THE NORMALITY OF ERROR

(Forthcoming, *Philosophical Studies*)

Abstract

Formal models of appearance and reality have proved fruitful for investigating structural properties of perceptual knowledge (Williamson 2013, Goodman 2013, Cohen and Comesaña 2013, Dutant and Rosenkranz 2019). This paper applies the same approach to epistemic justification. Our central goal is to give a simple account of The Preface, in which justified belief fails to agglomerate. Following Greco 2014, Goodman and Salow 2018, Beddor and Pavese 2019 amongst others, we understand knowledge in terms of normality. An agent knows p iff p is true throughout all relevant normal worlds. To model The Preface, we appeal to the normality of error. Sometimes, it is more normal for reality and appearance to diverge than to match. We show that this simple idea has dramatic consequences for the theory of knowledge and justification. Among other things, we argue that a proper treatment of The Preface requires a departure from the internalist idea that epistemic justification supervenes on the appearances and the widespread idea that one knows most when free from error.

1 Introduction[1]

How things are and how things appear need not be the same. Mirages, magic tricks and myopia can intervene between the two. What we know about how things are depends on how much reality departs from appearance.

A number of recent papers illuminate questions about knowledge in terms of the relationship between appearance and reality (Williamson 2013, Weatherson 2013, Goodman 2013, Cohen and Comesaña 2013, Williamson 2014, Carter 2019, and Dutant and Rosenkranz 2019).[2] This body of work shares a simple assumption: the greater the difference between how things are and how they appear, the less is known. We never know more than we do in the good case, where reality and

[1]We are grateful to Bob Beddor, Julien Dutant, Jeremy Goodman, John Hawthorne, Dan Hoek, Ben Holguin, Nico Kirk-Giannini, Lauren Lyons, and Ezra Rubenstein, along with audiences at Rutgers and the Melbourne Logic Seminar. We are particularly grateful to an anonymous referee at *Philosophical Studies*.

[2]For related models of statistical knowledge, see Dorr et al. 2014 and Goodman and Salow 2018. As in economics, the hope is that these models will provide insight into the target phenomenon, not that they capture exceptionless generalizations:

In keeping with this literature, throughout we make at least three significant idealizations. First, as usual in epistemic logic, we assume the relevant agents are logically omniscient. Second, we assume that our agents are cognitively unimpaired and attentive. Finally, we focus primarily on cases of perceptual knowledge, where the agent receives perceptual information about the world and makes deductions on this basis.

appearance coincide. As we depart from the good case, we know increasingly little.

Reduction is the thesis that to be justified is to be in a state indiscriminable from knowledge. Defenders of Reduction include Lenzen 1978, Stalnaker 2006, Halpern et al. 2009, Williamson 2013, and Rosenkranz 2018; for related views, see Bird 2007 and Ichikawa 2014.

Reduction $Jp \dashv\vDash \neg K \neg Kp.$

S is justified in believing p iff it is not the case that S knows that it is not the case that S knows that p.

On the above picture of knowledge we know most when reality matches appearance. Against this background, Reduction offers a simple account of justification. It is always consistent with what we know that reality matches appearance. So Reduction implies that one is justified in believing p iff one knows p in the good case (the case where appearance and reality coincide).

While simple, this picture has at least one shortcoming. Knowledge agglomerates: if one knows p and knows q, then one knows $p \wedge q$. So, if we know most when appearance and reality match, then justification agglomerates too. If one is justified in believing p and justified in believing q, then in the good case, one knows both. If in the good case one knows both, then in the good case, one knows their conjunction. Hence one is justified in believing $p \wedge q$.

Justification Agglommeration $Jp, Jq \vDash J(p \wedge q).$

If S is justified in believing p and justified in believing q, then S is justified in believing p and q.

The problem is, Justification Agglomeration is widely taken to be untenable. Consider The Preface.

The Preface Alex is a historian. She is just finishing a book about Napoleon. The main body of the book contains 999 carefully researched claims about his life and times. For each claim, Alex has gathered several pieces of evidence. Yet despite her rigor, some claim in the book is wrong. Furthermore, reflecting on the fallibility of historical investigation, Alex recognizes that even the most carefully researched books, if sufficiently long, tend to contain some errors. She adds a final claim as a preface of her book: 'Each claim in this book is carefully researched, but at least one is false.' (Makinson 1965)

The agent in The Preface gathers information from a variety of sources. Each source is reliable, but no source is perfectly reliable. Since each source is reliable, she is justified in believing each of the 999 claims in the body. Yet since no source is perfectly reliable (and, indeed, some source is wrong), she is justified in believing the claim in the preface.[3] Accordingly, she is justified in believing each of the 1000 claims in the preface and body of her book. Yet the conjunction of the claims in the preface and body of her book is a contradiction and no agent is justified in believing that. The preface agent is a counterexample to Justification Agglomeration.

Reduction and the above account of knowledge are jointly incompatible with judgements about The Preface. One response would relinquish Reduction. Yet Reduction has certain advantages over other theories of justification. Consider The Lottery.

The Lottery Billy buys a single ticket in a 999 ticket fair lottery. Reflecting that her ticket has a chance of $\frac{1}{999}$ to win, Billy concludes, regretfully, that it is a loser. Reflecting that all the tickets in the lottery have an equal chance to win, Billy concludes, regarding each, that it is also a loser. Yet Billy knows that not every ticket will lose. Hence she also concludes that some ticket is a winner. (Kyburg 1961)

We suggest that the lottery agent lacks some justification which the preface agent possesses. Whereas the preface agent is justified in believing, of each of the 999 claims in the main body of her book, that it is true, the lottery agent is not justified in believing, of each of the 999 tickets in the lottery, that is a loser. (See Nelkin 2000, Bird 2007, (p. 101), Sutton 2007 (p. 48-50), Smith 2010, 2016, Littlejohn 2012, and Ichikawa 2014 among others, for discussion.) This might strike some as puzzling—what can distinguish the Preface and Lottery cases?

Reduction draws the right line. The Lottery agent knows that she does not know of any ticket that it is a loser (cf. Hawthorne 2004). Not so the preface agent. For each claim in the preface and body of her book, it is compatible with what Alex knows that she knows that claim. Hence, Reduction rules out justification in The Lottery while granting it in The Preface. For this reason, we would like to retain Reduction.

Since The Preface gives rise to failures of Justification Agglomeration, this requires revising the simple picture of how knowledge depends on the relationship between appearance and reality. In particular, it requires us to deny the principle that an agent always knows strictly more when appearance and reality coincide than she would otherwise. When appearance and reality diverge, an agent is subject to error.[4] On our view, unlike on the simple picture, being subject to

[3]In §4.3, we will entertain (though not endorse) the view that whether one is justified in believing one is subject to error depends on how much error one is subject to. We specify that in fact some source is wrong in order not to prejudge this issue.

[4]There is second, common sense in which an agent can be described as being in error when

error need not decrease what an agent can know.

We take it to be relatively uncontroversial that the preface agent cannot epistemically rule out that she knows any claim in the body or preface of her book. Nevertheless, it remains an open question what claims, if any, the preface agent in fact knows.[5] Our investigation suggests an answer. What knowledge is available in cases like The Preface varies depending on how much error the agent is subject to.

We propose that the relationship between appearance and reality can be productively framed in terms of normality. Every appearance can be thought of as the output of an instrument which provides information of some kind about the world. When an appearance is considered in isolation, it is abnormal for it to differ from reality. As we discuss below, the basic appearance/reality model in Williamson 2013 then says that an agent knows that the world is not significantly less normal than it is. The Preface involves a case where an agent possesses multiple appearances from different sources. In each case, considered in isolation, it would be abnormal for the outputs of the instrument (say, historical records) to differ from reality. Equally, however, when all sources of information are considered together, it would be abnormal for none to be subject to any error.

Our model of The Preface implies that there is no single ordering of worlds with respect to normality which is relevant for determining what an agent knows. Rather, we must consider multiple such orderings. An agent knows a proposition iff every world at which it is false is significantly less normal than the world she occupies according to some ordering or other.

