Inquiry, 2013
Vol. 56, No. 1, 30–53, http://dx.doi.org/10.1080/0020174X.2013.775013

Inexact Knowledge without Improbable Knowing

JEREMY GOODMAN

University of Oxford, UK

(Received 6 February 2013)

ABSTRACT *In a series of recent papers, Timothy Williamson has argued for the surprising conclusion that there are cases in which you know a proposition in spite of its being overwhelmingly improbable given what you know that you know it. His argument relies on certain formal models of our imprecise knowledge of the values of perceptible and measurable magnitudes. This paper suggests an alternative class of models that do not predict this sort of improbable knowing. I show that such models are motivated by independently plausible principles in the epistemology of perception, the epistemology of estimation, and concerning the connection between knowledge and justified belief.*

Gettier cases are a welcome prediction of Williamson's models of perceptual knowledge.[1] But his models also make some more surprising predictions that many philosophers will find unwelcome. For example, once enriched in the natural way to model epistemic probabilities, they predict *improbable knowing*: cases in which you know a proposition in spite of its being improbable given what you know that you know it. Even if we should ultimately accept this prediction (as Williamson has argued elsewhere that we should),[2] it nevertheless ought to give us pause.

Another surprising prediction of Williamson's models is *pervasive illusion*: in almost all cases your surroundings are not the way that they perceptually appear. This prediction is clearly unwelcome, since it is inconsistent with the attractive view that, normally, the way that your surroundings perceptually appear to you is a way that you perceive—and thereby know—them to be.

Correspondence Address: Jeremy Goodman, University of Oxford, Oxford, UK. Email: goodman.jeremy@gmail.com

[1]Williamson, 'Gettier Cases in Epistemic Logic'.

[2]Williamson, 'Improbable Knowing', 'Very Improbable Knowing'.

Elsewhere Williamson has expressed sympathy for something like this view, conjecturing that 'in normal circumstances . . . our perceptual beliefs (those in which we take perceptual appearances at face value) count as knowledge'.[3] This conjecture is reinforced by Williamson's general picture of knowledge as central to our cognitive lives. According to this picture, perceiving is a kind of knowing, belief aims at knowledge, perceptual experience stands to perceiving as belief stands to knowing, and knowledge is commonplace. In *Knowledge and Its Limits* he writes:

> While belief aims at knowledge, various mental processes aim at more specific factive mental states. Perception aims at perceiving that something is so; memory aims at remembering that something is so. Since knowing is the most general factive state, all such processes aim at kinds of knowledge. If a creature could not engage in such processes without some capacity for success, we may conjecture that nothing could have a mind without having a capacity for knowledge.[4]

The last sentence of this passage even hints at a transcendental argument to the effect that the capacity to know by perception that things are as they perceptually appear is a necessary condition for having any perceptual experiences whatsoever. Whether or not Williamson would in fact endorse such an argument, he certainly seems to be committed to our normally being able to know that things are as they appear.[5]

But can this commitment be reconciled with the inexactness of our perceptual knowledge? It can, if we abandon Williamson's simplifying assumption that perceptual appearances are maximally specific. In Section I, I demonstrate the consistency and plausibility of this reconciliation by presenting and defending a natural refinement of Williamson's models in which perceptual appearances are not maximally specific. In Section II, I show that these refined models also fail to predict improbable knowing. In Section III, I use these models to develop a general strategy for resisting Williamson's argument for improbable knowing, even in non-perceptual cases. In Section IV, I consider three other arguments for improbable knowing and argue that they are inconclusive. In Section V, I show that Williamson's models of justified belief conflict with the attractive principle that a belief is justified only if it is the manifestation of a disposition to know. Modifying Williamson's models of imprecise knowledge so that they respect this connection between knowledge and justified belief yields the very models that I defend in the previous sections on independent grounds.

---

[3]Williamson, 'Knowledge and Skepticism', 697.

[4]Williamson, *Knowledge and Its Limits*, 48.

[5]Throughout, read 'appear' as outside the scope of 'know' in 'know that things are as they appear'.

I. Unspecific Appearances

Let us review Williamson's model $<W, R>$. The set of 'worlds' W is the set of ordered pairs of real numbers; these pairs' first and second members respectively represent the real and apparent values of some perceived magnitude, such an object's height or a surface's brightness. We model propositions as subsets of W. Let $R(<e, f>)$ abbreviate $\{<e^*, f^*>: <e, f>R<e^*, f^*>\}$; $R(<e, f>)$ is the strongest proposition known at $<e, f>$. In the model, knowledge is closed under entailment: p is known at $<e, f>$ just in case $R(<e, f>) \subseteq p$. Let Kp abbreviate $\{<e, f>: R(<e, f>) \subseteq p\}$; Kp is the proposition that p is known. Finally, let $P(<e, f>)$ abbreviate $\{e^*$: for some f^*, $<e, f>R<e^*, f^*>\}$; $P(<e, f>)$ is the set of values that, for all you know, are magnitude's real value.

The 'epistemic accessibility relation' R is defined as follows, relative to an arbitrarily chosen positive constant c: '$<e, f>R<e^*, f^*>$ if and only if $|e^* - f^*| \leq |e - f| + c$ and [. . .] $f = f^*$.' Williamson glosses this definition as follows: 'Thus the worlds accessible from a given world are those where the apparent value is exactly the same and the gap between it and the real value exceeds the gap in the given world by at most the constant c.'[6] More perspicuously, we can define R to be the smallest reflexive binary relation on W satisfying the following conditions:

Margin for Error: $[e - c, e + c] \subseteq P(<e, f>)$.
Any value within a radius c of the magnitude's real value is one that, for all you know, is the magnitude's value.

Luminous Appearances: $<e, f>R<e^*, f^*>$ only if $f = f^*$.
You know how the magnitude appears.

Appearance Centering: For some x, $P(<e, f>) = [f - x, f + x]$.
For every world, there is number x such that the strongest thing you know at that world about the magnitude's value is that it is within a radius x of its apparent value.

This definition is equivalent to Williamson's.

How might we modify this model to account for unspecific appearances? The natural proposal is to interpret the second member of a world not as a maximally specific value that the magnitude appears to have but instead as the midpoint of the smallest interval in which the magnitude's value appears to lie. In particular, we assume that, for some positive constant d, a world $<e, f>$ represents a situation in which the magnitude has the value e and unspecifically appears to lie in the interval $[f - d, f + d]$. Unspecific appearances

[6] Williamson, 'Gettier Cases in Epistemic Logic', 6.

have the potential to make an epistemic difference if we impose the following natural constraint:

Appearance Constraint: $[f - d, f + d] \subseteq \mathrm{P}(<e, f>)$.
Any value of the magnitude compatible with its (unspecific) appearance is compatible with our perceptual knowledge.

We can now define, for positive constants c and d, a new accessibility relation R^* as the smallest reflexive binary relation on W satisfying Margin for Error, Luminous Appearances, Appearance Centering, and the Appearance Constraint, and explore the resulting epistemic model $<\mathrm{W}, \mathrm{R}^*>$.[7]

For simplicity, these models treat the 'radius' d of a subject's unspecific perceptual appearances as constant across worlds. We have reason to think that this assumption is at least approximately correct. For example, it is an extremely plausible psychophysical and phenomenological hypothesis that the similarity structure of normal color experiences preserves the three-dimensional similarity structure of colors, in terms of hue, brightness and saturation.[8] But if there were appreciable variation from case to case in the unspecificity of perceptual appearances of hue, brightness and saturation, then color experience would have a six-dimensional similarity structure, since intervals are characterized by two numbers rather than one. The point generalizes: if the unspecificity of our perceptual appearances of a perceptible magnitude varied appreciably from case to case, then we would expect the similarity structure of our perceptual appearances of the magnitude to have twice the magnitude's dimensionality. We will revisit this assumption in Section IV.

