails, an agent learns less from repeated appearances with the same content. For example, imagine an agent who, once she has a certain perception of the water's temperature, will continue to have a similar perception of the temperature upon repeated experiences unless she receives a radically different sensory input. Such an agent will learn less from repeated experiences that feel the same. Similarly, using one thermometer twice instead of using two different thermometers would make a difference to the plausibility of independence.

abbreviates $P(\cdot \mid \{(r', A') : A' = \langle a_1, \ldots, a_n \rangle\})$:

a_1	a_2	r	$P(\cdot \mid A = \langle 1, 5 \rangle)$
1	5	1	.067
1	5	2	.111
1	5	3	.643
1	5	4	.111
1	5	5	.067

a_1	a_2	r	$P(\cdot \mid A = \langle 3, 3 \rangle)$
3	3	1	.233
3	3	2	.193
3	3	3	.149
3	3	4	.193
3	3	5	.233

Now imagine that in the actual world, the real value is 3. Probabilistically, appearances of 1 and 5 give the agent a huge amount of information. Given the prior, these appearances make it very probable that the real value is 3. By contrast, if the agent gets the appearances 3 and 3, she learns way less. These appearances provide little information about reality, because so many different real values could have produced these appearances.[20]

For this reason, the agent knows more in the case of divergence than in the case of match. This example suggests that knowledge isn't just a matter of whether an agent's evidence matches reality. When $A = \langle 3, 3 \rangle$, the agent's appearances perfectly match reality. When $A = \langle 1, 5 \rangle$, the agent's appearances don't match reality at all. But the agent learns way more about reality in the latter case than in the former case. The moral is that when it comes to multiple appearances, 'the good case' is not about appearances matching reality. The crucial feature is the likelihood of receiving various evidence about the world given various hypotheses.

On the basis of this case, we reject Match. What an agent knows does not depend in any simple way on the extent to which the real values match her appearances. Rather, they depend on the probabilities. To see this point precisely, we introduce a bridge principle connecting knowledge and probability, which says that the knowledge an agent gains from a piece of evidence should respect what the appearances make probable. We show that this principle is violated by Centrism, because Centrism demands that an agent's knowledge of a real value always be convex.

Our principle is that whenever one real value is possible at a world and another is impossible, the appearances at that world must make the former real

[20] An anonymous referee observes that one surprising feature of this particular example is that when there is a pair of appearances of 3, the agent considers 1 likelier than 3. One might worry that this is incompatible with the agent's evidence playing the functional role of appearances. Notice, however, that this example is derived from a prior distribution on which any single appearance makes it most likely that the real value matches reality. Divergence between reality and appearance can only be probabilified by multiple pieces of evidence, not by any one alone. Finally, DIVERGENCE does not force a scenario where certain pairs of matching appearances probabilify values other than themselves. For example, imagine a version of our original hot/warm/cold case, in which the probability of the water being hot conditional on appearing hot is roughly .8; the probability of the water being warm conditional on appearing hot is roughly .2, and the probability of the water being cold conditional on appearing cold is negligible. Suppose though that the probability of the water being warm conditional on appearing warm is roughly .6, while the probability of the water being hot conditional on appearing warm is roughly .2. In this case, receiving experiences of cold and hot will provide stronger evidence of warmth than receiving experiences of warm and warm; but each of warm/warm, hot/hot, and cold/cold will favor warm, hot, and cold respectively.

value as likely as the latter.[21] Let $P(r \mid A)$ abbreviate $P(\{(r', A) : r' = r\} \mid \{(r', A') : A' = A\})$.

(16) **Respect.** If $r' \notin \mathsf{Real}^t(r, A)$ and $r'' \in \mathsf{Real}^t(r, A)$, then $P(r'|A^t) < P(r''|A^t)$.

To see why Centrism violates Respect, return to our final example from the introduction, where knowledge is not convex. Imagine that the agent dips her hands in a bowl of water. The water's temperature may be very cold, warm, or very hot. When the water is very cold, it is extremely likely to appear very cold, slightly likely to appear warm, and slightly less likely to appear very hot. When the water is warm, it is extremely likely to appear warm, slightly likely to appear very cold, and slightly likely to appear very hot. When the water is very hot, it is extremely likely to appear very hot, slightly likely to appear warm, and slightly less likely to appear very cold. Now imagine that the agent dips her hands in a bowl of very cold water on two different occasions, and the water appears first very cold and then very hot. In this case, the agent can know that the water is either very cold or very hot, but that it is not warm. This is because her evidence makes being very cold or very hot extremely likely, but makes being warm very unlikely.

We now supply an abstract version of this case. Again imagine we are measuring only 5 values of a quantity. Imagine the real value is 3, but that it appears to be 1 and 5. With a margin of $m = 2$, Centrism predicts that the agent knows the real value is between 2 and 4.