2 Appearance/Reality Models

Imagine a thermometer that reports the temperature in a room. Outputs of the thermometer can be treated as apparent values of the room's temperature. In the good case, the thermometer is free from error. Its output coincides with the room's real temperature. In a bad case, however, the thermometer is subject to some error. Its output and the room's real temperature differ.

On a simple model of the situation, what an agent knows about the temperature by consulting the thermometer can be treated as a function of how bad the case is. As the thermometer's level of error increases, the amount that is known from it goes down. This idea is implemented by the basic appearance/reality model which we introduce in this section, from Williamson 2013.[6]

she has beliefs which are false. We will we employ 'error' exclusively in the former sense, to denote the failure of appearances to match reality.

[5] Littlejohn and Dutant 2020 offer the only extended discussion of this issue that we know of.

[6] Throughout, we follow Williamson 2013 in making the simplifying assumption that the agent knows everything that they are in a position to know. Readers who dislike this aspect of the idealization can treat the model as a model of what we are in a position to know.

2.1 The basic model

A possible world is a pair of real numbers, (r, a). r represents the real value of a quantity (e.g., the temperature in a room, the weight of an object, the height of a tree,...). a represents the apparent value of that quantity according to some instrument (e.g., a thermometer, a set of scales, the human visual system,...). What an agent knows about the real value of the quantity, given its apparent value, is represented by an accessibility relation, R. $R(r, a)$ is the set of worlds R-accessible from (r, a). This is the strongest proposition known at (r, a).

In the basic model, epistemic accessibility is determined by the distance between reality and appearance, $|r - a|$. An agent can rule out that a quantity takes the real value r' iff the distance between r' and the quantity's apparent value is significantly greater than the distance between the actual real value of the quantity and its apparent value. How much greater is significantly greater? We can introduce a margin for error constant m to represents this magnitude (with the assumption that $m > 0$). This gives us the accessibility relation in Distance:

Distance $\quad R(r, a) = \{(r', a) : |r' - a| \leq |r - a| + m\}.$

A world (r', a) is accessible from (r, a) iff the level of error at (r', a) is no more than m greater than the level of error at (r, a).

Before considering what an agent is justified in believing, we review some of the structural properties of the basic model. First, Distance implies that appearances are luminous. If the value of a quantity appears to be a, it is known to appear to be a.

Appearance Luminosity $\quad$ If $(r', a') \in R(r, a)$, then $a = a'$.

Appearance Luminosity is an idealization. It generates a simple model of perceptual knowledge by abstracting away from ignorance about perceptual appearances. Where apparent values are identified with the output of a measuring device, such as a thermometer or set of scales, it appears especially innocuous. As we'll see, Appearance Luminosity remains compatible with the failure of various internalist principles about knowledge and justification.

Since we assume Appearance Luminosity, we focus on knowledge of the real value of a quantity. $\mathsf{Real}(r, a)$ is the set of real values r' which are compatible with what is known at (r, a):

$$\mathsf{Real}(r, a) = \{r' : (r', a) \in R(r, a)\}.$$

We can represent knowledge of real values spatially. Figure 1 represents what an agent knows about the temperature in a room, given an apparent value of 75°F and a margin for error of 2.5°F. The axis plots different real values. Colored lines indicate the upper and lower limits of accessibility given a specified real value.

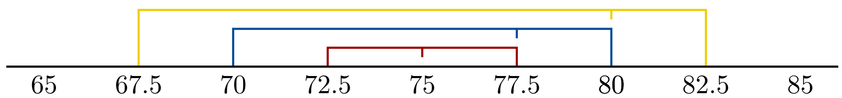

Figure 1: $a = 75$, $m = 2.5$

Holding the apparent value fixed, R returns a nested system of spheres when applied to worlds which vary regarding the real value of a quantity. In Figure 1, the strongest proposition known is what is known at the good case, where $a = r = 75$ (indicated by the red line). For any value $k \neq 75$, $R(75, 75) \subset R(k, 75)$. For example, where $r = 77.5$, the strongest proposition known is that the temperature is between 70 and 80 degrees (indicated by the blue line).

It is easy to see that knowledge is not luminous in the basic model.

Knowledge Luminosity $\quad Kp \vDash KKp$

Knowledge Luminosity fails because the basic model constrains knowledge with a margin for error. In Figure 1, $\mathsf{Real}(75, 75) = [72.5, 77.5]$. When the temperature is 75, the agent knows it is between 72.5 and 77.5. But $\mathsf{Real}(77.5, 75) = [70, 80]$. If the temperature were 77.5, the agent wouldn't know it was between 72.5 and 77.5. So at $(75, 75)$ the agent knows the temperature is between 72.5 and 77.5 degrees. But she doesn't know that she knows this.

2.2 Justification

We now turn to justification. As noted in the introduction, Williamson 2013 accepts Reduction (repeated below). Let M represent the dual of K:

Reduction $\quad Jp \dashv\vDash MKp$

An agent is justified in believing p just in case there is a world compatible with what she knows in which she knows p.

Given Distance, Reduction offers an appealingly simple account of justification. Distance implies that every world can see a corresponding good case with the same apparent value. Amongst worlds which agree in their apparent value, the agent knows strictly more in the good case where reality matches that appearance than in any case where the two diverge. Accordingly, Reduction implies that one is justified in believing p iff p is known in the corresponding good case. One is justified in believing all and only what one knows when the instrument producing one's appearances is free from error.

In the basic model, justification can be characterized in terms of an accessibility relation. An agent is justified in believing p at (r, a) iff p is true throughout $B(r, a)$. B can then be defined in terms of knowledge.

$$B(r, a) = R(a, a) = \{(r', a) : r' \in [a - m, a + m]\}$$

For instance, in Figure 1, at $(82, 75)$ the agent is justified in believing the temperature is between 72.5 and 82.5, because it is epistemically possible they know this, since it is possible they are at $(75, 75)$ and at $(75, 75)$ they know the temperature is between 72.5 and 77.5.

This provides an elegant model of justification. In particular, it implies that justification is consistent. An agent is never justified in believing p and justified in believing $\neg p$.

Consistency $\quad Jp \models \neg J \neg p$

Given Reduction, Consistency is equivalent to the Geach rule.[7]

Geach $\quad MKp \models KMp$

Geach corresponds to the condition that epistemic accessibility be convergent: the worlds accessible from any two accessible worlds must overlap.[8]

Convergence $\quad$ If wRv & wRu, then $\exists z : vRz$ & uRz

The basic model implies that epistemic accessibility is convergent because whenever (r, a) can access two other worlds (r', a) and (r'', a), each of these worlds can access the good case (a, a).

Unfortunately, this elegance is achieved at a cost. The basic model cannot model The Preface. As noted in the introduction, it validates Justification Agglomeration (repeated below). If an agent is justified in believing p and justified in believing q, then she is justified in believing $p \wedge q$.

Justification Agglommeration $\quad Jp, Jq \models J(p \wedge q)$

[7] Suppose S is justified in believing p, so that MKp holds. Then Consistency says that S is not justified in believing p, so that $\neg MK\neg p$ holds, which is just KMp.

[8] Suppose that MKp. Then at some v accessible from w, p holds throughout. But by Convergence, every world u accessible from w will see some world in Rv, itself contained in p. So at every world u accessible from w, p is possible. So KMp.

The agent is justified in believing p at (r, a) iff p is known at (a, a). But the set of propositions known at (a, a) is closed under conjunction, since knowledge itself agglomerates. So whenever p and q are known at (a, a), $p \wedge q$ is too.[9]

Consistency is plausibly a property of justification. If an agent is justified in believing a proposition, she is not justified in believing its negation. However, The Preface strongly suggests that agglomeration is not. An agent may be justified in believing two propositions, without being justified in believing their conjunction. Our task below is to develop a model of justification in terms of appearance and reality which rejects Justification Agglomeration while retaining Consistency. Before doing so, however, it will be helpful to consider a more general, philosophical characterization of the basic model.

2.3 Normality

Appearance/reality models offer a productive paradigm within which to theorize about perceptual knowledge. However, they do not readily generalize to knowledge of other kinds. In this brief section, we show how such models can be subsumed within a broader class of models which characterize what an agent knows in terms of what is normal. This will guide our models of The Preface in the remainder of the paper.

