Normality

Abstract

The modality of normality distinguishes states of affairs which are normal from those which are abnormal. Existing work on the modality of normality assumes that it is a restriction of metaphysical modality. In this paper, we argue that this assumption is inappropriate and explore the consequences of abandoning it.

After preliminary discussion (§1), we introduce the dominant framework for reasoning about normality (§2) and argue that it ascribes implausibly strong structural properties to the modality. In its place, we propose a new framework, which avoids this commitment (§§3-5). This account has a number of interesting features, which we explore in both an informal and formal setting. If correct, it implies that the modality of normality occupies a distinctive place in the space of modalities. Before concluding, we consider some of the wider implications of our account (§6), focusing on the role normality has played in epistemic theorizing.

1 Introduction

Normality is commonplace. We routinely distinguish normal weather from hurricanes and droughts; normal prices from bargains and rip-offs; and normal behavior from eccentricity and oddity. Not only do we form judgments about normality with ease, but facts about what is normal serve as a reliable guide to navigating the world. Knowing the normal presentation of a disease can help a doctor to successfully diagnose a patient. Knowing the normal migration patterns of birds can help ornithologists to identify species. And knowing the normal level of rush hour traffic can help a commuter to arrive on time.

Our subject in this paper is the modality of normality. Items of a variety of types can be evaluated for normality. For example, we can readily compare the normality of individuals, kinds and properties. Gerald Ford was more normal, for a president, than Richard Nixon. Weasels are more normal, for mammals, than wombats. And being tall is more normal, for basketball players, than being short. The modality of normality, in contrast, has to do with properties of states of affairs.[1] There may be interesting connections between the properties of individuals, kinds and properties and the modality of normality. However, in what follows, our attention will be focused exclusively on facts about the latter.

English is not particularly well-equipped to report such facts. As observed by Loets (2022), in sentences of the form ⌜ Normally, x would be F⌝, 'normally'

[1] We will use 'state of affairs' to refer to what others sometimes use 'proposition' to refer to. We don't intend anything to hang on the choice of terminology; we will simply use it to refer to 0-arity properties.

does not function as sentential operator, predicating normality of states of affairs, but rather as an adverbial quantifier with the power to bind variables in its scope (Lewis (1975)).[2] In common with other adverbial quantifiers, the quantificational force and domain of the quantifier are highly context sensitive. While we do not wish to rule out that sentences involving 'normally' may sometimes express facts of the kind we are interested in, the relationship between the two is opaque and we will avoid such constructions for present purposes.

Sentences of the form $\ulcorner x$ being F would be normal$\urcorner$ and $\ulcorner x$ being F would be abnormal$\urcorner$ appear better suited in this regard.[3] The former expresses—roughly—that x being F is compatible with with things being normal. However, it does not entail that x not being F is incompatible with things being normal. In this respect, it appears to denote an existential modality, akin to that denoted by adverbs such as 'possible' and 'permissible'. The latter expresses—again, roughly—that x being F is incompatible with things being normal. In this respect, it appears to denote a negative universal modality, akin to that denoted by adverbs such as 'impossible' and 'impermissible'. In what follows we will restrict ourselves to constructions like these, along with closely related sentences of the form $\ulcorner$It would be normal [/abnormal] for x to be $F\urcorner$ in providing informal glosses of our observations. Notably, English does not appear to have a dedicated sentential operator denoting the corresponding positive universal modality akin to that denoted by 'necessary' or 'obligatory'. It is nevertheless possible to express such a property, either as the dual of the property expressed by 'normal' or the internal negation of the property expressed by 'abnormal'. While this adds an additional layer of complexity, it provides the most faithful way of articulating claims about our target modality.

Williamson (2016, 2017) distinguishes between objective modalities (including metaphysical, nomic and practical modalities) and non-objective modalities (including deontic, epistemic and teleological modalities).[4] Where metaphysical

---

[2] As evidence of this, Loets (2022) notes the contrast between (†.a-b):

(†) a. Normally, I would be at work.
b. ?? Normally, I would be at work at on 6th April, 2023.

On the assumption that 'normally' is an adverbial quantifier, this contrast is attributable to the general infelicity of vacuous quantification.

[3] Observe that, in comparison with (†.b), (‡.a-b) are each fine:

(‡) a. My being at work at 1pm on 6th April, 2023 would be abnormal.
b. My being at work at 1pm on 6th April, 2023 would be normal.

[4] 'Objective modality' is used in a range of closely related but distinct ways by different authors (Williamson (2016, 2017); Ismael (2017); Strohminger & Yli-Vakkuri (2019); Bacon & Zeng (forthcoming); cf Rosen (2006)). We will use it strictly for those modalities which are restrictions of metaphysical modality, in the sense of Roberts (2020) (sometimes also called 'relative' modalities (cf. Hale & Leech (2017))). Some authors restrict the use of the term 'modality' to refer to the objective modalities or, more generously, to only properties of states of affairs. We will employ it more broadly, but we intend nothing to hang on this. Those who prefer a narrower use should feel free to insert the word 'putative' before phrases such as

modality is identified with the broadest objective modality, we can understand each of the objective modalities as restrictions of metaphysical modality by different (and potentially contingent) conditions (Hale & Leech (2017); Strohminger & Yli-Vakkuri (2019); Roberts (2020)).[5] Non-objective modalities are then distinguished from the objective modalities by their failure to fit this model.