If $d \leq c$, then $\mathrm{R} = \mathrm{R}^*$ and unspecific appearances make no epistemic difference. For concreteness, consider the model obtained by letting $d = c$. This model avoids pervasive illusion, since our unspecific appearances only fail to be veridical when their midpoint diverges from the magnitude's real value by more than our margin for error c. If our motivation for positing unspecific appearances were simply to avoid pervasive illusion, we could therefore reasonably claim vindication.[9] But we want more. We want to vindicate not merely the normality of things *being* as they appear, but the normality of our *knowing* that things are as they appear. In models in where $d = c$, taking appearances at face value yields knowledge only in the overwhelmingly improbable event that the magnitude's real value is exactly equal to the midpoint of its unspecific appearance. (It is impossible to know that things are as they appear in models where $d < c$.) Therefore, if unspecific appearances are to vindicate the normality of knowing that things are as they appear, their unspecificity must exceed our margin for error.

[7]Equivalently, $<e, f>\mathrm{R}^*<e^*, f^*>$ if and only if $|f^* - e^*| \leq \mathrm{Max}\{|f - e| + c, d\}$ and $f = f^*$.

[8]See Morrison, 'Color in a Physical World' for discussion.

[9]Hellie ('Noise') and Stazicker ('Attention and Indeterminacy') posit unspecific appearances on precisely such grounds.

Benj Hellie draws a different moral. He thinks that adding the Appearance Constraint allows us to drop Margin for Error.[10] The resulting model would predict that, whenever things are as they appear, we are able to know that things are as appear. In this sense, knowing that things are as they appear would be not merely normal but *luminous* (i.e., known to occur whenever it does occur). But although Margin for Error and the Appearance Constraint both share the consequence that our perceptual knowledge is inexact, the latter is no substitute for the former. Margin for Error is motivated not by the inexactness of perceptual knowledge, but rather by the idea that perceptual knowledge, like all knowledge, requires belief that is safe from error. Because our perceptual systems are *noisy*, this 'safety' condition entails that there are epistemically accessible cases in which the magnitude takes a value slightly different from its actual one and yet we enjoy the same appearances that we actually do.[11] We parameterize this range of cases by c: if x is the degree to which the midpoint of appearances diverges from the perceived magnitude's true value, then there is an epistemically accessible case with the same appearances in which the magnitude of this divergence is $x + c$. In effect, c is a measure of the epistemic impact of noise in our perceptual systems.[12] This is simply a different phenomenon from the one captured by the Appearance Constraint, which measures the epistemic impact not of perceptual noise but of perceptual unspecificity. Without a margin for error, there would be cases in which a magnitude with value x is known, for some y, to have a value in the interval $[x, y]$. Such knowledge is implausible in light of noise in our perceptual systems, assuming a safety condition on knowledge. We should therefore retain Margin for Error. (In their contribution to this symposium, Stewart Cohen and Juan Comesaña propose a model consistent with Hellie's in which Margin for Error also fails, drawing on recent work by Robert Stalnaker.[13] Their proposal is therefore implausible for the same reason that Hellie's is.)

If unspecific appearances are to vindicate the normality of our knowing that things are as they appear, we must let $d = c + b$, for some positive constant b that represents an upper bound on the size of normal deviations between the magnitude's real value and the midpoint of its unspecific appearance. Having imposed this constraint, $R^* \neq R$: the unspecificity of appearances makes an epistemic difference when it exceeds our margin for error. We can characterize this difference as follows: if the magnitude's real value is within a radius b of the midpoint of its appearance, then we know

[10] Hellie, 'Noise', 495 ff.

[11] Making good on this claim is beyond the scope of this paper; see Williamson, 'Probability and Danger'.

[12] As Williamson writes: 'The value of [c] will be determined by features of the visual mechanism, such as the long-run probability of a given distance between the real and apparent [values] of the [perceptible magnitude]' ('Very Improbable Knowing').

[13] Cohen and Comesaña, 'Williamson on Gettier Cases'; Stalnaker, 'On Hawthorne and Magidor'.

that things are as they appear; otherwise, the model agrees with Williamson's. Formally: $R^*(<e, f>)$ is $\{<e^*, f>: |e^* - f| \leq d\}$ if $|f - e| \leq b$, and is $R(<e, f>)$ otherwise. Intuitively, our appearance of width $2d$ is built out of a 'safe haven' of width $2b$ with 'buffer zones' of width c on either side. As long as the magnitude's real value is in the safe haven (as it normally will be), taking our perceptual appearances at face value yields knowledge. In particular, if the magnitude's value is in the safe haven, then what we know is that it is either in the safe haven or in the surrounding buffer zones—perceptual appearances and perceptual knowledge will coincide. The resulting model reconciles the normal knowledge-conduciveness of taking our perceptual appearances at face value and the inexactness of our perceptual knowledge owing to perceptual noise.

While this reconciliation is surely welcome, one might have reservations about the model's psychological plausibility. Ned Block, when considering the hypothesis that perceptual appearances are 'abstract relative to other contents, as determinables are to determinates, for example as red is to scarlet', objects that 'the variation of 6% due to attention is way above the "just noticeable difference" threshold, which for stimuli at these levels is approximately 2%'.[14] He is referring to recent work by Marisa Carrasco and her collaborators showing that the apparent brightness contrast of attended stimuli is amplified by exogenously captured attention.[15] Similar points can be made regarding other sources of visual noise, such as small differences in apparent brightness resulting from noise in our pupillary response. Put in terms of our model $<W, R^*>$, Block's objection is that these data show that d is *smaller* than c, yet, on my proposal, it is larger.

Block's objection that our discrimination thresholds place an upper bound on the unspecificity of perceptual appearances seems tacitly to assume something like the principle that, if it is consistent with appearances that o_1 has magnitude m with value x and consistent with appearances that o_2 has magnitude m with value x, then it is consistent with what we perceive that both o_1 and o_2 have magnitude m with value x. But we should reject this principle. For familiarity, let us focus on spatial magnitudes. Suppose I am looking at two trees t_1 and t_2. I perceive that t_1 is just the slightest bit taller than t_2. Therefore, there is no height h such that it is consistent with what I perceive that both t_1 and t_2 have height h. If the aforementioned principle were true, it would follow that the appearances of the trees' heights must be extremely specific, since every height not ruled out for t_1 must be ruled out for t_2 and vice versa. But this is not what normally happens in such cases. The perception of differences with respect to some perceptible magnitudes is a distinct cognitive achievement from, and often does not go by way of, one's perceptions

[14]Block, 'Attention and Mental Pain', 52.

[15]Carrasco, Ling, and Read, 'Attention Alters Appearance'.

of the magnitudes that differ. We might naturally describe the tree case as follows: t_1 appears to be between 20 and 25 feet tall, t_2 appears to be between 19 and 24 feet tall, t_1 appears to be about a foot taller than t_2, and taking all of these appearances at face value yields knowledge. Our powers of discrimination often significantly outstrip our powers of detection, and there is nothing mysterious about this phenomenon.[16]

II. Epistemic Probability

A proposition's *epistemic probability* is how likely it is given what you know. Let us now consider what predictions <W, R> and <W, R*> make about epistemic probabilities.