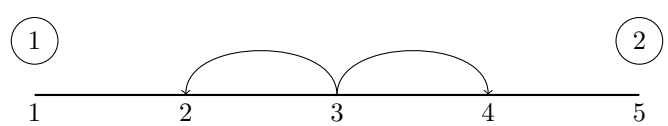

Figure 7: $r = 3, A = \langle 1, 5\rangle$, $m = 2$

Centrism implies the agent's knowledge of the real value is convex. We'll now see that this prediction violates Respect given a natural prior. So imagine that each real value has a likelihood of producing various appearances, weighted heavily towards itself. Crucially, after a steep initial drop-off, each distribution flattens out. When the real value is 1, there is a high likelihood of the value appearing 1 or 2, and a low likelihood of appearing 3, 4, or 5. When the real value is 5, there is a high likelihood of appearing 5 or 4, and a low likelihood

[21] On the other hand, we invalidate the extension of this principle from worlds to propositions. This principle says that if p is epistemically possible and q is epistemically impossible, then the probability of p on the appearances is higher than the probability of q on the appearances. To see the problem, imagine a model with 6 worlds, w_1 through w_6. w_1 has a probability of .4, w_2 has a probability of .2, and the remaining worlds have a probability of .1. Later we develop a model where if w_1 is actual, the agent knows she is in w_1 or w_2. But the agent knows she is not in w_3 through w_6. So the proposition that she is in w_2 is epistemically possible, and the proposition that she is in w_3 through w_6 is epistemically impossible, and yet the former proposition has a lower probability than the latter.

of appearing 1, 2, or 3. When the real value is 3, there is a high likelihood of appearing 2, 3, or 4, and the same small likelihood of appearing 1 or 5.

a	r	P
1	1	.110
2	1	.080
3	1	.003
4	1	.002
5	1	.001
1	2	.058
2	2	.080
3	2	.058
4	2	.003
5	2	.001
1	3	.002
2	3	.058
3	3	.080
4	3	.058
5	3	.002

a	r	P
1	4	.001
2	4	.003
3	4	.058
4	4	.080
5	4	.058
1	5	.001
2	5	.002
3	5	.003
4	5	.080
5	5	.110

Now our indifference and independence assumptions produce the following predictions about how likely various real values are given that the appearances are 1 and 5:

a_1	a_2	r	$P(\cdot \mid A = \langle 1, 5\rangle)$
1	5	1	.324
1	5	2	.171
1	5	3	.012
1	5	4	.171
1	5	5	.324

We can now see the failure of Respect. Centrism predicts that $\mathsf{Real}^2(3, 1, 5) = [2, 4]$. But we can see that $P(1 \mid A = \langle 1, 5\rangle) = P(5 \mid A = \langle 1, 5\rangle)$ is significantly larger than $P(3 \mid A = \langle 1, 5\rangle)$. So the theory predicts that the quantity might be 3 and can't be 1 or 5, even though 1 and 5 are more probable than 3 given that the agent had these appearances.[22]

In the next section, we validate Respect by defining knowledge directly in terms of probability and evidence.

5 Probabilism

We develop our model in several stages, considering progressively more complex models with progressively better predictions.

[22]To clarify, there are sometimes cases where the appearances are on opposite sides of the real value, and the agent can come to know that the real value is centered on the midpoint of these appearances. Everything depends on the underlying probabilities. We can't in general say which worlds are possible merely on the basis of facts about the distance between reality and appearance. We need more than just these distances; we need to know how likely each real value was of producing various appearances.

5.1 Defining probabilism

We begin by letting epistemic accessibility be the smallest reflexive relation that validates Respect. On this proposal, an agent knows as much as is possible to know consistent with the factivity of knowledge, and consistent with epistemic accessibility respecting the probabilities. Since this relation is reflexive, $E(r, A)$ is guaranteed to access (r, A). Since the relation satisfies Respect, it also accesses any world at least as probable as (r, A). The smallest relation with this property will relate a real value r only to itself and to any other real value as probable as r.

(17) $E^t(r, A) = \{(r', A) : P(r' \mid A^t) \geq P(r \mid A^t)\}$

This model makes the strange prediction that an agent can know p, even though p is unlikely on the apperances. Consider a simple case where an agent observes a quantity with five possible values. Imagine the quantity appears to have a value of 3. Let $P(r = n \mid A = \langle m, \ldots, m' \rangle)$ abbreviate $P(\{(r', A') : r' = n\} \mid \{(r', A') : A' = \langle m, \ldots, m' \rangle\})$. Suppose the probabilities are as follows:

a	r	$P(\cdot \mid A = \langle 3 \rangle)$
3	1	.15
3	2	.15
3	3	.4
3	4	.15
3	5	.15

(17) implies that the agent knows the real value is 3. But this is strange, since the real value is less likely to be 3 than to be some other value conditional on the appearances.