A central concern of this paper is where the modality of normality belongs in this picture. By considering its structural properties, we will aim to situate it within the space of modalities. Our conclusion will be that normality occupies an interesting location outside the objective modalities. While not an objective modality itself, it is more closely related to them than other, familiar non-objective modalities are.

The modality of normality has been put to work in a variety of ways and in a variety of areas: in epistemology, giving theories of justification (Goldman (1986); Leplin (2009); Smith (2010, 2013, 2016); Goodman (2013)) and knowledge (Greco (2014); Goodman & Salow (2018, 2021); Beddor & Pavese (2018)); in philosophy of language, giving theories natural representation (Stampe (1977); Dretske (1981, 1988); Millikan (1984, 1989); Stalnaker (1999)) and generics (Asher & Morreau (1995); Asher & Pelletier (1997, 2012); Eckardt (1999)); and in philosophy of science, giving theories of ceteris paribus laws (Pietroski & Rey (1995); Schurz (2001a,b, 2002); Spohn (2002); Smith (2007)) and biological function (Boorse (1977); Millikan (1984, 1989); Wachbroit (1994)). Getting clear on its place within the space of modalities is important if we want to assess the viability of these applications.

We start, in §2, by introducing the prevailing framework for theorizing about normality. This framework, we argue, generates implausibly strong commitments. Giving up these commitments requires denying that the modality of normality is an objective modality. In §3 we propose an alternative and show how it avoids the unwanted commitments. §4 supplements our informal proposal with a formal model for a first-order modal language, and investigates in more detail the predictions it makes. §5 considers objections to the proposal made in §§3-4. Finally, §6 consider the implications of our argument for the growing literature on normality and epistemology. §7 concludes.

2 The Standard Model

Where ϕ is a state of affairs, we'll let $\ulcorner \blacksquare\phi \urcorner$ express that $\neg\phi$ would be abnormal. That is, the operator denotes the property instantiated by a state of affairs iff

'epistemic modality' or 'deontic modality'. On our positive account, the modality of normality will not be a modality in the first, most restrictive sense, but will be a modality in the second, intermediate sense (i.e., a property of states of affairs).

[5] More carefully, (where $\Box$ is the unique broadest objective modality) O_i is an objective modality iff there is some condition R_i such that: (i) $\Box\exists!q : R_i(q)$ and (ii) $\Box(O_i p \equiv \exists q : R_i(q) \wedge \Box(q \rightarrow p))$.

it would not be normal for it to fail to obtain. Though details differ, there is substantial consensus over the basic properties of this operator.

According to the standard model of normality—which we will call the Standard Model—possible worlds can be ordered according to how normal they are (either absolutely, or relative to an index world). At all worlds, $\blacksquare\phi$ obtains iff ϕ obtains at all worlds for which no world is strictly more normal.[6]

The Standard Model is assumed throughout the majority of existing work on normality (Delgrande (1987); Boutilier (1994c,b); Asher & Morreau (1995); Asher & Pelletier (1997, 2012); Eckardt (1999); Smith (2007); Booth *et al.* (2012); Yalcin (2016); see Loets (2022), in particular, for a comprehensive overview of the Standard Model). As a result, while such theories differ in the conditions they impose on the relation of comparative normality, they share a common commitment to many of the structural properties of the modality it determines.

Importantly, $\blacksquare$ is an objective modality according to the Standard Model. Specifically, it is the result of restricting metaphysical modality by the (presumably contingent) condition of being maximally normal. As shown by Roberts (2020), the objective modalities, understood as restrictions of metaphysical modality, are all and only those with a logic characterized by some accessibility relation in a Kripke semantics. For instance, in the case of the modality of normality (according to the Standard Model), this would be the accessibility relation which maps each world to the set of worlds which are maximally normal by its lights.

A core property of objective modalities is that they are agglomerative (Williamson (2016)). If ϕ and ψ are necessary, in some objective modality, then so too is $\phi \wedge \psi$. Stated for normality:

Agglomeration $\quad \blacksquare\phi, \blacksquare\psi \models \blacksquare(\phi \wedge \psi)$

In our preferred locution for talking about normality, **Agglomeration** says that if it would be abnormal for x to be F and it would be abnormal for y to be G, then it would be abnormal for x to be F or y to be G. Or, equivalently, if it would be normal for x to be F or y to be G, then it would either be normal for x to be F or it would be normal for y to be G.

Agglomeration follows directly from the fact that the Standard Model is committed to understanding the modality of normality as a restriction of metaphysical modality. If ϕ and ψ obtain at every maximally normal world, then $\phi \wedge \psi$ will obtain at every maximally normal world.[7] Where **Agglomeration** has

[6] This gloss assumes that domain of worlds ordered by comparative normality has one or more maximal elements. This is assumption is helpful, at least as an idealization. Where it is abandoned, (as in, e.g., Boutilier (1994b,c)), the Standard Model must be amended so that $\blacksquare\phi$ obtains at a world iff at every world w_1 at least as normal, there is some world w_2 at least as normal as w_1, and ϕ obtains at all worlds at least as normal as w_2. For further discussion of how this bears on the question of whether normality is an objective modality, see footnote 7.

[7] **Agglomeration** continues to hold in the absence of the assumption that the ordering of worlds for comparative normality has a maximal element. However, its infinitary variant will

received explicit discussion (in particular, Boutilier (1994c, 112-113), Thompson (2008, 69-70), Smith (2010, 15-16), Smith (2016, §4), Smith (2016), Smith (2017) Smith (2018, 3859-3862)), it has been treated as a desirable consequence of the Standard Model. And it is perhaps understandable why. From the fact that it would be abnormal for Ana to come to the party and that it would be abnormal for Bob to come to the party, it is certainly somewhat tempting to infer that it would be abnormal for either Ana or Bob to come to the party.