We can appeal to normality to frame the relationship between appearance, reality and knowledge. Following a number of recent authors (Goodman 2013, Greco 2014, Goodman and Salow 2018, Beddor and Pavese 2019, Carter 2019), say an agent knows p when p is true throughout all sufficiently normal worlds. More precisely, an agent knows p at w when p is true at any world that is not significantly less normal than w.[10] Where w' is no more normal than w, $w' \leq w$. Where w' is significantly less normal than w, $w' \ll w$.

Let $w' \not\ll w$ iff it is not the case that $w' \ll w$ and either $w \leq w'$ or $w' \leq w$. Then $R(w)$ is the set of $\leq$-related worlds that are not significantly less normal than w.

Knowledge of Normality $\qquad R(w) = \{w' : w' \not\ll w\}$

[9] Reduction implies that Justification Agglomeration is equivalent to the requirement that MKp and MKq entail $MK(p \wedge q)$. Within epistemic logic, this condition is equivalent to the following.

Agg. $\forall w \forall v \forall u(v, u \in R(w) \supset \exists z \in R(w) : R(z) \subseteq R(v) \cap R(u))$

After all, suppose MKp and MKq are true at w. Then w can access a world v where p is known, and a world u where q is known. Agg implies that w can also access a world z where at least as much is known as at both v and u, and so Agg implies that $MK(p \wedge q)$ is true at w.

[10] See Goodman and Salow 2018 and Beddor and Pavese 2019 for alternative conceptions of knowledge in terms of sufficient normality that differ from Williamson 2013 and our own model by validating Knowledge Luminosity. Goodman and Salow 2018, fn10 observe that their theory can be made to satisfy a margin-for-error constraint, thereby invalidating KK, in a manner very similar to that described in this section. We are grateful to an anonymous referee for pointing this out to us.

Distance is a special case of Knowledge of Normality, given two assumptions. First, error is abnormal. The greater the distance between appearance and reality, the less normal conditions are.

$$(r, a) \leq (r', a') \text{ iff } |r - a| \geq |r' - a'| \text{ and } a = a'.$$

It is more normal for reality to match appearance than for reality and appearance to diverge slightly. And its more normal for reality and appearance to diverge slightly than to diverge dramatically (for discussion, see Goodman 2013 and Carter 2019).

Second, significant differences in error amount to significant differences in normality. One world is significantly less normal than a second iff the level of error at the first is significantly greater than the level of error at the second. As above, we define significant difference in terms of the margin for error m.

$$(r, a) \ll (r', a') \text{ iff } |r - a| > |r' - a'| + m \text{ and } a = a'.$$

Given these assumptions, Distance and Knowledge of Normality are equivalent.

In the basic model, we can define justified belief in terms of the most normal worlds. An agent is justified in believing p when p is true at any world that is not significantly less normal than the most normal worlds. Let $\mathsf{MostNormal}(w) = \{w' : w \leq w' \wedge \forall w'' : w' \leq w'' \supset w'' \leq w'\}$. That is, $\mathsf{MostNormal}(w)$ is the worlds which are $\leq$-related to w which are maximally normal. Given our constraints, there is precisely one such world (a, a), for each (r, a). Then B can be defined as follows:

$$B(w) = \{w' : \forall w'' \in \mathsf{MostNormal}(w) : w' \not\ll w''\}$$

Alternatively stated, one is justified in believing that things are no more than a little bit abnormal (Leplin 2009, Smith 2016, §5.2, Goodman and Salow 2018). Clearly, this agrees with the basic model that $B(r, a) = R(a, a)$.[11]

[11] An alternative interpretation of the basic model appeals to modal distance instead of normality. Here, an agent knows only those propositions which are the case throughout the set of worlds which fall within a specified modal distance of her location. According to a standard gloss, w is closer to v than it is to z just in case what is the case at w could have been the case at v more easily than it could have been the case at z. Given this gloss, the ideology of modal distance offers a simple way of conceptualizing of a safety-theoretic account of knowledge, on which S knows that p iff p could not easily have been false (Sosa 1999, Williamson 2000, 2011, 2014, Pritchard 2005, 2007).

Unfortunately, it seems that Distance cannot be understood in terms of modal distance. Let the modal distance between (r, a) and (r', a') be equal to the sum of the difference between r and r' and the difference between a and a'.

$$d((r, a), (r', a')) = |r - r'| + |a - a'|.$$

Distance implies that what an agent knows cannot be characterized in terms of d. The

This concludes our presentation of the basic model. We now turn to considering how to extend it to give an adequate model of justification.

3 A new model of justification

Where does the basic model go wrong with respect to modelling The Preface? The basic model represents what we know about some feature of the world on the basis of some source of information. In The Preface, an agent receives information from multiple, independent sources about multiple, independent features of the world. At first glance, then, one might think the basic model just isn't designed to characterize cases with the level of complexity required to generate preface-style phenomena.

We agree that failures of justification agglomeration arise only in cases with greater complexity than the basic model. But this does not fully explain why the basic model is inadequate. In this section, we demonstrate that the most conservative extension of the basic model to incorporate multiple quantities continues to validate Justification Agglomeration.

To model The Preface, we need to represent the epistemic position of an agent who receives reports on the value of multiple quantities, each from a different source. We do this by introducing sequences of appearance/reality pairs, which now play the role of possible worlds.

Each sequence $s = \langle (r_{s_1}, a_{s_1}), \ldots, (r_{s_n}, a_{s_n}) \rangle$ is an n-tuple of pairs of real values. Each pair in the sequence represents an agent's epistemic position with respect to some quantity. We can think of a sequence as characterizing a system made up of multiple distinct instruments, each reporting on different quantities. For example, a system might represent both the temperature and the volume of a fluid or the color and the weight of an object.

We introduce a new accessibility relation, $\mathbf{R}$, over sequences of appearance/reality pairs. $\mathbf{R}(s)$ is the strongest proposition that the agent knows at s. We begin by considering the simplest extension of the basic model to knowledge of multiple quantities. This validates Justification Agglomeration and so we reject it.

3.1 Intersectivism

A natural idea is that when an agent receives information about multiple quantities, she does not know less about any given quantity than she would have if she received only information about that quantity. This reflects the fact that in the

set of worlds R-accessible from (r, a) does not correspond to a region of points within some distance of (r, a). Consider the model in Figure 1. (83,75) is d-closer to (90,75) than it is to (62,75). But Distance implies that though (83,75) can see (62,75), it cannot see (90,75). For this reason, we embrace a normality interpretation of Distance instead of an interpretation in terms of modal distance.

cases we are modeling, the values of various quantities are independent. Learning about the value of one quantity cannot lead an agent to lose knowledge about another quantity. We represent this constraint as a bridge principle connecting $\mathbf{R}$ with R. Call this Pointwise Accessibility:

Pointwise Accessibility $\quad s' \in \mathbf{R}(s)$ only if $\forall i : (r_{s'_i}, a_{s'_i}) \in R(r_{s_i}, a_{s_i})$

Pointwise Accessibility says that if s' is $\mathbf{R}$-accesible from s, then each pair in s' must be R-accessible from the corresponding pair in s.

The null hypothesis for extending the basic model is that $\mathbf{R}$ is the weakest relation satisfying Pointwise Accessibility. On this proposal, an agent's knowledge about multiple quantities is the intersection of her knowledge of each individual quantity.

Intersectivism $\quad s' \in \mathbf{R}(s)$ iff $\forall i : (r_{s'_i}, a_{s'_i}) \in R(r_{s_i}, a_{s_i})$

One sequence s is accessible from a second s' iff Distance implies that each pair in s is accessible from the corresponding pair in s'.[12]

Consider an agent who receives independent reports on the temperature from thermometers in two different rooms. Suppose the temperature appears 75˚F in both rooms, and suppose the margin for error $m = 2.5$˚F. Figure 2 represents this case. The x-axis represents the real value of the first pair (the temperature in room 1), and the y-axis the real value of the second pair (the temperature in room 2).