It helps to think spatially. Think of elements of W as points in the Cartesian plane, the horizontal and vertical axes of which correspond, respectively, to the perceived magnitude's real value and to its apparent value (or the midpoint of its appearance, as appropriate). Think of propositions as regions in this plane. We have defined R in such a way that $R(<e, f>)$ always corresponds to a horizontal line segment in this plane centered on $<e, f>$; the same is true of $R^*(<e, f>)$. This fact allows us to use the natural length measure on line segments to define epistemic probabilities. For example, for any proposition p whose intersection with the line segment $R(<e, f>)$ is also a line segment, let the epistemic probability of p at $<e, f>$ in <W, R> equal the ratio of the length of this segment to the length of $R(<e, f>)$. In general, the epistemic probability of a proposition p at a world $<e, f>$ in <W, R> is the ratio of the total length of p's intersection with $R(<e, f>)$ to the total length of $R(<e, f>)$.[17] The same goes, *mutatis mutandis*, for R^*.

So interpreted, Williamson's model <W, R> predicts improbable knowing. Consider an arbitrary world $<e, e>$ in which appearances exactly match reality. $R(<e, e>) = \{<e^*, e>: |e^* - e| \leq c\}$, which in the Cartesian plane corresponds to the horizontal line segment with midpoint $<e, e>$ and length $2c$. This is the strongest proposition known at $<e, e>$. Now consider $KR(<e, e>)$, the proposition that $R(<e, e>)$ is known. It is straightforward

[16]The normality of appreciably unspecific appearances has non-trivial consequences in the philosophy of perception. For example, it is inconsistent with Michael Tye's identification of blurry vision with appreciable unspecificity in the visual appearances of the locations of objects' boundaries (since normal vision need not be blurry) (Tye, 'Blurry Image'). There are also strong independent grounds for rejecting Tye's view; see Allen, 'Blur'.

[17]In general, we let the epistemic probability of p at $<e, f>$ in <W, R> be $\lambda(p \cap R(<e, f>))/\lambda(R(<e, f>))$, where λ is the one-dimensional Lebesgue measure on the horizontal line through $<e, f>$. (If this quantity is undefined at any world, we let the epistemic probability of p be undefined at every world.) This induces a function from worlds to probability functions on the sub-algebra of propositions all of whose horizontal cross-sections are measurable, which includes all of the propositions that we will be concerned with.

to verify that KR($<e, e>$) = $\{<e, e>\}$. So the strongest proposition known at $<e, e>$ is known only at $<e, e>$. By the factivity of knowledge, the intersection of KR($<e, e>$) and R($<e, e>$) is just KR($<e, e>$) itself, i.e., $\{<e, e>\}$, which has length 0. Therefore, the epistemic probability of KR($<e, e>$) at $<e, e>$ in $<$W, R$>$ = $0/2c = 0$. R($<e, e>$) is known at $<e, e>$ *even though* at $<e, e>$ it has epistemic probability 0 of being known. So $<$W, R$>$ predicts not just improbable knowing but maximally improbable knowing.

Here is a more intuitive gloss on what is going on. In Williamson's model, as appearance and reality converge you get to know more and more about the perceived magnitude. As your epistemic achievement grows, it becomes less and less likely that you have managed such a feat. When appearances and reality match perfectly, what you know about the value of the magnitude could not be known in any other circumstances. Since infinitely many other circumstances are still epistemically accessible, your knowledge is maximally improbable. Williamson gives the following helpful heuristic: 'propositions known in only one world generate cases where the agent knows a truth *p* even though it is virtually certain on her own current evidence that she does not know *p*'.[18] Here is another helpful heuristic: in these models, when appearances match (or are centered on) reality, you get to know as much as it is possible to know about the perceived magnitude, and your knowledge is therefore as improbable as one's knowledge of the perceived magnitude can be.

Let us now turn to our model $<$W, R$^*>$. Consider again $<e, e>$, which now represents a world in which the magnitude's value lies halfway between the minimal and maximal values consistent with its appearance. R*($<e, e>$) = $\{<e^*, e>: |e^* - e| \leq d\}$ and KR*($<e, e>$) = $\{<e^*, e>: |e^* - e| \leq b\}$, as discussed in Section I. In other words, the perceived magnitude is known to be as it appears in $<e, e>$—i.e., to have a value somewhere in the interval $[e - d, e + d]$—just in case its value e^* is in the safe haven $[e - b, e + b]$. By the factivity of knowledge, the intersection of R*($<e, e>$) and KR*($<e, e>$) is just KR*($<e, e>$) itself, which in the Cartesian plane corresponds to the horizontal line segment with midpoint $<e, e>$ and length $2b$. Therefore, the epistemic probability of KR*($<e, e>$) at $<e, e>$ in $<$W, R$^*>$ = $2b/2d = b/(b + c)$.

The same holds for any world in which reality lies in the safe haven of the appearance: if $|f - e| \leq b$, then the probability of KR*($<e, f>$) at $<e, f>$ in $<$W, R$^*>$ = $b/(b + c)$. For all other $<e, f>$, the probability of KR*($<e, f>$) at $<e, f>$ in $<$W, R$^*>$ is equal to the probability of KR($<e, f>$) at $<e, f>$ in $<$W, R$>$: outside the safe haven, the two models agree on the probability of knowing the proposition that is in fact the strongest proposition known. We can think of $<$W, R$^*>$ as dividing worlds into the normal, 'good' cases where we know things to be as they appear, and the abnormal, 'bad' cases

[18] Williamson, 'Gettier Cases in Epistemic Logic', 8.

where we do not. The model agrees with Williamson's on the bad cases as regards the strongest proposition known and the probability that it is known. But in the good cases, the strongest proposition known is almost always weaker than in Williamson's model, owing to unspecific appearances, and the probability that it is known is $b/(b + c)$.[19]

So, in general, if $R^*(<e, f>) \subseteq p$, then the epistemic probability of Kp at $<e, f>$ in $<W, R^*>$ is at least $b/(b + c)$ (assuming it is defined). In other words, our model predicts a lower bound of $b/(b + c)$ on the epistemic probability of knowing a proposition that is in fact known. Therefore, so long as $b \geq c$, our model predicts no improbable knowing—no known proposition has less than 0.5 epistemic probability of being known. Let us now explore how b and c might be related.

As we observed earlier, both our margin for error c and the upper bound b on the size of normal deviations between the magnitude's real value and the midpoint of its appearance in effect parameterize the same quantity—namely, the amount of noise in our perceptual system. It is therefore natural to suppose, as an initial working hypothesis, that $b = c$. On this hypothesis, our model predicts no improbable knowing. Instead, it predicts that, in normal cases, the strongest proposition that we know about the magnitude's value has epistemic probability 0.5 of being known.

It is unclear how much comfort this will bring to those strongly inclined to reject improbable knowing. Is the compatibility of knowledge and epistemic probability 1 of ignorance predicted by Williamson's model significantly less plausible than the compatibility (and indeed normality) of knowledge and epistemic probability 0.5 of ignorance predicted by mine? I think it is. A proposition's having 0.5 epistemic probability is compatible with it being rational to invest significantly more confidence in it than in its negation; the connection between epistemic probability and rational confidence is not straightforward.[20] Indeed, there is reason to expect such divergence in this case, since it is plausible that our 'confidence density' should be centrally 'peaked', in contrast to the 'flat' probability density function used to model epistemic probabilities.[21]

Those who think that we can know only propositions that have at least epistemic probability x of being known, for some $x > 0.5$, could of course achieve this result by replacing our working hypothesis that $b = c$ with a stipulation that $b/(b + c) = x$. Williamson might complain that such stipulations are objectionably ad hoc—that such a blatant reverse engineering of our philosophy of mind to fit an antecedently preferred epistemology of

[19] Notice that this dichotomy classifies some veridical experiences as bad cases; compare Mark Johnston's discussion of 'veridical illusion'. Johnston, 'Better Than Mere Knowledge?'.