Nevertheless, we think that **Agglomeration** fails. In fact, we think that the kinds of counter-instances which demonstrate its failure are easy to find. Consider a collection of 100 biased coins, each of which has a $\frac{19}{20}$ chance of landing heads on any given flip. For each of the biased coins, it would be abnormal for that coin to land tails on a given flip.[8] Now suppose that all 100 coins are flipped simultaneously. It would not be abnormal for at least one of the coins to land tails.[9] Indeed, something stronger appears true. For no coin to land tails would be abnormal. Afterall, the chance of this occurring is approximately $\frac{59}{10{,}000}$.

These two observations, when combined, provide us with an immediate counter-example to **Agglomeration**. For some coins to land heads just is for the first coin to land heads or the second coin to land heads or the third coin... and so on. But **Agglomeration** says that if it would not be abnormal for this disjunctive state to obtain, then, for some coin or other, it would not be abnormal that coin to land heads.

Some may be suspicious about cases of the kind above. One possible source of worry is that coin flips are chancy in some way that precludes attributions of normality.[10] We do not share this worry. Nevertheless, this sort of chanciness is inessential to the basic observation.

For a person born in the US in 2023, it would be abnormal for that person to die before reaching the age of 30. Yet, it would not be abnormal for someone born in the US in 2023 to die before reaching the age of 30.[11] Indeed, something stronger appears true. For no-one born in the US in 2023 to die in the next 30 years would be (highly) abnormal. Again, these observations, when combined, provide us with a counter-example to **Agglomeration**. For someone born in the US in 2023 to die before the age of 30 is just for the youngest person born in

---

fail.

Infinitary Agglomeration $\bigwedge\{\blacksquare\phi | \phi \in \Gamma\} \models \blacksquare(\bigwedge \Gamma)$

The question of whether a modality which satisfied **Agglomeration** but not **Infinitary Agglomeration** should be classified as an objective is vexed and beyond the scope of our paper (but see Bacon & Zeng (forthcoming, 9)). Observe that, if some modalities which fail to satisfy **Infinitary Agglomeration** are categorized as objective, the gloss on what counts as a restriction in §1 will need to be revised.

[8] If you don't like this, lower the chance of the coin landing tails as much as you want.

[9] If you don't like this, increase the number of coins as much as you want.

[10] See Smith (2010, 15) and (2017) for an expression of this kind of worry.

[11] For reference, the SSA puts the chance of reaching the age of 30 at $\sim$ 97% for men, and $\sim$99% for women (https://www.ssa.gov/oact/STATS/table4c6.html)

the US in 2023 to die before 30 or the second youngest or the third youngest... and so on.

The kinds of cases which motivate rejecting **Agglomeration** are widespread and easy to generate.[12] However, accommodating such cases requires us to make substantial revisions to the way we think about normality. **Agglomeration** is a direct consequence of identifying normality with the property of obtaining across all maximal worlds, for some proposed ordering of comparative normality. Denying **Agglomeration** amounts to denying that $\blacksquare$ has a normal modal logic.[13] Since any modality which can be characterized by an accessibility relation in a Kripke semantics has a normal modal logic, it therefore also amounts to denying that $\blacksquare$ is an objective modality.

Vindicating judgments about the kinds of cases which motivate **Agglomeration** failure will require some departure from the Standard Model. Those who wish to retain the Standard Model could try to explain our judgments about these cases without giving up **Agglomeration**. The most obvious way to do this would be by appealing to the context sensitivity of the language we use to talk about normality (something along these lines is proposed in recent work by Goodman & Salow (2021)).

It is generally assumed by proponents of the Standard Model that the property picked out by predicates like 'normal' and 'abnormal' can vary according to the context in which they are used (Loets (2022, §4.2)). Whether someone dying at 30 can be correctly can be correctly described using 'normal' will vary according to whether we are talking the 21st century or the entirety of human history. Whether its being 20°C in winter can be correctly described using 'abnormal' will vary according to whether we are talking Melbourne or the entirety of Australia.

For context sensitivity to explain why it appears to us that **Agglomeration** can fail, it is not sufficient to posit that different orderings over worlds are elicited in different contexts of utterance. We also need a mechanism which would explain why specific utterances are only ever evaluated relative to contexts with the kind of ordering which would yield the judgments we observe. The proponent of this strategy needs to provide some mechanism which would explain why, in any context in which someone utters ⌜It would be abnormal for the nth coin to land tails⌝, in no maximally normal worlds in the contextually relevant order does the nth coin land tails. Simultaneously, this mechanism needs to explain why, in any context in which someone utters 'It would be abnormal for every coin to land tails', all maximally normal worlds in the contextually relevant ordering

[12] Some have claimed that there are two notions of normality: a statistical notion and a qualitative notion (Wachbroit (1994); Schurz (2001b, 2002, 2004); Smith (2010, 2016); Loets (2022)). Irrespective of what one thinks of this proposal, it will not offer an explanation of the current cases which would allow one to preserve **Agglomeration** for one or either of the notions. There is just no sense of normality on which it would be normal for every healthy person born in the US in 2023 to survive the next 30 years.

[13] That is, a logic satisfying Necessitation, Modus Ponens and the K axiom.

contain at least one coin which lands tails.