We require that whenever an agent knows p, the probability of p on the evidence is at least g, for some threshold g. In order to know p, one must be appropriately guided by one's evidence.

(18) **Guidance.** If S knows p at (r, A) at t, then $P(p|A^t) \geq g$.

To validate Guidance, we could let epistemic accessibility be the smallest reflexive relation that validates Respect and Guidance.[23]

This relation has a particular form. Since the relation is reflexive, $E(r, A)$ can access (r, A). Since it validates Respect, it can also access any world

[23] Respect does not imply Guidance, since (17) validates Respect but not Guidance. Similarly, Guidance does not imply Respect. Consider a model with five worlds, w_1 through w_5, where worlds with higher indices are more normal than worlds with lower indices. Consider a simple model where the epistemic possibilities at a world are those worlds that are not significantly less normal, and this set is found by admitting more and more abnormal worlds until the resulting set of worlds has a sufficiently high probability. Imagine the probabilities of w_1 through w_5 are .3, .25, .2, .15, and .1 respectively. If the actual world is w_5, the agent's epistemic possibilities are w_2 through w_5. In that case, every known proposition has a high probability. But w_1 is epistemically impossible while w_5 is epistemically possible, even though w_1 has a higher probability than w_5.

more probable than (r, A). But in order to validate Guidance, this relation must sometimes access worlds less probable than (r, A). To continue validating Respect, however, the relation is constrained so that whenever it accesses a world w less probable than the actual world, it also accesses any world v as probable as w.

In this way, validating Respect and Guidance together involves considering a system of spheres of possible worlds. Order all worlds by their probability conditional on the appearances. Now generate a system of spheres centered on whichever real values are most likely conditional on the appearances. The innermost sphere contains the real values that are most likely on the appearances. The next sphere contains the real values that are slightly less likely, and all real values more likely than them.

For example, consider again a simple case where an agent observes a quantity with five possible values, with an apparent value of 1. Suppose we have the following probabilities:

a	r	$P(\cdot \mid A = \langle 1 \rangle)$
1	1	.3
1	2	.25
1	3	.2
1	4	.15
1	5	.1

This induces a system of increasingly probable spheres centered on the real value 1, as in Figure 8.

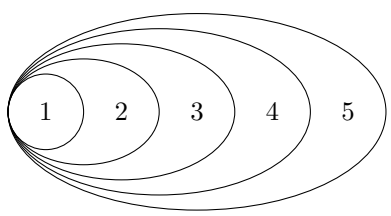

Figure 8: $A = \langle 1 \rangle$

Each sphere in Figure 8 has a different probability. From innermost to outermost, the spheres have the probabilities .3, .55, .75, .9, and 1. At any world, an agent's epistemic possibilities are the smallest sphere containing that world whose probability is at least g. So if $g = .7$, then $E(1, \langle 1 \rangle)$ includes the real values from 1 through 3. After all, five spheres contain the real value 1. But the sphere containing just real values 1 through 3 is the smallest one whose probability is greater than .7. By contrast, $E(4, \langle 1 \rangle)$ includes the real values from 1 through 4. Smaller spheres than this have a probability greater than g. But these smaller spheres don't include the real value 4.

More precisely, building on Hawthorne 2021, we can define epistemic accessibility in several steps. First, we use the probabilities and appearances to define a

system of spheres. Then we filter down this system to the guided spheres, which are the spheres in the system whose probability is at least g. Then we find the smallest guided sphere that contains the base world. The epistemic accessibility relation says that at (r, A) the epistemic possibilities are the smallest g-probable sphere that contains (r, A). In this way, an agent knows as much as is possible consistent with Respect, Guidance, and the reflexivity of epistemic accessibility.

(19) a. $\mathsf{Spheres}(r, A) = \{\{(r'', A) : P(r'' \mid A^t) \geq P(r' \mid A)\} : P(r' \mid A) \leq P(r \mid A)\}$

b. $\mathsf{GuidedSpheres}(r, A) = \{p \in \mathsf{Spheres}(r, A) : P(p \mid A^t) \geq g\}$

c. $\mathsf{Strongest}(r, A) = \iota p : p \in \mathsf{GuidedSpheres}(r, A) \;\&\; \forall q \in \mathsf{GuidedSpheres}(r, A) : p \subseteq q$

d. $E^t(r, A) = \mathsf{Strongest}(r, A^t)$

(19-d) is an elegant theory of knowledge that validates Respect and Guidance.[24] But it has other controversial consequences, such as Positive Introspection.[25]

(20) **Positive Introspection.** If S knows p, then S knows that S knows p.

The epistemic accessibility relation induced by Guided Probabilism is transitive, and so validates Positive Introspection. Return to Figure 8. When the real value is 1, the agent knows the real value is 1 through 3. But when the real value is 2 or 3, the agent knows the same thing. Generalizing, the key observation is that for any world w, every member v of the strongest g-probable w-containing sphere S itself has S as the strongest g-probable v-containing sphere.