[20] See Williamson's reply to Mark Kaplan. Williamson, 'Replies to Critics'.

[21] Morrison, 'Perceptual Confidence'.

perception is not credible. We need to be careful how we construe this objection. Epistemology can inform the philosophy of mind, as it does prominently in *Knowledge and Its Limits* and centrally in Section I above. ('Knowledge first', to use Williamson's slogan.) So the objection cannot be merely a complaint about reverse engineering. There is something else problematic about such stipulations. If we simply identify b and c, then our model has only one free parameter, c, the choice of which makes no difference to the model's structure. But if we allow ourselves to independently choose b and c as we see fit, the structure of our model will depend on how we make the choice. In general, we should prefer models that limit such choice points, both on grounds of simplicity and because the more such choices we make the more likely it is that they will turn out to conflict with each other. I leave readers to judge for themselves how compelling such considerations are in this case.

At any rate, there are non-ad hoc reasons to think that b should be at least somewhat greater than c. I will now give an argument that b should be approximately twice c.[22] If this argument is sound, it predicts that the lower bound on (and normal value of) the epistemic probability of knowing the strongest proposition that one in fact knows is approximately 2/3—an 'improvement' on 0.5, many might say. The argument proceeds from the same thought that motivated the hypothesis that $b = c$—namely, that both parameters should encode the same degree of deviation from cases in which the parameter's appearance is centered on its real value—but advocates an alternative way of representing this thought formally. The informal idea is to understand 'the same degree of normal deviation' in probabilistic terms. Instead of simply equating b and c, we should choose their relative values so that both parameters make the same probabilistic contribution as regards the objective chances of different degrees of divergence between appearance and reality.

Assume that good cases—those in which we know that things are as they appear—are objectively probable. Let π be the small objective chance of bad cases. Since the good case is the normal case, we count a case as normal just in case π is less than or equal to the objective chance of the magnitude's real value and the midpoint of its appearance diverging by the amount that they do in that case. In terms of our model, we choose b so as to make the objective chance of the magnitude's real value differing by more than b from the midpoint of its appearance equal to π. We want c, our margin for error, to make the same probabilistic contribution as b, namely π. While b concerns the normal range of divergence from circumstances in which appearances are centered on reality, c concerns the potential for this divergence to be more extreme than it in fact is. We model this idea by letting c vary from world to world as a function of the difference between the magnitude's real value and

[22]This and the following three paragraphs can be skipped without loss of continuity.

the midpoint of its appearance. For every world $<e, f>$, we choose c at that world so as to make π be the objective chance of the magnitude's real value and the midpoint of its appearance differing by at least $|f - e| + c$ *given that* they differ by at least $|f - e|$.

To represent this proposal formally, we need to enrich our models with a probabilistic threshold π and a probability density function Ch representing the objective chances of different degrees of divergence between the magnitude's real value and the midpoint of its appearance. We assume that Ch approximates a Gaussian distribution (a 'bell curve') centered on 0. Assume further, for simplicity, that this distribution does not depend on the magnitude's value. Let δ be a random variable standing for the difference between the magnitude's real value and the midpoint of its appearance. We can now represent our stipulation about b by the equation $Ch(|\delta| > b) = \pi$. We capture the above proposal about c by letting it vary from world to world in accordance with the equation $Ch(|\delta| > |f - e| + c(<e, f>) \mid |\delta| \geq |f - e|) = \pi$. We keep our definition of R^* as before. The only difference is that c and d are now parameterized by π and Ch, and c is now a function of $|f - e|$. In this enriched model, c makes the same probabilistic contribution as b in the sense that, however objectively improbably the midpoint of appearances diverges from reality, our margin for error ensures that, for all we know, something even more objectively improbable (by a factor of π) has happened.

As before, we choose the radius d of the magnitude's appearance so that $\mathrm{R}^*(<e, f>) = \{<e^*, f>: |e^* - f| \leq d\}$ just in case $|f - e| \leq b$. This requires letting $d = b + c(<e, e + b>)$, for some e. (The choice of e is arbitrary since $c(<e, f>)$ depends only on $|f - e|$.) Equivalently, and more perspicuously: $Ch(|\delta| > d) = \pi^2$. In this class of enriched models, given reasonable (small) values for π, b will be approximately twice c, and therefore the lower bound on the epistemic probability of knowing a proposition that one in fact knows will be approximately 2/3.[23] For those familiar with confidence intervals in statistics, the intuition behind this result is this: for Gaussian distributions and reasonable 'p-values' π, the width of a confidence interval with p-value π is approximately 2/3 the width of a confidence interval with p-value π^2.

Before moving on, I should mention that one could hold a view according to which the lower bound x on the probability of knowing a proposition that is in fact known is subject- and/or context-sensitive—perhaps because what counts as 'normal' is a subject- and/or context-sensitive matter. For example, on the subject-sensitive version of the proposal, margins for error vary with the circumstances. In 'anti-skeptical' circumstances, c will be much less than d, x is close to 1, and we can knowledgeably take almost all veridical experiences at face value. In the anti-skeptical limit where our margin for error $c = 0$, we

[23] For example, for $\pi = 0.1$, $c = 0.57b$ and this lower bound is 0.64; for $\pi = 0.01$, $c = 0.51b$ and the lower bound is 0.66.

end up with the Hellie/Stalnaker/Cohen-Comesaña model mentioned previously, according to which being in a good case is a luminous condition. As we move to more 'skeptical' circumstances, our margin for error c grows and x shrinks. Indeed, as discussed in Section I, once c reaches (and after it exceeds) d we end up with Williamson's original models in which $x = 0$; i.e., in which there is maximally improbable knowing. Whether merely confining maximally improbable knowing to skeptical circumstances strikes philosophers leery of the phenomenon as an acceptable result I leave for them to judge. Parallel considerations apply to contextualist versions of the proposal.

III. Beyond Perception

The considerations discussed thus far were specific to perception. Can we use them to construct a general strategy for resisting improbable knowing? The prospects for such a generalization initially look dim. Williamson's models are intended to characterize not only our perceptual knowledge but also our knowledge of the values of continuous magnitudes (like temperature) gained by using measurement devices (like thermometers) whose readings deviate from the magnitudes' true values as a result of random noise. The magnitude of this noise often far outstrips the specificity of such readings—consider the fluctuations in the rightmost digits of many digital scales. In other words, the precision of such scales clearly outstrips their accuracy. The degree of imprecision in their measurements, if any, is swamped by the margin for error governing our measurement-based knowledge. Now recall that, when we were modeling perceptual knowledge, $<W, R> = <W, R^*>$ whenever c was greater than d. Unspecific appearances made no epistemic difference when they were swamped by perceptual noise: maximally improbable knowing was preserved. There is therefore a strong *prima facie* case for thinking that, whatever our account of perceptual knowledge, the fact that noisy scientific instruments sometimes yield extremely improbably accurate readings guarantees at least some cases of extremely improbable knowing.[24]

While there are undeniably superficial differences between taking unspecific perceptual appearances at face value and gaining inexact knowledge of the temperature by reading an overly precise thermometer, these differences do not mandate asymmetry at the level of deep epistemic structure. We need to think more carefully about the causal processes underlying the sort of noise that pervades measuring instruments. I will proceed in two stages. First, I argue that Williamson's model makes incorrect predictions about

[24]Some will resist improbable knowing in these non-perceptual cases by denying that we can ever gain knowledge of the temperature by reading a thermometer. Perhaps they follow Hume in holding that, although we can have non-inferential empirical knowledge, induction is always a matter of mere probabilities. I agree with Williamson that such views are objectionably skeptical, and so will set them aside.

our instrument-based knowledge. Second, I offer a different account of our instrument-based knowledge, motivating it with informal examples, and drawing out its formal parallels with our model of perceptual knowledge $<W, R^*>$.