As far as we can see, the only plausible way of doing this is to posit a connection between what ordering is contextually relevant and what objects are mentioned in an utterance. According to this kind of picture, by mentioning the nth coin, an ordering is made salient in which any worlds in which the nth coin lands tails is at least somewhat abnormal. Meanwhile, in quantifying over all the coins, an ordering must be made salient which counts as maximally normal at least some worlds in which some coin lands tails.

The problem is that the latter commitment leaves the account incapable of explaining why we judge that the sentence ‘For each coin, it would be abnormal for that coin to land heads’ is true. More generally, any account which holds that what ordering is contextually salient is determined on the basis of what is mentioned lacks the resources to explain divergences in our judgments about sentences which differ only with respect to the scope of a universal quantifier.[14]

Perhaps, faced with this obstacle, the proponent of the Standard Model will simply stipulate that the contextually relevant ordering varies according to whether the quantifier takes wide or narrow scope. It is unclear whether, at this point, the proponent of the Standard Model would still count as offering an *explanation* of our judgments. However, they face more significant problems. The conjunction ‘It would not be abnormal for some coin to land tails, but, for each coin, it would be abnormal for that coin to land tails’ appears a perfectly accurate way of describing the kinds of cases in which **Agglomeration** appears to fail. Yet absent shifts in context mid-utterance, the contextualist strategy has no way of explaining this.

§§3-4 offer a positive proposal about what makes a state of affairs abnormal which vindicates, rather than explaining away, apparent failures of **Agglomeration**. In doing so, it departs from the Standard Model in a number of ways.

3 A Theory of Normality

Our theory combines two core ideas. First, that normality is determined relative to one or more subject matters—sets of mutually exclusive states of affairs. What is normal is what obtains across the maximally normal states in the relevant subject matters. Second, that the comparative normality of states of affairs is proportional to their probability. We’ll introduce each of these ideas separately in §3.1 and §3.2.

States of affairs, as we will think about them, are the kinds of things which obtain (or fail to obtain) at worlds. Each state of affairs partially settles how

[14]The problem is not that such an account is committed to claiming that a speaker mentions everything she quantifies over. Rather, it is that, at least in the cases we are considering, what she mentions will not vary with the scope of the quantifier.

things are at the worlds at which it obtains and the same state of affairs can obtain at multiple different worlds. Understood this way, states of affairs can be ordered for strength. Where one state of affairs obtains at every world at which a second obtains, we will say that the second necessitates first. For example, a die landing 2 necessitates the die landing prime, but not *vice versa*. Correspondingly, we'll say that one state of affairs is at least as strong as another iff the first necessitates everything necessitated by the second. We will assume that for each world there is some maximally strong non-absurd state of affairs which obtains at that world and at no other.

3.1 Subject Matters

Imagine 100 fair coins are tossed simultaneously. Some outcomes would be abnormal, such as all the coins landing tails. No outcome in which $\frac{1}{2}$ of the coins landed heads would be abnormal, however. But suppose, now, that exactly 50% of the coins are painted red on the heads side, with the other 50% being painted red on the tails side. It's hard to resist accepting that it would be abnormal for all 100 coins to land with their red face up. Yet for all 100 coins to land red face up is for some outcome in which $\frac{1}{2}$ of the coins to land heads to obtain.

Similar judgments can be elicited across a wide variety of cases. For a commercial airline to cancel half a dozen flights in a six month period would not be abnormal. However, it would be pretty abnormal for them to cancel the next half a dozen flights *you* are scheduled to take. Or another: there is no card in a 52 card deck which it would be abnormal to draw. However, if you are first asked to name a card, it would be at least somewhat abnormal for someone to draw that card at random from the deck.

Here is one *prima facie* lesson of these cases: in determining whether ϕ is abnormal, it matters how the space of possibilities is divided. Depending on what subject matter is being considered, different ways of dividing up possibilities will be made relevant. We can think of a subject matter as corresponding to a set of mutually exclusive and jointly exhaustive states of affairs (Lewis (1988a,b); Yablo (2014); Yalcin (2011); Plebani & Spolaore (2021)). Intuitively, a state of affairs is included in a subject matter iff it fully settles how things are with respect to that subject matter and there is no strictly weaker state which also fully settles how things are with respect to that subject matter. Characterized this way, a subject matter determines a partition on the space of possible worlds. Two worlds are cell-mates iff some state in the subject matter obtains at both. Since the states are mutually exclusive (i.e., necessarily fail to co-obtain), cells will be disjoint. Since they are jointly exhaustive (i.e., necessarily one of the states obtains), each world will be a member of some cell. Our guiding idea is that, in evaluating ϕ for normality, we consider what is the case across the most normal states in subject matters which ϕ is associated with.[15] Here is a first pass (to be supplemented below):

[15] In recent work, Goodman & Salow (2021) also suggest relativizing normality to a contextually supplied partition. Crucially, the two proposals differ in whether the subject matters

Normality $\blacksquare\phi$ iff ϕ is necessitated by each of the maximally normal states of affairs in each subject matter ϕ is associated with.

To understand what **Normality** says, we obviously need to supplement it with an account of what it is for a state of affairs to be associated with a subject matter.

Say that a state of affairs ϕ is about a subject matter iff there is some subset of the subject matter such that, necessarily, ϕ obtains iff some member of that set obtains. Equivalently, ϕ is about a subject matter iff the set of worlds at which it obtains is the union of some set of cells of the partition the subject matter induces.