We distinguish different good and bad cases. $s' = \langle(75, 75), (70, 75)\rangle$ is a good case with respect to room 1, but a bad case with respect to room 2. The agent knows that room 1's temperature is in [72.5,77.5], and knows that room 2's temperature is in [67.5, 82.5] —that is, the agent at s' knows she is in the red region. $s'' = \langle(70, 75), (75, 75)\rangle$ is the reverse of s'. It is a bad case with respect to room 1, but a good case with respect to room 2. The agent at s'' knows she is in the yellow region. Finally, $s = \langle(70, 75), (70, 75)\rangle$ is a bad case with respect to both locations. An agent at s knows she is in the blue region.

Unfortunately, Intersectivism is incapable of modelling the kinds of cases we are interested in. Consider the principle Dominance, which is validated by Intersectivism.

Dominance $\quad$ If, for all i: $a_{s_i} = a_{s'_i}$ and $a_{s_i} = r_{s_i}$, then $\mathbf{R}(s) \subseteq \mathbf{R}(s')$.

Dominance says that an agent knows at least as much in the global good case as she does in any other case with the same appearances. Put another way,

[12]See Stalnaker 2006, §7 for a similar proposal.

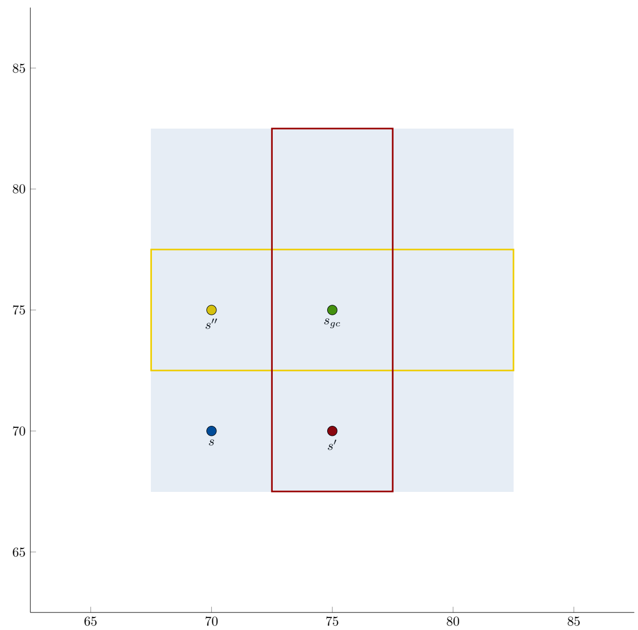

Figure 2: $a_1 = a_2 = 75, m = 2.5,$

you never know more than you do when you're free from error. Intersectivism implies Dominance. The region accessible from any global good case is strictly contained within the region accessible from any distinct sequence with the same apparent values.

However, given that **R** is reflexive, Dominance implies Justification Agglomeration. Every case can see the corresponding global good case with the same apparent values. Suppose that an agent is justified in believing p and justified in believing q at s. Then p and q are both known at some world with the same appearances. Accordingly, by Dominance, both are known at the corresponding global good case. Since knowledge agglomerates, $p \wedge q$ is also known at the global good case. Yet s can see the global good case. So an agent at s is justified in believing $p \wedge q$.

For example, consider Figure 2. s can access s' and s'', where the temperature is known to fall in the red and yellow regions, respectively. So at s the agent is justified in believing the temperature is within [72.5,77.5] in room 1, and justified in believing the temperature is within [72.5,77.5] in room 2. Yet, s can also access the global good case, s_{gc}. At s_{gc}, the agent knows the temperature is within both the red and the yellow regions. So, at s, is justified in believing the temperatures at locations 1 and 2 are both within [72.5,77.5].

Since we reject Justification Agglomeration, we must reject Dominance. But

Dominance is implied by Intersectivism. So we must reject Intersectivism.[13]

3.2 Internalism

Intersectivism satisfies a pair of characteristically internalist features. First, what an agent is justified in believing supervenes on how things appear to her.

Supervenience If, for all i: $a_{s_i} = a_{s'_i}$, then $s \models Jp$ iff $s' \models Jp$.

Supervenience says that what you are justified in believing is a function of how things appear to you. All systems which are alike with respect to appearances are alike with respect to what beliefs are justified.

Supervenience holds because what an agent is justified in believing at a sequence is what is known at the corresponding global good case where all of her appearances match reality. Supervenience is a precisification of the internalist intuition regarding the new evil demon problem (Lehrer and Cohen 1983, Cohen 1984). According to this intuition, what you are justified in believing in the global good case is the same as what you are justified in believing in any bad case in which things appear the same.[14]

Surprisingly, there is tension between Supervenience and The Preface. Justification Agglomeration follows from Supervenience given one other assumption: that where reality and appearance match, what an agent knows is exactly what she is justified in believing.

Optimism If, for all i: $r_{s_i} = a_{s_i}$, then $s \models Kp$ iff $s \models Jp$.

Optimism is the conjunction of two appealing seeming principles. First, that when one is subject to no error, one does not have misleading justification. That is, p is never false in a case where one is subject to no error and justified in believing p. Second, that when one is subject to no error, one is not Gettierized. That is, as long as appearance coincides with reality, justified true belief coincides with knowledge.[15] Optimism is validated by Intersectivism, since

[13] We have focused on cases where each quantity is measured on a single scale (in our case, temperature) and has the same margin for error. But shouldn't we also be interested in systems which measure quantities of different kinds or which are subject to different margins? For example, we might want to model a system comprising instruments for measuring temperature, humidity and light levels within a single room, where each instruments differs in its reliability. To do this, we could simple map each quantity to a common scale, renormalised relative to a (arbitrary) margin constant. While this presents no particular issue in theory, to avoid the added complexity, we will continue to focus on systems measuring quantities associated with a single scale and assume they have the same margin.

[14] See Bird 2007, p. 10 and Ichikawa 2014 for solutions to the new evil demon problem in Reduction-friendly frameworks.

[15] In some Gettier cases in the literature, appearance and reality coincide. However, these examples are not characterized by appearance/reality models. Gettier cases in these models

at every sequence what you are justified in believing is what you know at the corresponding global good case.

Interestingly, Supervenience and Optimism imply Justification Agglomeration, given the assumption that knowledge agglomerates.[16] This points to a general problem for modeling The Preface in an internalist account of justification. Any model of The Preface which rejects Justification Agglomeration must pick between Supervenience and Optimism.

There is a second respect in which the basic model is internalist. It predicts that the agent is always justified in believing that the world is not very different from how it appears to be. Where $[m,n]_i$ says the ith quantity's real value is between m and n ($\{s : n \leq r_{s_i} \leq m\}$):

Entitlement $\qquad \exists k \leq m : s \models J[a_{s_i} - k, a_{s_i} + k]_i.$

The agent is justified in believing that the real value of a quantity falls within a fixed range of its apparent value, at least as small as the constant, m. Increases in error limit what an agent can know, but do not affect an agent's entitlement to justifiedly believe that things are like they appear.[17]

Supervenience and Entitlement are internalist principles. The first says the way things appear to be to an agent determines what she is justified in believing. The second says she is always justified in believing things are close to how they appear to be. Our model will embrace Entitlement and a restricted form of Supervenience. But it will reject Supervenience in general. For certain claims, whether an agent is justified in believing them will depend on more than her appearances.

3.3 Conservatism

Pointwise Accessibility says that what the agent knows about any individual quantity is no weaker than what she would know about it in isolation, if she hadn't received any information about other quantities. In addition to Pointwise Accessibility, we want our model to satisfy a stronger condition. What an agent knows about any individual quantity should be the same as what she would know about it in isolation.

are identified in Williamson 2013, and do require the agent to be subject to a non-zero level of error.

[16] Suppose the agent justifiedly believes p and justifiedly believes q at s. Then by Supervenience she justifiedly believes p and justifiedly believes q in the corresponding good case s_{gc}. But then by Optimism she knows p and knows q at s_{gc}; so since knowledge agglomerates, she knows $p \wedge q$ at s_{gc}; so by Optimism again she is justified in believing $p \wedge q$ at s_{gc}; so by Supervenience again she is justified in believing $p \wedge q$ at s.