III.i. Against Williamson's model

Suppose that our noisy digital scale is unbiased and subject to approximately Gaussian noise, in the following sense. The scale produces a new reading every second, and its reading on any given occasion of measurement is probabilistically independent of its previous readings, where the probabilities in question are objective chances. These reading are characterized by a Gaussian probability density function—a 'bell curve' centered on the object's true weight. In other words, every time the scale gives a reading, the probability that the reading is in the interval $[x, y]$ is equal to the area under this curve between x and y. If one waits long enough, then with probability 1 the normalized histogram of accumulated readings will approximate the bell curve to an arbitrary degree of accuracy.

The tails of the bell curve never drop off to zero, so errors of any size have *some* positive objective chance.[25] Yet, on pain of skepticism, we surely know *something* about what the scale will read when we weigh a tennis ball—e.g., we know that it won't register over 1,000 pounds. Knowing a proposition is therefore compatible with knowing that its negation has some positive objective chance. Furthermore, how much we can know depends in part on how sharply the bell curve is peaked. This 'width' of a bell curve is characterized by its *standard deviation*: the distance from its center to the inflection points on either side where the curve switches from being concave-down to concave-up. Insofar as what we can know about the scale's next reading is sensitive to the probabilistic profile of the setup, we can assume that it is sensitive only to the distribution's mean (which will be equal to the weighed object's true weight, since the scale is unbiased) and standard deviation, since a Gaussian probability density function is fully characterized by its mean and standard deviation.

I am assuming an approximately Gaussian distribution because such distributions are what we normally find in practice. This is no coincidence: processes that aggregate the results of a large number independent random processes will be characterized by approximately Gaussian probability distributions even if their component processes are not characterized by even approximately Gaussian distributions.[26] The best demonstration of this principle is the Galton board (Figure 1). A ball is dropped into a lattice

[25]This is an idealization, since, for example, there may be no chance of the scale indicating a negative weight. But the idealization is harmless for present purposes.

[26]This is a consequence of the Central Limit Theorem.

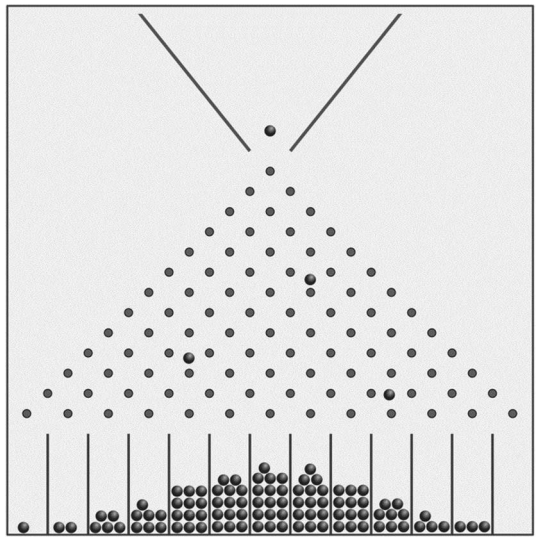

Figure 1. The Galton Board.

Source: Wikipedia.

of pegs. At every row, the ball has an equal probability of being deflected to the left or to the right. How it is deflected at one row is independent of how it was deflected on previous rows. The chance distribution characterizing the accumulation of these random deflections will be an approximately Gaussian distribution, as evidenced by the histogram generated by accumulated balls dropped from the same location. Indeed, in the infinite limit, the resulting distribution is exactly Gaussian. We approach this limit by repeatedly halving the spacing between the pegs while quadrupling the number of rows; this preserves the distribution's standard deviation.[27]

The source of the Gaussian noise we find in our scientific instruments is likewise the aggregation of a large number of independent random processes. So let us take the Galton board as our model. We will model our answer to the question what we can know about the next reading of our noisy scale on our answer to the question what we can know about where a ball will land when dropped into the Galton board. The cases are admittedly causally asymmetric. In the case of the scale, we infer causes from effects, while, in the case of the Galton board, we infer effects from causes. But this causal difference does not make an epistemic difference.[28] In this setting, Williamson's constraint

[27]In the infinite limit the balls' trajectories become Wiener processes, the mathematical idealization of Brownian motion.

[28]If you are worried that it might make a difference, substitute the following setup for the Galton board: Zeus throws down light bolts and they bounce off the clouds like balls bounce off the pegs

that 'any increase in the gap between appearance and reality has an epistemic cost for the agent: more knowledge is lost'[29] corresponds to the following prediction about the Galton board: how much we can know about where the ball will land is a monotonically decreasing function of the horizontal displacement between where it will in fact land and where it is dropped.

For Galton boards with enough rows, Williamson's anti-skepticism will presumably lead him to think that in cases in which the ball will fall straight down the middle we know something non-trivial about where it will land. (For simplicity, we can assume that the board has an even number of rows and ignore where the ball's trajectory passes odd numbered rows. If you are worried that knowledge of the future raises special difficulties, we could instead consider a ball that has already been dropped but whose landing location has not yet been observed.) More generally, if Williamson's model <W, R> aptly characterizes our knowledge of balls dropped into Galton boards, then the strongest thing we can know about where a given ball will land is that it will land within a radius $x + c$ of where it was dropped, where x is the (horizontal) displacement between where it is dropped and where it will in fact land and the margin for error c is determined by the standard deviation of the approximately Gaussian chance distribution over bins at the bottom of the board. Since the noise in normal scientific instruments is the result of processes relevantly analogous to the many deflections a ball takes as it falls through a lattice of pegs, Williamson's model must correctly describe our knowledge of where a ball dropped into a Galton board will land if it is to correctly describe the unspecific knowledge we gain through the normal use of scientific instruments.

But Williamson's model *does not* correctly describe our knowledge in the case of the Galton board. It incorrectly predicts that what we can know about where the ball will land is a function of where it will in fact land. But there is no such function. What we can know about where the ball will land depends not only on where it will land but also on the details of its trajectory. Williamson's model fails to account for this dependence.

This trajectory-sensitivity is a consequence of the extremely plausible assumption that, if a falling ball passes any row at a horizontal displacement of n pegs from where it was dropped, then, for all we know, it lands n pegs from where it was dropped. For example, if on row 82 the ball will fall through the 11th gap to the right of where it was dropped, then for all we know the ball will land 11 gaps to the right of where it was dropped.

To see the problem for Williamson's model, consider a case in which the ball will fall straight down and land directly below where it was dropped. By the aforementioned anti-skeptical principle, in this case we know

of a Galton board. We see that the tree next to us was just struck by lightning. What are we now in a position to know about where in the sky Zeus is, assuming we antecedently had no idea?