In order to be associated with a subject matter, a state of affairs must be about it. However, not every subject matter which a state of affairs is about will be associated with it. We propose a distinction between subject matters which are relevant and those which are not. In order to be associated with a state of affairs, a subject matter must be relevant. Being relevant (as we intend it) is not an intrinsic property of subject matters; rather different subject matters will be relevant at different contexts. Our idea is that, in a given context, normality talk expresses the property of being (ab)normal relative to the relevant subject matters in that context. As a result, across contexts the property picked out in normality ascriptions will co-vary with which subject matters are relevant.

Association π is associated with ϕ iff $\begin{cases} \text{(i)} & \phi \text{ is about } \pi\text{; and} \\ \text{(ii)} & \pi \text{ is relevant.} \end{cases}$

To see how **Normality** and **Association** can allow for failures of **Agglomeration**, it suffices to observe that the subject matters associated with a conjunction need not overlap with any of the subject matters associated with its conjuncts.

For a simple example, consider the four possible ways two coins could land. Different subject matters will divide up these possibilities in different ways. For instance, take the following three subject matters:

π_1: `How did coin 1 land?`

π_3: `How many coins landed heads?`

π_2: `How did coin 2 land?`

Where each shaded region corresponds to a different state of affairs, each partition in **Figure 1** corresponds to a different subject matter. Thus, π_1 and π_2 are the subject matters which distinguish between possibilities based on how the first and second coins land, respectively; in contrast, π_3 is the subject matter

considered in determining facts about normality vary according to the state of affairs being evaluated. As a result, while **Agglomeration** fails in our proposal, it remains valid in Goodman and Salow's.

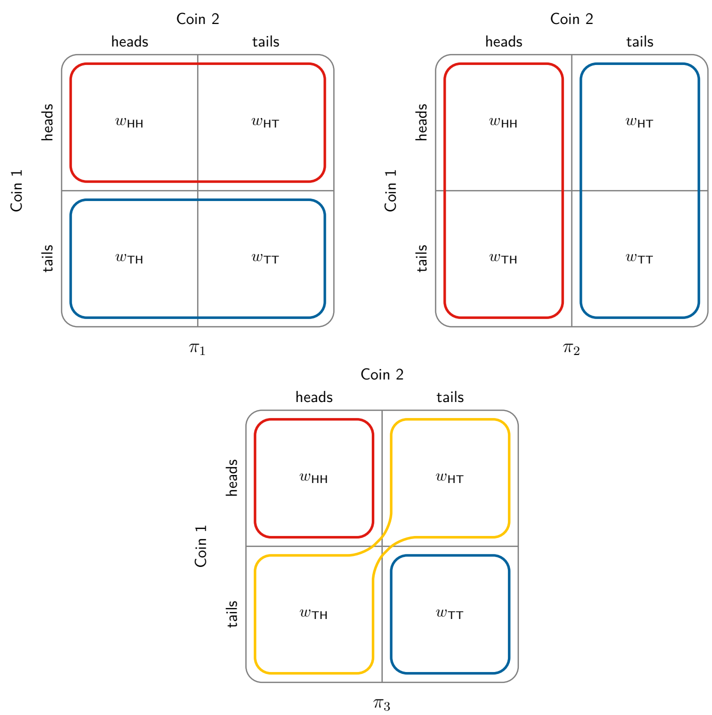

Figure 1: A four world model.

which distinguishes between possibilities based on what proportion of the coins land heads (but not on how any particular coin lands).

Suppose that all and only these three subject matters are relevant.Then the states of affairs of the first and second coins landing heads will be uniquely associated with π_1 and π_2, respectively. By **Normality**, then, it would be abnormal for the first coin not to land heads (i.e., to land tails) iff the state of affairs corresponding to the red region in π_1 is more normal than the state of affairs corresponding to the blue region. The same goes, *mutatis mutandis*, for the second coin.

The state of affairs of both coins landing heads, in contrast, will be uniquely associated with π_3. Accordingly, by **Normality**, it would be abnormal for both coins not to land heads (i.e., for some coin to land tails) iff the state of affairs corresponding to the red region in π_1 is more normal than the states of affairs corresponding to the yellow region and the blue region. Yet, for all we have said so far, the comparative normality of the states in π_1 and π_2 fails to impose any requirements on the comparative normality of states in π_3. Accordingly, absent further constraints on the relation of comparative normality, **Normality** and **Association** will be compatible with failures of **Agglomeration**.

These observations generalize to more complicated cases, such as our biased coins example. It is relatively natural to think that, within our setup, it will be relevant how each individual coins lands and also what proportion of coins land heads and land tails.

With this in mind, for $1 \leq n \leq 100$, let $\hat{\pi}_n$ be the subject matter comprising the state of affairs of the nth coin landing heads and the state of affairs of the nth coin landing tails. Intuitively, $\hat{\pi}_n$ is the subject matter `How did the nth coin land?`. Let $\hat{\pi}_{\%}$ be the subject matter comprising, for each $0 \leq k \leq 100$, the state of affairs of exactly $\frac{k}{100}$ coins landing heads. Intuitively, $\hat{\pi}_{\%}$ is the subject matter `How many coins landed heads (exactly)?`

Wherever $\hat{\pi}_1 - \hat{\pi}_{100}$ and $\hat{\pi}_{\%}$ exhaust the relevant subject matters, failures of agglomeration of the kind we discussed above will be possible. For $1 \leq n \leq 100$, the state of affairs of the nth coin landing heads will be associated with $\hat{\pi}_n$. In contrast, the state of affairs of all the coins landing heads will be associated with $\hat{\pi}_{\%}$. Yet nothing we have said so far requires that, if for any n, the nth coin landing heads would be more normal than the nth coin landing tails, then for any for $k \geq 1$, it would be more normal for all the coins to land heads than for exactly $\frac{k}{100}$ to land tails.