The model in (19-d) also permits stronger instances of introspection. Drawing on Goodman 2013, Williamson 2013b considers and rejects the possibility of cliff-edge knowledge at a world (r, A), where either S knows that the real value is at least r or S knows that the real value is at most r.[26]

(21) **Cliff-edge knowledge.** S has cliff-edge knowledge at (r, A) at t iff $\forall n > 0 : r + n \notin \mathsf{Real}^t(r, A)$ or $\forall n > 0 : r - n \notin \mathsf{Real}^t(r, A)$.

[24] Goodman and Salow 2018 develop a quite similar model that does not validate Guidance. They let the epistemically accessible worlds at w be the union of (i) the worlds not significantly less normal than the most normal worlds consistent with the evidence and (ii) the worlds at least as normal as w. In some applications, they understand normality in terms of probability. In Figure 8, for example, they might say that the accessible worlds at w_i when w_1 is consistent with the evidence are the union of $\{1, 2, 3\}$ with any w_j where $j \leq i$. This theory differs from ours in the case where an agent's evidence eliminates worlds besides the most normal world. For example, if the appearances rule out w_2, our theory can allow w_4 to be accessible at w_3. Once w_2 is eliminated, even w_1 may access w_4. By contrast, in Goodman and Salow 2018 the elimination of the second most normal world does not affect which worlds are significantly less normal than the most normal world. Guidance can then fail at the most normal world in a case where the evidence eliminates every world less normal: only the most normal world will be accessible, which may not itself be very probable.

[25] Of course, Positive Introspection only holds here in the presence of our simplifying assumptions, including Appearance Luminosity. It also only holds if the ordering over worlds (in our setting, the probabilities) is not sensitive to which world one is in.

[26] Stalnaker 2009, p. 406 defends cliff-edge knowledge. For a response, see Hawthorne and Magidor 2010, p. 1092. See also Weatherson 2013, p. 67.

In Figure 8, with $g = .7$ the agent has cliff-edge knowledge when the real value is 3 or 4. Generalizing, the agent has cliff-edge knowledge of a real value whenever the set of real values at least as probable as the real value r is itself g-probable.

To avoid validating Positive Introspection, we constrain epistemic accessibility with a margin for error requirement. This margin consists of the real values whose probability isn't significantly less than the actual real value. Here, we face a choice point regarding how to define significant drops in probability. One option is to consider the distance $|P(r' \mid A) - P(r \mid A)|$. This predicts that .1 is much further from .2 than .01 is from .02. We prefer to require that the ratio of $P(r' \mid A)$ to $P(r \mid A)$ be greater than some threshold s. When the real value is r, the agent treats any other real value as possible whose probability is not less than s times the probability of r.

(22) **Probabilistic margin for error.** $\forall(r, A) : \{r' \mid \frac{P(r'|A^t)}{P(r|A^t)} \geq s\} \subseteq \mathsf{Real}^t(r, A)$

According to our preferred model, 'Probabilism', epistemic accessibility is the smallest relation satisfying Respect, Guidance, and Probabilistic Margin for Error. An agent knows as much as is possible to know consistent with respecting the appearances, and also respecting the margin for error. On this theory, the epistemic possibilities at real value r are any worlds in one of two categories. The first category is the set of worlds in the strongest g-probable r-containing sphere. The second category is the set of worlds whose probability isn't less than s of the probability of r.

(23) **Probabilism.** $E^t(r, A) = \mathsf{Strongest}(r, A^t) \cup \{(r', A^t) \mid \frac{P(r'|A^t)}{P(r|A^t)} \geq s\}$

Given the right prior, this theory violates Positive Introspection. Consider again the simple case where an agent observes a quantity with five possible values, with an apparent value of 1, and with the following probabilities from Figure 8:

a	r	$P(\cdot \mid A = \langle 1 \rangle)$
1	1	.3
1	2	.25
1	3	.2
1	4	.15
1	5	.1

Now suppose $g = .7$ and $s = \frac{2}{3}$. This means that an agent knows p only if the probability of p on the appearances is at least .7, and the agent treats r' as possible if the probability of r' isn't less than $\frac{2}{3}$ the probability of r.
Positive Introspection fails. When the real value is 2, the agent knows the real value is between 1 and 3. But when the real value is 3, the agent only knows that the real value is between 1 and 4. So when the real value is 2, the agent knows the real value is between 1 and 3, but doesn't know this fact.