[17] The strongest claim known about quantity i at a global good case is that the real value is within m of its apparent value. Yet the strongest claim an agent is justified in believing at a sequence is the strongest claim known at the corresponding global good case.

As before, $\mathsf{Real}(s_i) = \{r' \mid (r', a) \in R(s_i)\}$ is the set of real values of quantity i compatible with what is known at (r_{s_i}, a_{s_i}) according to R. Let $\mathsf{Real}_i(s) = \left\{r_{s'_i} : s' \in \mathbf{R}(s)\right\}$ be the set of real values of quantity i compatible with what is known at s according to $\mathbf{R}$. Conservatism is the principle that the possible real values of any given quantity at s are the same as the possible real values at (r_{s_i}, a_{s_i}) in the basic model:

Conservatism $\qquad \mathsf{Real}_i(s) = \mathsf{Real}(s_i)$.

Intersectivism implies Conservatism. However, Conservatism is weaker than Intersectivism. Conservatism allows that the agent may know interesting information about which real values of quantities are compossible.

One reason we accept Conservatism is that it guarantees that our enriched models coincide with the basic appearance/reality models in the limiting case involving a single quantity.[18] As long as sequences are one-dimensional, Conservatism requires that our models make the same predictions about what is known and justified as the models introduced in §2. Accordingly, given Conservatism, failures of justification agglomeration can arise only in models with two or more dimensions. In this way, Conservatism ensures that a model is properly characterized as extending the basic model, rather than merely revising it.

A second reason we accept Conservatism is that it guarantees several of properties of the basic model, even in multi-dimensional models. First, any conservative model validates Entitlement. Entitlement says that an agent is justified in believing that the real value of quantity i falls within k of its apparent value, for some $k \leq m$. Suppose $\mathbf{R}$ is conservative. It follows that, for any s, there is an accessible sequence, s', which is a good case with respect to the quantity i (i.e., $r_{s'_i} = a_{s'_i}$). Furthermore, it follows that at s', it is known that the value of i falls within m of a_{s_i}. Hence, at s, an agent is justified in believing that the value of i falls within m of a_{s_i}.

Second, Conservatism implies a restricted form of Supervenience. An agent's justified beliefs about any individual quantity are guaranteed to supervene on her appearances. On the other hand, Conservatism does not imply Supervenience unrestrictedly. An agent's justification regarding multiple quantities may not supervene on the appearances, since it may not be straightforwardly decomposable into her knowledge of individual quantities. This is as it should be: we saw above that Supervenience and Optimism imply Justification Agglomeration.

We want a theory that respects Pointwise Accessibility and Conservatism while allowing for violations of Justification Agglomeration. We develop such a theory in the next section, which we motivate in terms of failures of normality to aggregate.

[18]We are grateful to an anonymous referee for encouraging us to address the issue of one-dimensional models in our framework.

4 The Normality of Error

In The Preface, what an agent is justified in believing about the total level of error outstrips what she is justified in believing about the level of error from any individual source. To model this, we give up two key properties of Intersectivism. First, an agent may know more about the total level of error than she knows about the error from any individual source. Second, what an agent knows in each global good case need not be strictly stronger than what she knows at any other case with the same apparent values. Given Reduction, this generates failures of Justification Agglomeration.

We accept Conservatism. What an agent knows about a given quantity is independent of what she knows about any other quantity. Accordingly, given Distance, we accept that for any source of information, an agent cannot rule out that it is free from error. For all the agent knows about each quantity, she is in a good case with respect to it. But this doesn't settle an agent's collective knowledge of multiple quantities. An agent may know that some source of information is subject to error, even if she does not know which. Sometimes, she can know that she is not in the global good case—the case which is good with respect to every quantity. This is precisely the kind of knowledge which is ruled out by Intersectivism, and which leads to the failure of Justification Agglomeration.

Imagine a weather station comprised of 100 measurement instruments. For each instrument, error is abnormal. Normally, the thermometer reports the correct temperature; normally, the barometer reports the correct atmospheric pressure; normally, the hygrometer reports the correct level of humidity; and so on. However, it would be abnormal for all 100 instruments to be reporting an accurate value. Normally, either the thermometer or the barometer or the hygrometer (and so on) is reporting an incorrect measurement.

Cases of this kind have the structure of The Preface. An agent consulting the weather station is justified in believing the reports of each individual instrument. But she is also justified in believing that one of the instruments is reporting an incorrect value. Since she is not justified in believing a contradiction, what she is justified in believing fails to agglomerate.

Our model sets out to implement this idea. For any source of information, considered individually, it is abnormal for it to be subject to error. However, when an agent receives information from multiple distinct sources, it may nevertheless be normal for some source or other to be subject to error.

4.1 A model

To model these ideas precisely, we need to talk about the total level of error at a world. We calculate the amount of error at a world, $\mathsf{Error}(s)$, by summing the distance between reality and appearance for each quantity.[19]

$$\mathsf{Error}(s) = \sum_{1 \le i \le n} |r_{s_i} - a_{s_i}|$$

In our model, the total level of error which is normal is an independently specified constant in the model, $\mathsf{NormalError}$. Where $\mathsf{NormalError}>0$, it is abnormal for a sequence to be entirely free from error.[20]

To incorporate normal error into our theory of knowledge, we propose that an agent knows that the total level of error is not significantly less normal than it actually is. We introduce a constant, e, that constrains how much an individual knows about the level of total error. We then say that if s' is $\mathbf{R}$-accessible from s, then the total error at s' is not significantly less normal than the total error at s. Call this Error Margin:

Error Margin $s' \in \mathbf{R}(s)$ only if $|\mathsf{NormalError} - \mathsf{Error}(s')| \le |\mathsf{NormalError} - \mathsf{Error}(s)| + e$.

We let $\mathbf{R}$ be the weakest relation satisfying Pointwise Accessibility and Error Margin. This relation can be equivalently defined as follows:

$s' \in \mathbf{R}(s)$ iff:

a. $\forall j : a_{s_j} = a_{s'_j}$.
b. $\forall j : |r_{s'_j} - a_{s'_j}| \le |r_{s_j} - a_{s_j}| + m$.
c. $|\mathsf{Error}(s') - \mathsf{NormalError}| \le |\mathsf{Error}(s) - \mathsf{NormalError}| + e$.

On this proposal, any world s' that is compatible with what is known at s must satisfy three conditions, (a)-(c). (a) says that the apparent value of each quantity is luminous. (b) is a local constraint on how much is known about the value of each quantity. It says that we know that no quantity diverges from its apparent value by much more than it actually does. (c) is a global constraint on how much is known about the total level of error. It says that we know that the total level of error does not deviate from the normal level by much more than it actually does.

4.2 The Preface

To model The Preface, we distinguish our knowledge of our overall performance on a series of tasks from our knowledge of our performance on each task individually. An agent's knowledge about her aggregate performance can outstrip the aggregation of her knowledge of her individual performances. This will occur

[19] Here another natural option would appeal to squared distance. We are sympathetic to this alternative, but stick to the simpler definition above throughout.

[20] As a simplifying idealization we assume that this constant is a real number, but we could equally identify it with a real interval instead, to model the case where the total quantity of error is normal as long as it falls within some range.

whenever the normal error is sufficiently high, and the actual error is sufficiently close to the normal error. In this case, the agent cannot rule out being in the good case with respect to any quantity in particular, but can rule out being in the good case with respect to every quantity. Equally, the agent can rule out being in a particularly bad case with respect to every quantity, even if she cannot rule out being in a bad case that is especially bad for any particular quantity.

To see the point, consider a simple kind of preface case involving two quantities. As above, imagine an agent receives reports on the temperature in two different rooms, both of which are reported to be 75°F. Suppose that $\mathsf{NormalError} = 10$°F. That is, under maximally normal conditions, the total error will amount to 10°F. Finally, suppose that $e = 7.5$°F$= 3m$. That is, the global margin for error is three times the local margin for error, which is 2.5°F.