Predicting that **Agglomeration** can fail in the kinds of cases which motivated its rejection is not the same as predicting that it does fail in those cases. In order to do the latter, our proposal needs to be supplemented with a substantive account of what makes one state of affairs more normal than another. This is the focus of the next subsection.

3.2 Probability

Comparing the normality of entire worlds is complicated. Any order over worlds will inevitably need to settle trade-offs between different dimensions along which worlds can be abnormal. While such questions may not pose an insurmountable challenge for the proponent of the Standard Model, we would be happier not to have to adjudicate them.

Comparing the normality of states of affairs is, in many instances, comparatively simple. Two dice landing with a total of 12 is less normal than two dice landing with a total of 7; an American family having 2 children is more normal than their having $2 + n$ children (for all $n \geq 1$); the closer a person's temperature to 37°C, the more normal; and so on.[16]

We propose the following principle as a simple gloss on the comparative normality of states of affairs.[17]

[16]Crucially, where ϕ is more normal than ψ, it does not follow that, for any χ, $\phi \wedge \chi$ will be more normal than $\psi \wedge \chi$. Someone having septicemia and a temperature of 37°C may well be less normal than their having septicemia and a temperature of 42°C.

[17]See Goodman & Salow (2021) for a related proposal which also proposes to reduce comparative normality to comparative probability.

Probability ϕ is more normal than ψ iff ψ is much less probable than ϕ.

Probability says that the ordering of states for comparative normality is a coarsening of their ordering for comparative probability; for any ϕ, the set of states which are less normal than ϕ will be a subset of the states which are less probable than ϕ. Importantly, **Probability** is not intended as an account of the pre-theoretical notion of comparative normality (or of the locution $\ulcorner\phi$ would be more abnormal than $\psi\urcorner$). Rather, it is simply intended to provide a gloss on the kind of ordering over states of affairs which is invoked in **Normality**. It is ultimately unimportant to our account of the unary modality whether this ordering coincides with an intuitive notion of comparative normality.[18]

It is vague what it takes for one state of affairs to be much less probable than another. So, by **Probability**, it is vague what it takes for one state of affairs to be more normal than another. This is as we would expect. Since we are primarily interested in structural features of normality, vagueness regarding comparative normality will not present an obstacle to developing our account. We assume only that there will be some threshold $t \geq 1$ such ϕ is much less probable than ψ iff the probability of ψ divided by the probability of ϕ is (strictly) greater than t. We allow that the exact value of this threshold may be unknowable and dependent on the context in which we are situated. For any resolution of what it takes to be much less probable, however the structural properties of the resulting order over states of affairs will remain the same.

What kind of probability measure does **Probability** appeal to? Here, we want to remain as neutral as possible. There is, we claim, a pre-theoretical but respectable objective probability measure which is invoked in claims about the likelihood of a pair of dice landing a certain way, a family having a given number of children or your having a certain temperature. Whatever that measure is, it is the measure which is relevant to the comparative normality of states of affairs. (NB: The target measure would appear to be well-characterized by a propensity interpretation, according to which, e.g., the probability of a die landing a certain way is proportional to the degree to which it is disposed to be land that way (Popper (1957, 1959, 1990); Giere (1973); Fetzer (1982, 1983); Gillies (2000, 2016)). While we are sympathetic to the prospects of this kind of gloss on the relevant measure, **Probability** should not be taken to rest on the success of any particular version of propensity theory. Rather, it appeals directly to the judgments which such theories aim to capture.)

We are now in a position to see what predictions **Normality** and **Probability** make about our original case of **Agglomeration** failure when combined. In our case, the probability of any particular biased coin landing heads is 19 times the probability of its landing tails. Accordingly, assuming that the relevant

[18]Observe that **Normality** could be stated directly in terms of what is necessitated by the states of affairs in a subject matter for which no state of affairs in the subject matter is much more probable. We prefer the present formulation primarily because it allows us to separate the roles played by subject matter sensitivity and by the specific choice of an ordering over states of affairs.

threshold for being much less probable is lower than 19, the former will be more normal than the latter. In contrast, there is some $k > 0$ such that the probability of all 100 coins landing heads is lower than the probability of exactly k coins landing tails.Accordingly, on any way of fixing the relevant threshold, the latter will be at least as normal as the former. It follows that if the relevant subject matters are as described in the previous section, for any $1 \leq n \leq 100$, it would be abnormal for the nth coin to land tails. However, it would not be abnormal for some coin to land tails (since, in the associated subject matter, there are guaranteed to be maximally normal states which entail that not all the coins land heads).

The point generalizes to our other, qualitative examples. There is some $n \geq 30$ such that, for any $k \leq 30$, the probability that someone born in the US in 2023 will die at k years old is much lower than the probability that they will die at n years old. However, for every pattern of mortality over the entire population of people born in the US in 2023 in which none die before the age of 30, there is a much more probable pattern of mortality in which at least one person dies before the age of 30.

4 A Non-Standard Model of Normality

The previous section introduced the core components of our theory in an informal setting. As we saw, **Normality** and **Probability** are able to predict failures of **Agglomeration** in precisely the right cases.

In this section, we implement our theory more rigorously, by giving a class of models for a modal propositional language which implement the proposals of the previous section. Readers uninterested in the details of this implementation can skip the section. The most important observations will be that, on our theory, the modality of normality is (i) closed under necessary equivalence, modus ponens and necessitation, but (ii) not closed under single-premise closure.