Probabilism also rules out some of the above cases of cliff-edge knowledge. In Figure 8, the agent no longer knows the real value is between 1 and 3 when

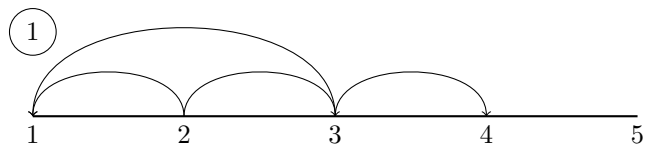

Figure 9: $A = \langle 3 \rangle$

the real value is 3. When the real value is 3, it wouldn't be significantly less likely for the real value to be 4, and so it is epistemically possible that the real value is 4.[27]

Probabilism is also compatible with the normality theory of knowledge. On this proposal, S knows p when p is true at any world that is not significantly less normal than actuality.

(24) $E(r, A) = \{(r', A) : (r', A) \not\ll (r, A)\}$

Say a world's degree of normality is its probability conditional on the appearances. Where $P(r \mid A)$ abbreviates $P(\{(r', A') : r' = r\} \mid \{(r', A') : A' = A\})$:

(25) $(r, A) \leq (r', A)$ iff $P(r \mid A) \leq P(r' \mid A)$

At any world, an agent's epistemic possibilities are whichever worlds are not significantly less probable than the actual world. Our model agrees with this theory, given the right definition of 'significantly less probable'. In particular, we can say that (r', A) is significantly less normal than (r, A) just in case the ratio of $P(r' \mid A)$ to $P(r \mid A)$ is less than s, and (r', A) is not in the smallest g-probable sphere containing (r, A). Probabilism is then a special case of the normality theory.[28]

[27] On the other hand, our model does permit cliff-edge knowledge in cases where the next real value would be significantly less likely than the actual one. We think this is the right prediction: in order for the probabilities to be this way, the agent would have to have experienced a large number of reliable appearances. In that case, they could plausibly discover the real value, at least in finite cases.

[28] We've been working with models where there are a finite number of possible values for the relevant quantity. It's natural to wonder how our theory might look if we were to generalize it to the infinite case. In particular, what of a setting where precise values are represented and the possible values form a continuum? (As before, we do not consider cases where the apparent value is a range.) Obviously something needs tweaking here, for the simple reason that it's natural to think that every precise value has probability 0, and so there won't be any interesting differentiation of worlds into spheres, given that they are all equiprobable. But there's a natural way to extend our theory to the infinite case, at least when the probability distribution is continuous. Each value will have the same probability, namely 0, but it is false that each value has the same probability density. By way of picture thinking, imagine one's probability distribution to be represented by a curve, where the probability for each value range is proportional to the area under the curve. Then the probability density at a point will correspond to the height of the curve at the point, which corresponds to the limit of the values found by taking increasingly smaller segments of the curve centered on that point, and dividing that segment's area by its width. Once we have probability densities for each point, we can proceed as before, building a sphere system on probability densities rather than probabilities.

With Probabilism stated precisely, we turn to its applications for multiple appearance.

5.2 Applying probabilism

First, Probabilism validates Respect. If one real value is possible and another isn't, then the former is supported by the appearances as much as the latter. To see Respect in action, consider our counterexample to convexity.

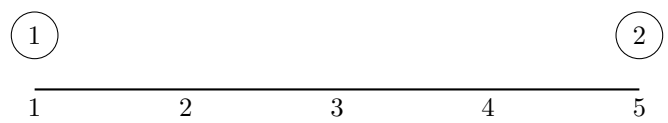

Figure 10: $A = \langle 1, 5 \rangle$

a_1	a_2	r	$P(\cdot \mid A = \langle 1,5 \rangle)$
1	5	1	.324
1	5	2	.171
1	5	3	.012
1	5	4	.171
1	5	5	.324

According to Probabilism, what an agent knows depends on the real value. When $r = 3$, the real value is extremely unlikely given the appearances. For this reason, they know nothing about the real value. But when the real value is anything else, the agent knows something interesting. Suppose for simplicity that the threshold g for guidance is $\frac{2}{3}$. Then when $r = 1$ or $r = 5$, the accessible real values are $\{1, 2, 4, 5\}$. When $r = 1$ or $r = 5$, the probability of the actual world on the evidence is .324, and the probability of the set of worlds at least as probable as this ($\{1, 5\}$) is just below $\frac{2}{3}$. The probability of the set of worlds almost as probable as this ($\{1, 2, 4, 5\}$) is almost 1, and so this is the set of epistemic possibilities. Since the proposition that the real value is 1, 2, 4, or 5 is supported sufficiently by the appearances, it is known. Epistemic accessibility is not convex, since value 3 is excluded. Since Probabilism allows failures of convexity, Probabilism rejects Appearance Centering. The possible real values are not just those centered on any point or range, let alone the appearances.