[29] Williamson, 'Gettier Cases in Epistemic Logic', 6.

something non-trivial about where the ball will land. By the analogue of Appearance Centering, there is a distance x such that, in this case, what we know about where the ball will land is that it will land no more than x from where it was dropped. Now consider a case in which the ball veers off to the right and then back to the left such that it lands directly below where it was dropped and, on some even-numbered row, the ball will be horizontally displaced from where it was dropped by more than x. The above principle of trajectory sensitivity, together with the analogue of Appearance Centering, now entail that in this case we do *not* know that the ball will land no more than distance x from where it was dropped. So what we know about where the ball will land is not a function of where it will land. *A fortiori*, it is not a monotonically decreasing function of the horizontal displacement between where it lands and where it was dropped: Williamson's model fails to accurately characterize our knowledge in the case of the Galton board. Since Galton boards are relevantly analogous to the instruments we normally use to measure physical magnitudes, Williamson's model fails to accurately characterize our inexact knowledge of physical magnitudes that results from the use of such instruments.

(One might try to get around this argument by stipulating that c—the radius of what we know about where the ball will land in cases where it falls straight down—is greater than half the number of rows of the Galton board. The idea is to rule out the existence of a trajectory where the ball veers to the right by more than c and then veers all the way back. This strategy fails because c is a function of the standard deviation of the Galton board's characteristic bell curve. And as mentioned above, quadrupling the number of rows while halving the distance between pegs preserves this standard deviation, but it doubles the total distance the ball travels as it is deflected back and forth. This operation guarantees the existence of a Galton board with enough room for the ball to make the necessary return trip.)

Williamson's model fails to do justice to the complexities characteristic of our knowledge of physical magnitudes gleaned through the use of noisy instruments, because it collapses all cases in which readings and reality agree into a single world. Such cases can differ epistemically: the inner workings of the measurement process can generate Gettier-like ignorance when, unbeknownst to us, our instrument's reading is accurate because two independent abnormal processes happened to cancel out. Such abnormalities can prevent our true predictions from amounting to knowledge. In so doing, they undermine the principle Williamson uses to argue for improbable knowing—namely, that an 'agent's ignorance increases as the gap between appearance and reality widens'. In the case of our noisy scale, this is simply not true.

Williamson could repair his argument for improbable knowing by finding some parameter other than the gap between appearance and reality such that our ignorance is a monotonically increasing function of this parameter and, in some world where this parameter achieves a particularly low value,

the epistemic probability that it achieves such a low value is low. But it is hard to identify any natural parameter of Galton-board trajectories of which our ignorance is a function; one can always concoct pairs of trajectories that agree on the candidate parameter but intuitively differ as regards what we can know. Rather than speculating about such parameters, I now give an alternative account of our instrument-based knowledge of physical magnitudes. As it happens, this alternative account does not predict improbable knowing.

III.ii. Normality

I know that I have a next-door neighbor, but I have not met her. How much can I know about her height? Presumably I know something non-trivial about her height—for example, I know that she is not over eight feet tall. But it is very odd to think that the closer she is to the median adult female height the more I am in a position to know about how tall she is. It seems much more natural to suppose that, as long as she has a normal height, what I know is that she does not have an extraordinary height.

This thought—that we do not get extra epistemic credit for getting closer and closer to some paradigm case—is especially strong in situations where it is not clear what the paradigm case would even be. For example, how much can I know about how recently my friend Peter did laundry? On a simple way of forcing this case into something like the <W, R> model we end up predicting that the more recently Peter did laundry the more I can know about how recently he did it. But this prediction is implausible. We are strongly inclined to think that whether (unbeknownst to me) Peter did laundry yesterday or the day before makes no difference to what I can know about how recently he did laundry. It is much more natural to think that, as long as it has not been abnormally long since he last did laundry, what I know is that it has not been extraordinarily long since he last did laundry. Perhaps it is normal to have done laundry in the past three weeks and extraordinary not to have done laundry in the past two months. Assuming so, my proposal is that, as long as Peter has done laundry in the last three weeks, what I know about when he last did laundry is that it was some time in the past two months. Maybe in cases where he has not done laundry in the past three weeks my ignorance about the time since he last did laundry is a monotonic function of the time since he last did it. But all normal cases are epistemically alike as regards my knowledge of the time since he last did laundry.

This sort of picture generalizes. *If things are normal, then what you know is that they aren't extraordinary; if things aren't normal, you know less.* Notice that <W, R*> has just this character if we think of the 'good cases' as normal and of illusion as extraordinary. In the case of the Galton board, we might similarly classify trajectories as normal and extraordinary. If the ball will take a normal trajectory, what you know is that it will not take

an extraordinary trajectory; if not, you know less. Considerations of normality supersede considerations of the divergence between appearance and reality. Such divergence may *constrain* normality: for example, any trajectory by which the ball lands 10 standard deviations from where it was dropped is abnormal, and indeed extraordinary. But as we saw, there can also be extraordinary cases in which the ball ends up right where it started. The same goes for measuring devices. Certain deviations between real and measured values are always abnormal, and certain of these are always extraordinary. But there can also be extraordinary cases where real and measured values coincide, even if the underlying peculiarities are inaccessible to us.

As we saw in Section II, the upshot of such models for issue of improbable knowing depends on the ratio of the measure b of the set of normal accessible worlds to the measure d of the set of non-extraordinary accessible worlds. It is hard to address this issue in abstraction from any particular setup. But insofar as improbable knowing is antecedently implausible, it does not seem objectionably ad hoc to conjecture that b/d will always exceed 0.5—in effect, that when things are normal, it is as least as epistemically probable as not that they are normal.[30] If the foregoing account is correct, then, although forming beliefs about the temperature on the basis of a noisy thermometer is psychologically quite unlike taking perceptual appearances at face value, this difference does not preclude a similarity at the level of epistemic structure. Williamson's case for improbable knowing is at best inconclusive.[31]

IV. Resisting Improbable Knowing

Other arguments can be given in support of the possibility of improbable knowing. Suppose I know each of a collection of propositions, but for each of them I fail to know that I know it. From these propositions I competently deduce their conjunction. The epistemic probability of my knowing this conjunction is low, even though each of its conjuncts is such that the epistemic probability of my knowing it is high. Assuming that knowledge is closed under competent deduction, I will know the conjunction even though it will be epistemically improbable that I know it. Since such cases are possible, so is improbable knowing.

This argument is dialectically weak. Philosophers inclined to resist improbable knowing tend to be those who also deny that knowledge is closed under

[30] Williamson disagrees; see his 'Response to Cohen, Comesaña, Goodman, Nagel, and Weatherson'.

[31] One might be tempted to go further and *identify* the non-normal worlds with the extraordinary worlds; this is akin to the Hellie/Stalnaker/Cohen–Comesaña proposal, and would render normality luminous. Insofar as Williamson's influential anti-luminosity argument is successful, this temptation will have to be resisted; nothing I have said threatens the application of that argument to normality.

competent deduction. Indeed, they tend to be philosophers who are generally skeptical of epistemic logic because of the idealizations that it involves—for example, that we know every logical consequence of any set of propositions that we know. They will respond to this argument with a shrug.

By contrast, Williamson's argument for improbable knowing cannot be so easily dismissed. For while one might reasonably doubt whether knowledge is always closed under competent deduction, it is extremely hard to deny that, when looking at a tree (for example), I know the conjunction of every proposition that I know *about how tall the tree is*. It is bizarre to suppose that, although I perceive that the tree is between 90 and 105 feet tall and I perceive that the tree is between 95 and 110 feet tall, I fail to perceive that the tree is between 95 and 105 feet tall. For a particular perceived or measured magnitude, our knowledge of propositions about the magnitude's value is normally closed under conjunction. Denying closure is therefore not a plausible way to resist Williamson's argument for improbable knowing from the inexactness of simple perceptual and instrument-based knowledge. His argument is important because it does not assume that knowledge is generally closed under competent deduction.