> **Definition 1** (Language)**.** $\mathcal{L}$ is the smallest set containing the sentential atoms $\{\mathsf{A}, \mathsf{A}', ..., \mathsf{B}, \mathsf{B}', ...\}$ and which is closed under boolean connectives $(\neg, \wedge, \vee)$ and a unary modal operator $(\blacksquare)$.

A model for $\mathcal{L}$ is a tuple $\mathcal{M} = \langle \mathcal{W}, \Pi, Pr, t, [\![\cdot]\!] \rangle$. $\mathcal{W}_{\mathcal{M}}$ is a domain of worlds. We identify worlds with characteristic functions on the sentential atoms. $\Pi_{\mathcal{M}}$ is a non-empty finite set of partitions on $\mathcal{W}_{\mathcal{M}}$, such that $\{\{w\} | w \in \mathcal{W}_{\mathcal{M}}\} \in \Pi_{\mathcal{M}}$. Intuitively, we think of each $\pi \in \Pi_{\mathcal{M}}$ as representing a relevant subject matter. $Pr_{\mathcal{M}}$ is a function from worlds to probability measures over $\mathcal{P}(\mathcal{W}_{\mathcal{M}})$. $t_{\mathcal{M}} \geq 1$ is a constant, representing the magnitude by which the probability of two states of affairs must differ for one to be much more probable than the other. $[\![\cdot]\!]_{\mathcal{M}}$ is an interpretation function, which maps sentences in $\mathcal{L}$ onto $\mathcal{P}(\mathcal{W}_{\mathcal{M}})$. Where possible, we suppress indexation to a model.

We say that p is about π iff p is the union of some subset of π. We can then define a function, $|\cdot|$, which maps states to the set of relevant subject matters which they are about. Intuitively, we think of $|p|$ as representing the subject matters associated with p.

Definition 2 (Association)**.** $|p| = \{\pi \in \Pi : \exists X \subseteq \pi : p = \bigcup X\}$.

Finally, for each $w \in \mathcal{W}$ and $\pi \in \Pi$, we distinguish a privileged subset of π comprising the states which are sufficiently probable at w.

Definition 3 (Maximality)**.** $Max_w(\pi) = \{p : \nexists q \in \pi : \frac{Pr_w(q)}{Pr_w(p)} > t\}$.

$p \in Max_w(\pi)$ iff $p \in \pi$ and the most probable elements of π are no more than t times as likely as p at w. Given **Probability**, we can think of $Max_w(\pi)$ as the set of maximally normal states within π at w.

We are now in a position to introduce our semantics.

Definition 4 (Semantics)**.**

$$
\begin{array}{llcl}
\text{i.} & [\![\mathsf{A}]\!] & = & \{w : w(\mathsf{A}) = 1\} \\
\text{ii.} & [\![\neg\phi]\!] & = & \mathcal{W} - [\![\phi]\!] \\
\text{iii.} & [\![\phi \wedge \psi]\!] & = & [\![\phi]\!] \cap [\![\psi]\!] \\
\text{iv.} & [\![\phi \vee \psi]\!] & = & [\![\phi]\!] \cup [\![\psi]\!] \\
\text{v.} & [\![\blacksquare\phi]\!] & = & \{w : |[\![\phi]\!]| \neq \varnothing \wedge \forall \pi \in |[\![\phi]\!]| : \bigcup Max_w(\pi) \subseteq [\![\phi]\!]\}
\end{array}
$$

Each atom is mapped to a set of worlds. Negation, conjunction and disjunction behave in the usual boolean way. $\blacksquare\phi$ is true at w iff there is some subject matter associated with $[\![\phi]\!]$ and for every π associated with $[\![\phi]\!]$, ϕ is true throughout every maximally normal state in that π (relative to w). Observe that where, for some $w, v \in \mathcal{W}$ and $\pi \in \Pi$, $Max_w(\pi) \neq Max_v(\pi)$, we allow that $\varnothing \subset [\![\blacksquare\phi]\!] \subset \mathcal{W}$. This reflects the idea that what is abnormal is a contingent matter.[19]

We define entailment in the expected way.

Definition 5 (Entailment)**.**

i. $\Gamma \models_{\mathcal{M}} \phi$ iff $(\bigcap_{\psi_i \in \Gamma} [\![\psi_i]\!]_{\mathcal{M}}) \subseteq [\![\phi]\!]_{\mathcal{M}}$.
ii. $\Gamma \models \phi$ iff for all $\mathcal{M}$: $\Gamma \models_{\mathcal{M}} \phi$.

Where $\Gamma \models_{\mathcal{M}} \phi$, we say that the inference is valid-in-$\mathcal{M}$. Where, $\Gamma \models \phi$, we will say that the inference is valid (and, more loosely, we will say the same of a meta-rule, where the validity in all models of one inference implies the validity in all models of a second).

Fact 1. **Agglomeration** is invalid.

[19] Since, for any $\pi \in \Pi$ we impose no constraints between Pr_w and Pr_v, we also predict $\blacksquare$ will not obery the 4 axiom. See Carter (2019) for arguments that this is a desirable outcome.