Our model allows for failures of convexity. But nothing in our model requires that convexity fails in all or most cases of perceptual experience. In some cases, convexity failures are quite natural. For example, imagine that you are monitoring the temperature in a room with two different thermometers. When you check the readings, one thermometer reports 20 degrees Fahrenheit, and the other predicts 80. In this case, a convexity failure is natural. The most plausible explanation here is that one thermeter is broken. Similarly, imagine

Then our theory can proceed as before.

that you are interviewing two eyewitnesses to a crime, and suspect one of them may be an accomplice. If the two eyewitnesses report wildly divergent heights for the perpetrator, you may suspect that one of them is lying, in which case convexity may fail. By contrast, imagine you are at a state fair, and a hundred people have placed a guess about how many jelly beans are in a jar. Here, convexity may be natural, even if the guesses tend to cluster around certain answers (for example, round numbers). Our model can smoothly accommodate the range of cases. Whether convexity holds in a particular case depends on the shape of the underlying probabilities. Convexity will hold in cases where the slope of probabilities is suitably gentle rather than steep, so that a real value has a significant chance of producing apparent values which diverge from it significantly. Indeed, it is open to our models that many ordinary cases of perceptual knowledge have just this structure, and support convexity. This would explain the prevalence of convexity in existing discussions of models of appearance and reality.

Probabilism also rejects Match. Appearances which match reality can provide less knowledge than appearances which diverge from reality.

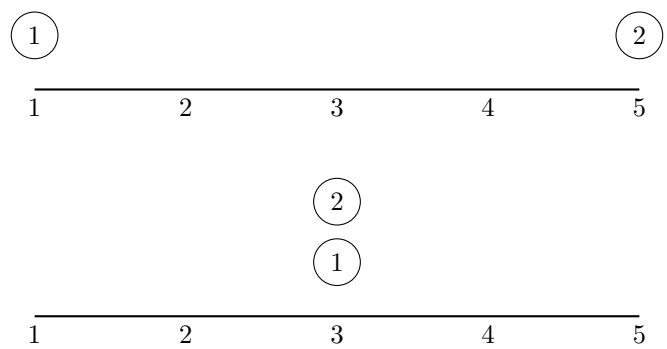

Figure 11: $A = \langle 1, 5 \rangle$

a_1	a_2	r	$P(\cdot \mid A = \langle 1,5 \rangle)$
1	5	1	.067
1	5	2	.111
1	5	3	.643
1	5	4	.111
1	5	5	.067

a_1	a_2	r	$P(\cdot \mid A = \langle 3,3 \rangle)$
3	3	1	.233
3	3	2	.193
3	3	3	.149
3	3	4	.193
3	3	5	.233

Imagine the real value is 3. In the first case, the agent receives the evidence $\langle 1, 5 \rangle$. In the second case, the agent receives the evidence $\langle 3, 3 \rangle$. In the second case the agent's evidence perfectly matches reality; in the first, it doesn't. Nonetheless, Probabilism predicts that the agent learns a great deal in the first case, and nothing at all in the second case. In the first case, the probability of the actual world given the evidence is .643, which is far more probable than any other world. In the second case, the probability of the actual world given the evidence is .149, which is lower than all other real values. So the agent doesn't know

anything about the real value. Probabilism again explains in a systematic way how an agent's knowledge depends on what is probable given her evidence.

Probabilism also allows knowledge defeat. For reasons of space, we won't provide a worked out example. But the basic diagnosis is by now familiar. Suppose the agent dips her hands twice in warm water, and the water first appears warm and then hot. After the first appearance, the agent's evidence may make it quite likely that the water is warm, and much less likely that it is cold or hot. So at that time the agent may know that the water is not hot. But after she checks again and the water appears hot, her evidence produces a smoother probability distribution that assigns a fairly high probability to both the hypothesis that the water is warm and the hypothesis that the water is hot. In this later state, she agent loses her earlier knowledge that the water is hot. Knowledge defeat is produced when new information affects the underlying probability measure over worlds.

Finally, Probabilism predicts that multiple matches can generate more knowledge than a single match. When the agent checks the water twice and it appears warm both times, her resulting probability measure can become concentrated on the hypothesis that the water is warm. Multiple appearances with the same content can have further epistemic effect by further concentrating the posterior probability distribution.[29]

6 Probability and Normality

Our discussion has made heavy use of the concept of probability, but has been deliberately non-committal about the kind of probability that is called for. And this is because we want to allow for different developments of the theory that differ according to the kind of probability that is appealed to. A theory that is more internalist in flavor might appeal to the kind of evidential or *a priori* probabilities that are at play in Williamson 2000 and objective Bayesians. But one might instead use a kind of probability distribution that is more tied to the physical niche in which the organism finds itself, and which might be appealed to by a psychophysics of the perceptual systems of that organism. These kind of probabilities would be the basis of the kind of objective estimations of reliability that a more externalist reliabilist might have in mind. Which kinds of probabilities might provide the most satisfying model for perceptual knowledge