Another argument for improbable knowing appeals to the anti-luminosity of appearances. The models we have been considering predict that we know the *exact* width and midpoint of our unspecific perceptual appearances. But this is an idealization. In a more realistic model, we might introduce epistemically accessible worlds in which our appearances have different widths and midpoints than they actually have. Consider a world w where taking appearances at face value yields knowledge. Let $p = \{w^*$: for all x, if it is consistent with appearances in w^* that the perceived magnitude has value x, then it is consistent with appearances in w that the magnitude has value $x\}$. By the Appearance Constraint, $\mathrm{KR}(w) \subseteq p$. But by plausible symmetry considerations concerning our ignorance of the width and midpoint of our unspecific appearance, the epistemic probability of p at w is less than 0.5. Therefore, together with the Appearance Constraint, the anti-luminosity of appearances generates improbable knowing.

This argument has the same weakness as the argument from closure. It relies on modeling knowledge as a universal quantifier over a set of epistemic possibilities—something that anyone who rejects closure already denies. Absent some further story, no one who rejects closure need feel compelled by it. To be clear, I am not denying the principle that knowledge is closed under competent deduction, nor am I denying that the principle's intuitive appeal lends some support to improbable knowing. I am merely noting a sociological fact about contemporary epistemologists: these two arguments are unlikely to win improbable knowing many converts.

A potentially more persuasive argument for improbable knowing attempts to rehabilitate Williamson's original argument. Suppose we concede that Williamson's model $<\mathrm{W}, \mathrm{R}>$ does not characterize the actual structure of our

perceptual and instrument-based knowledge, because of the unspecificity of our perceptual appearances and the internal structure of the processes responsible for the noise in our scientific instruments. Arguably, it is nevertheless possible that there be agents with maximally specific noisy appearances, and possible that there be scientific instruments subject to noise as a primitive matter of fundamental physical law without any underlying mechanism. In such circumstances, my strategies for resisting Williamson's argument would be unavailable. So improbable knowing is arguably possible.

This argument concedes a great deal. Williamson argues not merely that improbable knowing is possible, but that it is commonplace. That improbable knowing might be widespread is part of what makes Williamson's argument so striking. The mere possibility of improbable knowing is not uninteresting, but it is much less disruptive to our general epistemological worldview if it can be confined to counterfactual situations.

In any case, both the perceptual and instrumental horns of this argument can be resisted. Regarding perception, as mentioned at the outset, one might think that it is part of the nature of perceptual appearances that they normally yield knowledge when taken at face value. If so, perhaps one could give a transcendental argument against the possibility of maximally specific noisy appearances. Regarding instruments, while perhaps there could be a thermometer that was subject to noise due to some alien and capricious law of nature, it is debatable whether we could use such a thermometer to gain knowledge of the temperature (as opposed to knowledge merely of the probabilities of various ranges of temperatures). Perhaps such knowledge depends on the divisibility of measurement episodes into normal, abnormal, and extraordinary as sketched in the previous section, and such physical laws may not support such a division. These considerations are of course speculative, but so too are the proposals to which they are addressed.

In any case, a more principled reply can also be given. Even if our perceptual appearances were maximally specific and our measurement devices were subject to noise that lacked any mechanistic basis, there would still be reason to deny that the structure of our knowledge was captured by <W, R>. The reason has to do with certain plausible general principles connecting knowledge and justified belief. These principles are the subject of the next section. They support a general argument against improbable knowing, in contrast to the case-by-case arguments of the preceding sections.

V. Justification and Dispositions to Know

Consider the following general strategy for modeling justified belief. We begin by laying down some constraints on what an agent knows, such as Margin for Error and Appearance Centering. We then consider an agent who knows everything that she can. Formally, this is achieved by defining epistemic

accessibility as the *smallest* reflexive binary relation on worlds consistent with the constraints we have laid down. (We have to make sure that our constraints define such a relation.) We assume that what our agent has justification to believe depends only on her appearances, perhaps because she cannot help having beliefs that depend only on her appearances. Moreover, we assume that, necessarily, our agent has justification to believe everything that she knows. It follows that, in any world, our agent has justification to believe everything that she knows in some world with the same appearances. And this is all that she has justification to believe, since any further beliefs could not possibly amount to knowledge given her appearances, and so would be mere speculation.

Formally, we implement this strategy by defining the *doxastic accessibility* relation S as follows: $<e, f>\mathrm{S}<e^*, f^*>$ if and only if, for all e^{**}, $<e^{**}, f>\mathrm{R}<e^*, f^*>$. Informally, the doxastically accessible worlds are those that are consistent with what is known in every world in which our agent has the same appearances that she actually has. For the class of models with which Williamson is concerned, this definition of doxastic accessibility agrees with his. The agent has justification to believe all and only the propositions true in all doxastically accessible worlds, and she believes all and only the propositions that she has justification to believe. In this sense, she is perfectly rational. Notice that doxastic accessibility can be read off the structure of epistemic accessibility. Justified belief is thereby characterized in terms of knowledge.

The doxastic accessibility relation induced in this way by Williamson's model $<\mathrm{W}, \mathrm{R}>$ is the following: $\mathrm{S}(<e, f>) = \mathrm{R}(<f, f>) = \{<e^*, f>: |f - e^*| \leq c\}$. In every world, the strongest proposition that the agent believes is identical to the strongest proposition that she knows in the world with the same appearances in which appearances matches reality, which is in turn identical to the proposition that appearances are as they in fact are and the magnitude's value differs from its appearance by no more than the margin for error c. At the end of the last section, we were considering whether $<\mathrm{W}, \mathrm{R}>$ accurately describes any possible agents' perceptual or instrument-based knowledge. I will now give an argument that it does not.

A belief is justified if it is well directed at its aim. Since the aim of belief is knowledge, it is natural to think that a belief is justified just in case it manifests a disposition to know in normal circumstances. This thought has considerable appeal, especially within Williamson's framework in which knowledge is taken as the basic notion in terms of which justification is to be explained. Rather than defending this view here, I will simply explore some of its consequences.[32]

[32]The view has been defended by Maria Lasonen-Aarnio, 'Unreasonable Knowledge', 2. As she puts it: 'Reasonableness is at least largely a matter of managing one's beliefs through the adoption of policies that are generally knowledge conducive, thereby manifesting dispositions to know

Williamson's models have all the structure we need to formally model this proposal. Since we are modeling an agent whose beliefs are determined by appearances, we can treat her cognitive dispositions as pairs of an appearance and the strongest proposition that she believes when she has that appearance. A cognitive disposition yields knowledge just in case this proposition is known. Circumstances are normal just in case appearances and reality diverge by no more than c.

We can now formalize the view that a belief is justified only if it manifests a disposition to know in normal circumstances:

> *Disposition to Know*: If $|f - e| \leq c$, then $R(<e, f>) \subseteq S(<e^*, f>)$.
> What you justifiably believe is known in all normal worlds with the same appearances.

Williamson's model $<W, R>$ violates Disposition to Know. It models an agent who is brazen in her beliefs. She always believes a proposition that she knows in only one world; in almost all normal cases, this belief fails to amount to knowledge. Brazenness is not a knowledge-conducive cognitive disposition. By Disposition to Know, beliefs so formed are therefore not justified.