Consider Figure 3. We can start by observing what is known at $s = \langle (70, 75), (70, 75)\rangle$, where the temperature is 70°F in both room 1 and room 2. First, Pointwise Accessibility entails that the temperature in each room is known to fall within the interval [67.5, 82.5] (i.e., within $5{+}m$ of its apparent value). Second, $\langle (70, 75), (70, 75)\rangle$ is a world at which the level of error is normal. $\mathsf{Error}(s) = 2(|70 - 75|)$, which coincides the normal quantity of error, $\mathsf{NormalError} = 10$. So, given Error Margin, the total error is known to fall within the interval [2.5, 17.5]. Thus, at $\langle (70, 75), (70, 75)\rangle$, the real value at the two locations is known to be within the blue region.

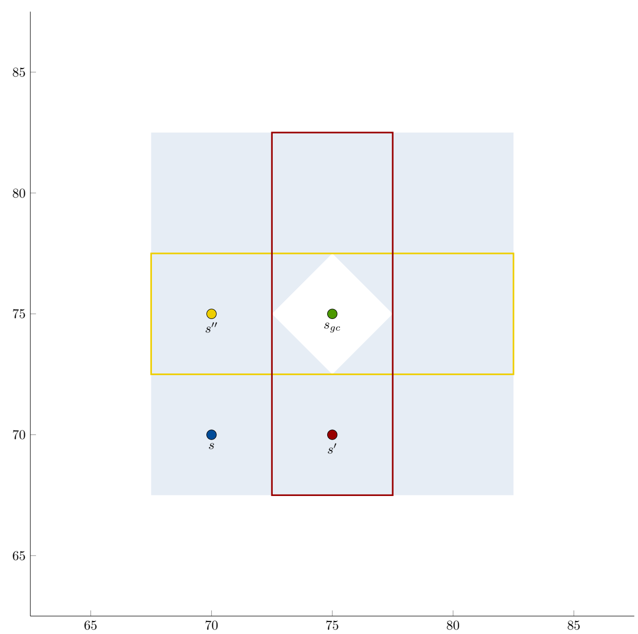

Figure 3: $a_1 = a_2 = 75, \mathsf{NormalError} = 10,\ e = 3m = 7.5$.

At $\langle(70,75),(70,75)\rangle$ the agent knows she is not in the good case with respect to the temperature in both rooms—$\langle(75,75),(75,75)\rangle$ is inaccessible. However, for all she knows, she is in a good case with respect to room 1 (since she can see $s' = \langle(75,75),(70,75)\rangle$), and for all she knows she is in a good case with respect to room 2 (since she can see $s'' = \langle(70,75),(75,75)\rangle$). At the former, the temperature at room 1 is known to be within [72.5,77.5], since the accessible worlds are the points within the red region. At the latter, the temperature in room 2 is known to be within [72.5,77.5], since the accessible worlds are the points within the yellow region. So, at $\langle(70,75),(70,75)\rangle$ for all she knows, she knows the temperature in room 1 is within [72.5,77.5] and for all she knows, she knows that the temperature in room 2 is within [72.5,77.5]. But she knows that she does not know that the temperatures in room 1 and room 2 are both within [72.5,77.5], since this would require her to be in the global good case —the world which is a good case with respect to the temperature at both locations. Accordingly, given Reduction, the agent is justified in believing, of each room, that the temperature is within [72.5,77.5] in that room (as Conservatism requires), but she is not justified in believing that the temperature is within [72.5,77.5] in both rooms. Justification Agglomeration fails.[21]

Crucially, Dominance also fails in our models. You don't always know most when you are free from error. Sometimes, being subject to error can put you in a position to know claims which you could not know at a global good case. As we saw in §3.1, giving up Dominance is a necessary condition on Justification Agglomeration failure. Note that while justification fails to agglomerate and justification is defined in terms of knowledge, knowledge does agglomerate. Failures of agglomeration in our models are confined in a desirable way.

The logic of justification generated by our models is not a normal modal logic. The K axiom fails. While we accept $J(p \supset (q \supset (p \wedge q)))$, we do not accept $Jp \supset (Jq \supset J(p \wedge q))$. As a result, it is not possible to define what an agent is justified in believing in terms of an accessibility relation, as it was in the basic model.

4.3 Structural Properties

In our model, what is known is constrained by multiple parameters: the level of error which is normal in the system ($\mathsf{NormalError}$) and the constants which specify the inexactness of the agent's knowledge of individual and total error (m and e). One advantage of this parametric approach is that we can now characterize various of the principles above within our model in terms of constraints on these

[21] Justification Agglomeration holds in the basic model because epistemic accessibility has a sphere structure, nested around the good case. By contrast, our model rejects Agg (that $\forall w \forall v \forall u(v, u \in R(w) \supset \exists z \in R(w) : R(z) \subseteq R(v) \cap R(u))$) in the above case. s can see s' and s''. The only world which sees a subset of the points seen by both s and s'' is the global good case. This is because the global good case is the only case at which the temperature in both locations is known to fall within [70, 80]. Yet s cannot see the global good case. So s cannot see a point which sees only points seen by both s' and s''.

parameters. In doing so, we'll see that our models impose interesting connections between the principles.

Above, we proposed that our model should involve a minimal departure from basic appearance/reality models in its predictions about knowledge of individual quantities. We gave this idea precise formulation with the principle of Conservatism, which requires that for every quantity i, we have $\mathsf{Real}_i(s) = \mathsf{Real}(r_{s_i}, a_{s_i})$: what is known about i at s according to an enriched model coincides with what is known about i at (r_{s_i}, a_{s_i}) according to the basic model with the same individual margin.

In our model, Conservatism holds just in case the following two conditions are satisfied, where n is the dimensions of the model (i.e., the number of quantities represented by a world s).[22]

$$\begin{array}{ll} \text{i.} & \mathrm{e} \geq m \\ \text{ii.} & \mathsf{NormalError} \leq (m \times (n-1)) + e \end{array}$$

We can characterize these constraints in informal terms. Recall again that the margins m and e measure the exactness of knowledge of individual and total error. Our first constraint then says that our knowledge of the total error cannot be more exact than our knowledge of the error regarding any individual quantity. The second constraint says that how much total error is normal imposes a lower bound on how inexact our knowledge is. As the level of normal error increases, the margins for individual and total error must increase accordingly.

With a characterization of Conservatism in place, we can consider its consequences. In our models, Conservatism implies that justification is locally consistent: no agent is both justified in believing p and justified in believing $\neg p$. As we saw above, given Reduction, local consistency is equivalent to the Geach rule. It holds iff accessibility is convergent: any two worlds which are both accessible from a third can see a common world. In our models, this will hold iff the following constraint is satisfied:

$$\mathsf{NormalError} \leq (m \times n) + e$$

This constraint trivially follows from Conservatism, since the latter requires that $\mathsf{NormalError} \leq (m \times (n-1)) + e$.

Our models turn out to impose further conditions on the epistemology of error. Consider the principle that justfiication is globally consistent: no agent is justified in believing each member of a set of inconsistent propositions. Global consistency holds iff accessibility is n-wise convergent: any n worlds which are all accessible for some world can see a common world. In our models, the local consistency of

[22] Proofs of observations in this section are confined to a supplementary appendix, available at https://www.dropbox.com/s/xe3aqh6lom4gan7/error_appendix_8.29.20.pdf?dl=0.

justification entails the global consistency of justification. This is because in our models, convergence implies n-convergence.[23]

How should we think about this result? In the scenarios represented by our models, global consistency has some plausibility. An agent who is justified in believing, for each of 10 rooms, that the temperature is within 2.5°F of its apparent value appears may nevertheless fail to be justified in believing that the total difference from the apparent temperature across all 10 combined is no greater than 25°F. That is, justification agglomeration may fail. Yet it is less obvious that she can be justified in believing that the total difference from the apparent temperature across all 10 combined is strictly greater than 25 °F. It is far from clear that global consistency may fail.

In this respect, the kinds of scenarios represented by our models contrast with the familiar preface scenario introduced in §1. Alex, the historian, is justified in believing each claim in the body of her book. But she is also justified in believing the claim in the preface, that some claim in the body is false. So the claims she is justified in believing are globally inconsistent.