[29] Before concluding, it is worth comparing our model with that in §2, where an agent receives a single appearance. Our model generates that model as a special case, provided that we impose a few constraints on the prior. First, suppose that $\frac{P(r-m|A)}{P(r|A)} = s$. Second, suppose that $P(\{(r', A) : r' \in [r-m, r+m]\} \mid A) = g$. Third, suppose that the probability of a real value conditional on an appearance decreases monotonically with distance from the apparent value. In that case, our model produces the same predictions as that in §1. When the real value is 50 and the apparent value is 50, the agent knows that the real value is within m of 50. When the real value is 45 and the apparent value is 50, the agent knows that the real value is anywhere from $45 - m$ to $55 + m$. In our view, then, the plausibility of the distance-based theory is explained by the way that the distance-based theory encodes natural assumptions about the priors.

is a question that we wish to leave open. It is worth remarking that there are plenty of models of perception that appeal to probability distributions but don't invoke the concept of knowledge at all. The standard Bayesian models update probability distributions by perceptual appearances in a way such that no external situation gets probability 0, and on a flat-footed theory according to which knowledge would require probability 0, every external situation would be epistemically live after a series of appearances. Similarly, empirical models of the Bayesian brain often appeal to various kinds of Gaussian distributions of external value hypotheses but are silent on the facts of knowledge.[30] Here again, since Gaussian curves trail off without hitting 0, a flat-footed way of tying the facts of knowledge to the facts of probability would yield highly skeptical results. It would be nice to have a non-skeptical model which didn't simply discard the kinds of probabilistic models that we've just alluded to, but attempted some non-skeptical marriage of the concept of knowledge with the kinds of probability distributions that these theories appeal to. That's what we've tried to accomplish in this paper.

Of course one thing that won't be helpful for this model are what we might call 'Williamsonian probabilities'. Williamsonian probabilities are what you get by conditionalizing some prior probability distribution on the propositions that are known. When p is known, the Williamsonian probability of p is 1. It's at best unhelpful to explain knowledge in terms of probabilities that are themselves directly explained in terms of knowledge.

Nor do we want to get into the weeds about the concept of evidence. Just as we've got no objection to the existence of Williamsonian probabilities, we don't particularly wish to argue here against the thesis that one's evidence or more generally what one has to go on is most fundamentally about what one knows. We don't wish to appropriate the concept of evidence. What we wish to say is that when it comes to giving an informative model of how the epistemology of multiple appearances works, one wants a notion of appearance as distinct from things known, and wants a probability distribution that is non-Williamsonian.

We've been focusing on the case of perceptual knowledge from multiple appearances. But it's natural to wonder whether the contours of our model might be helpful for modeling other kinds of epistemic phenomena. Here we are cautiously optimistic. The case of multiple reports about a single quantity from various informants seems plausibly analogous to the case of multiple appearances of a single quantity. Here again we suspect that many of the lessons will carry over nicely. Without a doubt, the tendency of epistemologists of testimony would be to think that the best case for maximal information is matching testimony. But we see no reason not to extend the conclusions of our previous discussion to provide a suitable antidote to these traditional inclinations. Similarly, though perhaps more mundanely, our model will naturally extend to understanding the epistemic import of multiple voltmeter or thermometer readings of a single quantity.

[30] For a helpful introduction to this literature, see Doya et al. 2011, Trommershauser et al. 2011.

On the other hand, there are cases where notions of relative probability of individual outcomes are not intuitively allied to judgments of relative normality. For example, consider a hundred ticket 'asymmetric' lottery where the first ticket is 10% likely to win, tickets 2 through 90 each have a roughly .91% chance of winning, and tickets 91 through 100 each have a roughly .9% chance of winning. Suppose that the agent's ticket is number 97, and ticket 1 is the winner. Intuitively the agent does not know that her ticket will lose. But Probabilism appears to make this prediction. The actual world's probability on the evidence is 10%, since ticket 1 wins. The set of worlds where the ticket is 1 through 90 has a very high probability, over 90%, and the actual world is significantly more likely than any world outside this set. So Probabilism predicts that the epistemic possibilities do not include worlds where the winning ticket is 91 through 100. In such cases, it is much more tendentious to try to use the facts of relative probability to model the evolution of knowledge.