The natural thing to do is to add Disposition to Know to our constraints on knowledge. Consider the model $<W, R'>$, where R' is the smallest binary relation on W consistent with Margin for Error, Luminous Appearances, Appearance Centering, and Disposition to Know. This is perfectly well defined, since S' is defined in terms of R' as above, and so Disposition to Know represents a coherence constraint on knowledge.[33] In $<W, R'>$, $S'(<e, f>) = R'(<f, f>) = \{<e^*, f>: |f - e^*| \leq 2c\}$. Indeed, it turns out that $<W, R'> = <W, R^*>$ for $c = b$—which, as we saw in Section II, fails to predict improbable knowing.

This result is not mysterious. The thought behind $<W, R'>$ is that believing what you have justification to believe normally yields knowledge. The thought behind $<W, R^*>$ was that taking perceptual appearances at face value normally yields knowledge. Moreover, the thought behind the normal/abnormal/extraordinary structures outlined in Section III was that, normally, the strongest thing that we know is something that it is normal to know. What these structures have in common is that they are models of iterated normality. In the case of $<W, R'>$, knowledge requires belief that is normally true, justification requires belief that is normally knowledge, and knowledge itself requires justification. These models are well confirmed: they can be motivated by seemingly unrelated considerations about perceptual knowledge, noisy instruments, bare estimations, and the nature of justification. Just as importantly, the models are non-disruptive. They do not

---

and avoid false belief across a wide range of normal cases.' See also Hawthorne and Srinivasan, 'Disagreement without Transparency'.

[33] i.e., $S'(<e, f>) =_{df} \{<e^*, f^*>: \text{for all } e^{**}, <e^{**}, f> R' <e^*, f^*>\}$.

predict skepticism, arbitrary iterations of knowledge, or violations of margin for error principles.

These models also vindicate our initial unease about the sort of improbable knowing predicted by Williamson's models. Such knowledge would involve brazen beliefs—beliefs that almost always fail to amount to knowledge. Brazen beliefs are not justified because an agent with healthy cognitive dispositions does not hold them, knowledge being the aim of her beliefs. Since knowledge requires justified belief, there cannot be such brazen improbable knowing. We therefore have strong abductive grounds to think that simple perceptual and instrument-based knowledge does not involve improbable knowing. Whether or not there are any cases of improbable knowing turns, in part, on the controversial question of whether knowledge is closed under competent deduction.[34]

(A more radical reply would be to reconcile <W, R> with Disposition to Know by adopting S' as a doxastic accessibility relation and thereby denying that knowledge requires justified belief. Maria Lasonen-Aarnio has recently defended such a view, arguing that once we accept Disposition to Know it would be unnatural and unprecedented to maintain a justified belief requirement on knowledge. But Williamson does not seem open to this sort of move, having elsewhere written: 'The idea of knowledge without justification should strike us as anomalous. . . . Knowing p is the central, unproblematic case of normative appropriateness in believing p.'[35])

VI. Conclusion

None of preceding discussion threatens the main conclusion of Williamson's paper: that general structural features of knowledge predict the existence of Gettier cases of both the 'original' and 'fake barn' varieties. Both sorts of cases are predicted by <W, R*>, for the same reason that they are predicted by <W, R>. This provides further confirmation of the robustness of the Gettier result.

Acknowledgements

Thanks to David Bennett, Cian Dorr, Peter Fritz, John Hawthorne, Harvey Lederman, and Tim Williamson for discussion of this material.

[34]Williamson defends this closure principle in his reply to Hawthorne and Lasonen-Aarnio. Williamson, 'Replies to Critics'; Hawthorne and Lasonen-Aarnio, 'Knowledge and Objective Chance'.

[35]Williamson, 'On Being Justified', p. 112; see also his discussion of 'knowledge-maximizing policies' in his 'Response to Cohen, Comesaña, Goodman, Nagel, and Weatherson'.

References

Allen, Keith. 'Blur'. *Philosophical Studies* 162 (2013): 257–73.

Block, Ned. 'Attention and Mental Pain'. *Philosophical Issues* 20 (2010): 23–63.

Carrasco, Marisa, Sam Ling, and Sarah Read. 'Attention Alters Appearance'. *Nature Neuroscience* 7 (2004): 308–13.

Cohen, Stewart, and Juan Comesaña. 'Williamson on Gettier Cases and Epistemic Logic'. *Inquiry* 56 (2013). doi: 10.1080/0020174X.2013.775012.

Hawthorne, John and Maria Lasonen-Aarnio. 'Knowledge and Objective Chance'. In *Williamson on Knowledge*, ed. Patrick Greenough and Duncan Pritchard, 192–208. Oxford: Oxford University Press, 2009.

Hawthorne, John, and Amia Srinivasan. 'Disagreement without Transparency: Some Bleak Thoughts'. In *The Epistemology of Disagreement: New Essays*, ed. David Christensen and Jennifer Lackey. Oxford: Oxford University Press, 2013.

Hellie, Benj. 'Noise and Perceptual Indiscriminability'. *Mind* 114 (2005): 481–508.

Johnston, Mark. 'Better Than Mere Knowledge? The Function of Sensory Awareness'. In *Perceptual Experience*, ed. Tamar Szabó Gendler and John Hawthorne, 260–90. Oxford: Oxford University Press, 2006.

Lasonen-Aarnio, Maria. 'Unreasonable Knowledge'. *Philosophical Perspectives* 24 (2010): 1–21.

Morrison, John. 'Color in a Physical World: A Problem Due to Visual Noise'. *Mind* 121 (2012): 333–73.

Morrison, John. 'Perceptual Confidence'. Unpublished.

Stalnaker, Robert. 'On Hawthorne and Magidor on Assertion, Context, and Epistemic Accessibility'. *Mind* 118 (2009): 399–409.

Stazicker, James. 'Attention and the Indeterminacy of Visual Experience'. Ph.D. Diss., University of California, Berkeley, 2011.

Tye, Michael. 'Blurry Image, Double Vision and Other Oddities: New Problems for Representationalism?'. In *Consciousness: New Philosophical Perspectives*, ed. Quentin Smith and Aleksandar Jokic, 7–32. Oxford: Oxford University Press, 2003.

Williamson, Timothy. 'Improbable Knowing'. In *Evidentialism and Its Discontents*, ed. Trent Dougherty, 147–164. Oxford: Oxford University Press, 2011.

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

Williamson, Timothy. 'Knowledge and Skepticism'. In *The Oxford Handbook of Contemporary Philosophy*, ed. Frank Jackson and Michael Smith, 681–700. Oxford: Oxford University Press, 2005.

Williamson, Timothy. 'On Being Justified in One's Head'. In *Rationality and the Good: Critical Essays on the Ethics and Epistemology of Robert Audi*, ed. Mark Timmons, John Greco, and Alfred Mele, 106–22. Oxford: Oxford University Press, 2007.

Williamson, Timothy. 'Probability and Danger'. *Amherst Lecture in Philosophy* (2009): 1–35. http://www.amherstlecture.org/williamson2009/ (accessed March 21, 2013).

Williamson, Timothy. 'Replies to Critics'. In *Williamson on Knowledge*, ed. Patrick Greenough and Duncan Pritchard, 279–384. Oxford: Oxford University Press, 2009.

Williamson, Timothy. 'Very Improbable Knowing'. *Erkenntnis* (forthcoming).

Williamson, Timothy. 'Gettier Cases and Epistemic Logic'. *Inquiry* 56 (2013). http://dx.doi.org/10.1080/0020174X.2013.775010.

Williamson, Timothy. 'Response to Cohen, Comesaña, Goodman, Nagel, and Weatherson on Gettier Cases in Epistemic Logic'. *Inquiry* 56 (2013). doi: http://dx.doi.org/10.1080/0020174X.2013.775016.