Our initial preface case is in some ways underspecified. It is crucial to The Preface that Alex's sources of information are fallible. Fallibility engender inexactness. Where a source of information is fallible, the knowledge it produces is subject to margin for error constraints. Yet, as it is normally presented, The Preface case does not immediately reveal the inexactness of Alex's knowledge.

Our own scenarios represent one way of filling in details about the case. Most importantly, in our models, error is quantifiable. At each world, the level of error regarding any individual quantity can be given a precise value. This allows us to explicitly represent in quantitative terms both the inexactness of the agent's knowledge and what she knows about the error she is subject to. It is in virtue of this increase in expressive resources that, we think, justification should be globally consistent in our models.

In our original presentation of The Preface, for Alex to believe she is subject to error just is for her to believe that some claim in the body of her book is false. In contrast, the agents in our models have greater flexibility in how they can represent their predicament. In addition to whether they take themselves

[23] Any world which is accessible from both the global good case and from a world with a normal level of error (where $\mathsf{Error}(s) = \mathsf{NormalError}$) is accessible from every world. Yet for accessibility to be convergent, each world with a normal level of error must be able to see some world which can also be seen from the global good case. Otherwise, at any world which could see both the global good case and that world with a normal level of error, an agent would be justified in believing two pairwise inconsistent propositions. As a result, as long as justification is convergent, we know that there are some worlds in the model which are accessible from every world. Accordingly, accessibility will also be n-wise convergent. Any set of n worlds will be able to see a world in common.

For illustration, consider Figure 3. Local consistency requires that the blue region has a non-empty overlap with the intersection of the red and yellow regions (i.e., the region accessible from s_{gc}). Yet the overlap of these three regions will be accessible from any point. So global consistency is also satisfied.

to be subject to error, they need to decide how much error they are subject to, both at the level of individual quantities and overall. This increased capacity for detail is accompanied by pressure to ensure global consistency; instead of merely believing that one of her beliefs are false, an agent can instead believe that she is subject to some quantifiable amount of error. Since an agent can believe she is subject to some error without lapsing into inconsistency, she should do.

This leaves open the question of how our models ought to be extended to less richly structured scenarios, like The Preface. While this falls beyond the scope of the present work, see Carter and Goldstein 2020 for an exploration of similar ideas within a non-quantitative setting.

Finally, let's consider the status of the internalist principles from above in our model. We already saw above that Conservatism implies Entitlement. So, any models satisfying the above constraints will also validate Entitlement. For each quantity, the agent is justified in believing that the quantity's real value is similar to how it appears. Similarly, we saw above that Conservatism implies a weakened form of Supervenience, which says that an agent's justified belief about the real value of any individual quantity supervenes on the appearances.

On the other hand, Conservatism alone neither implies nor rules out Supervenience. As we saw above, however, Supervenience is incompatible with Optimism, given Justification Agglomemeration failure.

Optimism says that an agent knows p in the global good case just in case she is justified in believing p in the global good case. It turns out that if Justification Agglomeration is invalid and Optimism is valid, then the following condition must hold:

$$\mathsf{NormalError} > m \times n$$

Holding fixed that $\mathsf{NormalError} > m \times n$, we can then characterize Optimism in terms of a constraint on e, m, and n. If $\mathsf{NormalError} > m \times n$, Optimism holds just in case:

$$e \geq m \times n$$

As long as these conditions are satisfied, an agent knows less at every world accessible from the global good case than she knows at the global good case itself. This ensures that at the global good case she is justified in believing all and only those propositions which are known at the global good case (thereby satisfying Optimism).

The conjunction of Conservatism and Optimism creates an elegant chain of dependencies between our parameters: $m \leq e \leq m \times n < \mathsf{NormalError} \leq (m \times (n-1)) + e$. If we accept Optimism and reject Justification Agglomeration, we must accept failures of Supervenience. However, in our models, Optimism turns out to imply that Supervenience fails in two particularly notable ways.

First, as one might have expected from consideration of the qualitative preface, whether one is justified in conjoining justified claims about individual quantities does not supervene on one's appearances. At a world with a normal level of error, one is justified in believing, of each individual quantity, its real value is within m of its apparent value, but one is not justified in believing that the real value of every individual quantity is within m of the apparent value. In contrast, in the global good case, one knows, of each individual quantity, its real value is within m of its apparent value. So, since knowledge agglommerates and entails justification one is justified in believing every individual quantity is within m of the apparent value. Accordingly, justification to believe the latter claim varies across worlds with the same appearances.

But Optimism also implies a second kind of failure of supervenience in our models, one which is potentially less desirable. Whether an agent is justified in believing that she is subject to a non-zero level of error will vary across worlds with the same appearances. Since Justification Agglommeration fails, at any world with a normal level of error, the agent is justified in believing that error is greater than 0. By Optimism, however, at the global good case the agent is not justified in believing that total error is greater than 0 (since she does not know this). So her justification for believing that she is subject to at least some error does not supervene on her appearances. Yet, we might have thought that in The Preface, an agent is always justified in believing she is subject to some error.

It turns out that we can remain neutral on whether to accept Optimism or whether, instead, an agent is always justified in believing that she is subject to some error. This is because, when combined with failures of Justification Agglommeration, both principles imply that Supervenience will fail in general. [24]

In one respect, the failure of Supervenience is unsurprising. In our models, ignorance of error has a different epistemology than first order ignorance. By Conservatism, knowledge about the real value of any individual quantity is centred on its apparent value. The strongest proposition about a quantity which can be known is that it falls within m of the value it appears to have. In contrast, knowledge about total error need not be centered either on the normal level of error or on 0 total error. There is no unique strongest proposition about total error which can be known. Correspondingly, the strongest propositions about total error which for all one knows one knows can vary across worlds. It is this structural feature of our model which ensures supervenience holds for claims about individual quantities, but permits failure for claims about total error.[25]

[24] To see why, suppose that the proposition that error is greater than 0 is justified at every world. Then, by Reduction, there must be at least some world at which error is known to be greater than 0. This will be a world which cannot see the global good case (since the total error at the global good case is 0). Since Justification Agglommeration fails, there is something known at the global good case which is not known at any other world (that is, there is no s such that $\mathbf{R}(s) \subseteq \mathbf{R}(s_{gc})$). What is known at the global good case is justified at the global good case, but not at any world which cannot see the global good case. So Supervenience fails.

[25] From a logical point of view, Rosenkranz 2018 comes closest to our model. Like Williamson

This concludes our discussion of structural properties of our models.

4.4 Normality, Reconsidered

Recall that the basic model in §2 could be understood in terms of normality: at any world w, the agent knows p just in case p is true at any world v that is sufficiently normal at w. Before concluding, we will show how to reinterpret our class of models directly in terms of normality.

Our own models relied on the distinction between global and local normality. In addition to whether a world is normal with respect to any given quantity, we have crucially appealed to the normality of a world with respect to all quantities at once. A world where each quantity's real value matches its apparent value is very normal with respect to each individual quantity. But it may still be abnormal in another sense, because it has an abnormally low amount of error.

We can represent these ideas more precisely by directly defining two different kinds of normality orderings over worlds. Our first family of orderings compare the normality of two worlds both with respect to individual sources of information.

i. $s \leq_i s'$ iff $|r_{s_i} - a_{s_i}| \geq |r_{s'_i} - a_{s'_i}|$ and $a_{s_i} = a_{s'_i}$.
ii. $s \ll_i s'$ iff $|r_{s_i} - a_{s_i}| \geq |r_{s'_i} - a_{s'_i}| + m$ and $a_{s_i} = a_{s'_i}$.

These orderings are simple extensions of the basic model's approach to normality. s' is no less normal than s with respect to quantity i just in case $(r_{s'_i}, a_{s'_i})$ is no

2013, this theory rejects Knowledge Luminosity while retaining Reduction and Consistency. In addition, Rosenkranz 2018 departs further from Williamson 2013 by denying Justification Agglomeration.