To model such cases, one promising option would be to introduce a normality ordering over worlds that is defined independently of probability.[31] One could then define knowledge in terms of the both normality and probability. In particular, we could generalize our two characterizing principles of Guidance and Respect to this richer setting. Guidance continues to require that the epistemic possibilities have a sufficiently high probability on the evidence. But we could now introduce another version of Respect, which says that when one world is epistemically possible and another is not, the first world is at least as normal as the other. Together, these two principles suggest a construction of epistemic possibilities in terms of normality and probability. The epistemic possibilities are those worlds not significantly less normal than actual. But this idea is itself cashed out in terms of probability: starting with the actual world, we admit increasingly abnormal worlds into the epistemic possibilities until the resulting probability of the epistemic possibilities exceeds the threshold imposed by Guidance (and Probabilistic Margin for Error is satisfied). This model can potentially explain asymmetric lottery cases, where normality and probability diverge. In addition, this model has the potential to extend our account to models with infinite numbers of worlds, where there may not be a straightforward assignment of probability to individual worlds, but there may nonetheless be natural things to say about when one world is more normal than another.[32]

In summary, our probabilistic models can do much the same as the models

[31] See Goldstein and Hawthorne 2021 for defense of the role of normality in the theory of knowledge.

[32] Both our original model and the revised normality model allow that agents can know that some possibilities do not obtain before having any experiences, assuming that the relevant probability distribution is sufficiently biased against these possibilities prior to evidence (and, in the case of the normality model, assuming that these possibilities are sufficiently abnormal). It is not hard to motivate such prior knowledge. Consider the hypothesis that our life occurs in a hundred year bubble of relative normality embedded in a sea of chaos. If this hypothesis is epistemically possible in advance of experience, it is hard to see how later evidence could rule it out. See Bacon 2020 for further discussion of whether it is appropriate to posit this kind of prior knowledge. It is beyond the scope of the present paper to explore this interesting issue further.

which we intend to supercede in the case of single appearances. But our probabilistic models are far better placed than any of those models to provide substantial insight into how knowledge evolves in the case of multiple appearance.[33]

[33]Thanks to Irene Bosco, Sam Carter, Stephanie Collins, Cian Dorr, and Jeremy Goodman.

References

Andrew Bacon. Inductive knowledge. *Nous*, 54(2):354–388, 2020.

Bob Beddor and Carlotta Pavese. Modal virtue epistemology. *Philosophy and Phenomenological Research*, 2019.

Sam Carter. Higher order ignorance inside the margins. *Philosophical Studies*, pages 1–18, 2019.

Sam Carter and Simon Goldstein. The normality of error. *Philosophical Studies*, 2021.

Stewart Cohen and Juan Comesaña. Williamson on gettier cases and epistemic logic. *Inquiry*, 56:15–29, 2013.

Cian Dorr, Jeremy Goodman, and John Hawthorne. Knowing against the odds. *Philosophical Studies*, 170(2):277–87, 2014.

Kenji Doya, Shin Ishii, Alexandre Pouget, and Rajesh P.N. Rao, editors. *Bayesian Brain: Probabilistic Approaches to Neural Coding*. MIT Press, 2011.

Julien Dutant and Sven Rosenkranz. Inexact knowledge 2.0. June 2019.

Daniel Garber. Field and jeffrey conditionalization. *Philosophy of Science*, 47 (1):142–145, 1980.

Simon Goldstein. Fragile knowledge. *Mind*, 2021.

Simon Goldstein and John Hawthorne. Counterfactual contamination. *Australasian Journal of Philosophy*, 2021.

Jeremy Goodman. Inexact knowledge without improbable knowing. *Inquiry*, 56 (1):30–53, 2013.

Jeremy Goodman and Bernhard Salow. Taking a chance on kk. *Philosophical Studies*, 175(1):183–96, 2018.

Jeremy Goodman and Bernhard Salow. Knowledge from probability. In *Proceedings of TARK XVIII: Eighteenth Conference on Theoretical Aspects of Rationality and Knowledge*, Beijing, 2021.

John Hawthorne. The epistemic use of 'ought'. In Lee Walters and John Hawthorne, editors, *Conditionals, Paradoxes, and Probability: Themes from the Philosophy of Dorothy Edgington*. Oxford University Press, 2021.

John Hawthorne and Ofra Magidor. Assertion and epistemic opacity. *Mind*, 119 (476):1087–1105, 2010.

Frank Hong. *Uncertain Knowledge*. PhD thesis, University of Southern California, In Preparation.

Robert Stalnaker. On hawthorne and magidor on assertion, context, and epistemic accessibility. *Mind*, 118(470):399–409, 2009.

Julia Trommershauser, Konrad Kording, and Michael S. Landy. *Sensory Cue Integration*. Oxford University Press, 2011.

Brian Weatherson. Margins and errors. *Inquiry*, 56(1):63–76, 2013.

Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.

Timothy Williamson. Gettier cases in epistemic logic. *Inquiry*, 56(1):1–14, 2013a.

Timothy Williamson. Response to cohen, comesaña, goodman, nagel, and weatherson on gettier cases in epistemic logic. *Inquiry*, 56(1):77–96, 2013b.

Timothy Williamson. Very improbable knowing. *Erkenntnis*, 79:971–999, 2014.

Timothy Williamson. Model-building in philosophy. In Russell Blackford and Damien Broderick, editors, *Philosophy's Future*. Oxford University Press, 2017.