One significant feature of Rosenkranz 2018 is the decision to define justification in terms of being in a position to know, rather than knowledge itself. In addition, Rosenkranz 2018 follows Heylen 2016 (and Rosenkranz 2016) in adopting a non-normal logic of being in a position to know, so that being in a position to know does not agglomerate. Rosenkranz 2018 suggests that an agent can be in a position to know p and be in a position to know $\neg Kp$ without being in a position to know the Fitch conjunction $p \wedge \neg Kp$.

For example, imagine that the agent is alone in a room with an even number of books on the wall. They could easily count the number of books, but have not. In this case, Heylen 2016 suggests that the agent is in a position to know that the number of books is even. But since they have not counted yet and know this, they know and hence are in a position to know that they don't know that the number of books is even.

Whether this argument is compelling depends on exactly how we understand the notion of being in a position to know. Schaffer 2007 embraces the normality of being in a position to know because he understands being in a position to know in terms of satisfying the evidential component of knowledge. Notice that in the case above, the agent does not actually possess evidence regarding the number of books, though they easily could have. More precisely, Schaffer 2007 follows Williamson 2000 and Hawthorne 2004 in understanding the notion of being position to know in terms of 'only needing a belief-based-on-competent-deduction-while-retaining-knowledge-of-the-premises in order to know' (249). In the case above, the agent needs to perform more than just a competent deduction in order to know that the number of books is even. The agent must also gather new evidence.

In this paper, we are perfectly happy to interpret justification in terms of being in a position to know, rather than knowledge. But our own investigation assumes that knowledge or being in a position to know is a normal modal operator.

less normal than (r_{s_i}, a_{s_i}) in the basic model. s is significantly less normal than s' with respect to quantity i just in case (r_{s_i}, a_{s_i}) is significantly less normal than $(r_{s'_i}, a_{s'_i})$ in the basic model. $s \not\ll_i s'$ iff it is not the case that $s \ll_i s'$, and either $s \leq_i s'$ or $s' \leq_i s$.

Crucially, we supplement these individual normality orderings with a second normality ordering that focuses on the total error at a world. s' is no less normal than s with respect to total error just in case the difference between $\mathsf{Error}(s')$ and $\mathsf{NormalError}$ is no greater than the distance between $\mathsf{Error}(s)$ and $\mathsf{NormalError}$. s is significantly less normal than s with respect to total error just in case the difference between $\mathsf{Error}(s)$ and $\mathsf{NormalError}$ is e greater than the distance between $\mathsf{Error}(s')$ and $\mathsf{NormalError}$. $s \not\ll_{\mathsf{NE}} s'$ iff it is not the case that $s \ll_{\mathsf{NE}} s'$, and for all i, either $s \leq_i s'$ or $s' \leq_i s$.

i. $s \leq_{\mathsf{NE}} s'$ iff $|\mathsf{Error}(s) - \mathsf{NormalError}| \geq |\mathsf{Error}(s') - \mathsf{NormalError}|$.
ii. $s \ll_{\mathsf{NE}} s'$ iff $|\mathsf{Error}(s) - \mathsf{NormalError}| > |\mathsf{Error}(s') - \mathsf{NormalError}| + e$

$\leq_{\mathsf{NE}}$ and $\ll_{\mathsf{NE}}$ order sequences with respect to what is normal regarding the sequence as a whole.

With our family of normality orderings in place, we turn to the construction of epistemic accessibility. In the basic model, we defined epistemic accessibility at w in terms of the worlds that were not significantly less normal than w. But now we have multiple different conceptions of normality.

There is no natural way of defining a single ordering over sequences which preserves the structural properties we are interested in and satisfies Knowledge of Normality. The problem is that our global normality ordering can conflict with the local normality orderings. Consider a world s_{NE} where the total error is the normal error. Suppose s_{gc} is a global good case with the same appearances. For any individual normality ordering $>_i$, we have $s_{gc} >_i s_{\mathsf{NE}}$, but in the global normality ordering we have $s_{\mathsf{NE}} >_{\mathsf{NE}} s_{gc}$. Since our two kinds of normality orderings conflict, we cannot simply define a new ordering where $s < s'$ just in case s is less normal than s' on some ordering; the result would not be asymmetric. So our model rejects a definition of epistemic accessibility in terms of any single conception of normality.

Instead, we define accessibility in terms of multiple conceptions of normality at once.

$$\mathbf{R}(s) = \{s' : \forall i : s' \not\ll_i s\} \cap \{s' : s' \not\ll_{\mathsf{NE}} s\}$$

That is, at s an agent knows that things are not much less normal than they actually are with respect to each individual source of information, and she also know that things are not much less normal than they actually are with respect to all of the sources of information taken as a whole. Thus, our model presents a form of normality-based theory of knowledge, but one which gives up the assumption that there is a single normality ranking which is relevant determining to what we know.

5 Conclusion

Reduction connects what an agent is justified in believing to what an agent knows. As we hope to have shown in this paper, a connection of this kind can be leveraged to shed light on the structural properties of both states. In particular, given Reduction, an adequate model of The Preface requires denying Dominance. The good case with respect to the relationship between appearance and reality is not always the best case with respect to what an agent knows. While failures of Dominance are striking, we have suggested they can be explained on the basis of normal error. Sometimes, it is abnormal for a system to be error-free.

References

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

Alexander Bird. Justified judging. *Philosophy and Phenomenological Research*, 74:81–110, 2007.

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

Sam Carter and Simon Goldstein. The more you know. 2020.

Stewart Cohen. Justification and truth. *Philosophical Studies*, 46:279–295, 1984.

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.

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

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.

Daniel Greco. Iteration and fragmentation. *Philosophy and Phenomenological Research*, 88(1):656–673, 2014.

Joseph Halpern, Dov Samet, and Ella Segev. Defining knowledge in terms of belief: the modal logic perspective. *Review of Symbolic Logic*, 2(3), 2009.

John Hawthorne. *Knowledge and Lotteries*. Oxford University Press, 2004.

Jan Heylen. Being in a position to know and closure. *Thought*, 2016.

Jonathan Jenkins Ichikawa. Justification is potential knowledge. *Canadian Journal of Philosophy*, 44:184–206, 2014.

Henry Kyburg. *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middletown, CT, 1961.

Keith Lehrer and Stewart Cohen. Justification, truth, and coherence. *Synthese*, 55(2):191–207, 1983. doi: 10.1007/BF00485068.

Wolfgang Lenzen. Recent work in epistemic logic. *Acta Philosophical Fennica*, 30, 1978.

Jarrett Leplin. *A Theory of Epistemic Justification*. Springer, 2009.

Clayton Littlejohn. Lotteries, probabilities, and permissions. *Logos and Episteme*, 3:509–514, 2012.

Clayton Littlejohn and Julien Dutant. Justification, knowledge, and normality. *Philosophical Studies*, 177(6):1593–1609, 2020. doi: 10.1007/s11098-019-01276-2.

David Makinson. The paradox of the preface. *Analysis*, 25(6):205–207, 1965.

Dana Nelkin. The lottery paradox, knowledge, and rationality. *The Philosophical Review*, 109:373–409, 2000.

Duncan Pritchard. *Epistemic Luck*. Oxford University Press UK, 2005.

Duncan Pritchard. Anti-luck epistemology. *Synthese*, 158(3):277–297, 2007. doi: 10.1007/s11229-006-9039-7.

Sven Rosenkranz. Being in a position to know and closure: Reply to heylen. *Thought*, 5:68–72, 2016.

Sven Rosenkranz. The structure of justification. *Mind*, 127(506):309–338, April 2018.

Jonathan Schaffer. Closure, contrast, and answer. *Philosophical Studies*, 133: 233–255, 2007.

Martin Smith. What else justification could be. *Noûs*, 44(1):10–31, 2010. doi: 10.1111/j.1468-0068.2009.00729.x.

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

Ernie Sosa. How to defeat opposition to moore. *Philosophical Perspectives*, 13: 141–154, 1999.

Robert Stalnaker. On logics of knowledge and belief. *Philosophical Studies*, 128: 169–199, 2006.

Jonathan Sutton. *Without Justification*. Oxford University Press, 2007.

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

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

Timothy Williamson. Improbable knowing. In Trent Dougherty, editor, *Evidentialism and its Discontents*. Oxford University Press, 2011.

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

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