As a counter-model, consider the four-world domain and set of partitions depicted in **Figure 1**. Assume that $[\![\mathsf{A}]\!] = \{w_{\mathsf{HH}}, w_{\mathsf{HT}}\}$ and $[\![\mathsf{B}]\!] = \{w_{\mathsf{HH}}, w_{\mathsf{TH}}\}$. Suppose further that, for all w in the domain: $Pr_w([\![\mathsf{A}]\!]) = Pr_w([\![\mathsf{B}]\!]) = \frac{3}{5}$ and that $[\![A]\!]$ and $[\![B]\!]$ are probabilistically independent. Accordingly, we have that for all w in the domain: (a) $Pr_w(w_{\mathsf{HH}}) = \frac{9}{25}$; (b) $Pr_w(w_{\mathsf{HT}}) = Pr_w(w_{\mathsf{TH}}) = \frac{6}{25}$; and (c) $Pr_w(w_{\mathsf{TT}}) = \frac{4}{25}$. Finally, let $1 \leq t < 1\frac{1}{4}$.

It follows that, for all w, $Max_w(\pi_1) = [\![\mathsf{A}]\!]$ and $Max_w(\pi_2) = [\![\mathsf{B}]\!]$. So $\blacksquare\mathsf{A}$ and $\blacksquare\mathsf{B}$ are true throughout the model. Yet, in contrast, $Max_w(\pi_3) = \{w_{\mathsf{HT}}, w_{\mathsf{TH}}\}$. So $\blacksquare(\mathsf{A} \wedge \mathsf{B})$ is false throughout the model. Hence **Agglomeration** is not valid in all models.

Since **Agglomeration** fails, the logic of $\mathcal{L}$ generated by our models is not a normal modal logic. Nevertheless, our logic retains a number of properties common to the objective modalities.

Fact 2. For all $\mathcal{M}$: if $\models_{\mathcal{M}} \phi \leftrightarrow \psi$, then $\models_{\mathcal{M}} \blacksquare\phi \leftrightarrow \blacksquare\psi$.

Fact 2 reflects the fact that $\blacksquare$ is not a hyper-intensional operator in our models. If $\phi \leftrightarrow \psi$ is valid in a model, then $\blacksquare\phi \leftrightarrow \blacksquare\psi$ will be valid in the model as well.

Modus Ponens	$\phi, \phi \rightarrow \psi \models \psi$
Necessitation	If $\models \phi$, then $\models \blacksquare\phi$.

Similarly, our models preserve **Modus Ponens** and **Necessitation** as valid inference and meta-inference rules, respectively.

Fact 3. **Modus Ponens** and **Necessitation** are valid.

These facts are reassuring. They demonstrate that our logic is not departing further than necessary from a normal modal logic. In fact, we can reformulate our semantics as an instance of a well-understood family of models for an interesting class of non-normal modal logics.

In neighborhood semantics (Segerberg (1971); Chellas (1980)), modality is characterized by a relation, $\mathcal{N}$, which associates a world, w, with a neighborhood $\mathcal{N}(w) \subseteq \mathcal{P}(\mathcal{W})$. Where $\mathcal{O}$ is a necessity modal, $w \in [\![\mathcal{O}\phi]\!]$ iff $[\![\phi]\!] \in \mathcal{N}(w)$. There are interesting relations between our models and neighborhood semantic models.

For any $\mathcal{M}$, let $\mathcal{N}_{\mathcal{M}}$ be a neighborhood relation defined such that $\mathcal{N}_{\mathcal{M}}(w) = \{p : |p| \neq \varnothing \wedge \forall \pi \in |p| : \bigcup Max_{w,\mathcal{M}}(\pi) \subseteq p\}$. Then we have the following correspondence between our semantics $\blacksquare$ and $\mathcal{N}_{\mathcal{M}}$:

Observation 1. $w \in [\![\blacksquare\phi]\!]_{\mathcal{M}}$ iff $[\![\phi]\!]_{\mathcal{M}} \in \mathcal{N}_{\mathcal{M}}(w)$.

Based on **Observation 1**, it is easy to define a map, $*$, from our models into the class of neighborhood semantic models such that, for all $\mathcal{M}$, $\Gamma \models_{\mathcal{M}} \phi$ iff

$\Gamma \mathrel{\substack{\models \\ \mathcal{M}^*}} \phi$.[20]

Under this map, various structural properties of the neighborhood relation will correspond different structural properties of the $\blacksquare$ modality. For example, in the corresponding neighborhood semantic models, it is not in general the case that if $p \in \mathcal{N}(w)$ and $p \subseteq q$, then $q \in \mathcal{N}(w)$. That is, $\mathcal{N}$ is not closed under the superset relation. **Fact 4**, which says that $\blacksquare$ is not closed under single-premise entailment within a model, follows as a corollary.

Fact 4. It is not the case that for all $\mathcal{M}$: if $\phi \models_{\mathcal{M}} \psi$, then $\blacksquare\phi \models_{\mathcal{M}} \blacksquare\psi$

We can also prove a number of facts about inference patterns valid across all our models in terms of corresponding features of $\mathcal{N}$. First, since Π is required to be finite, it follows that no world has a neighborhood containing the empty set; for any w, $\varnothing \notin \mathcal{N}(w)$.[21] This ensures that there is no world at which the tautology would be abnormal. Second, $\mathcal{N}$ is proper; for any w, if $p \in \mathcal{N}(w)$, then $(\mathcal{W} - p) \notin \mathcal{N}(w)$.[22] This ensures that there is no world at which both a proposition and its negation would be abnormal. On the other hand, it is not the case that, for any w: if $p, q \in \mathcal{N}(w)$, then $p \cup q \in \mathcal{N}(w)$.

Fact 5. (i)-(ii) are valid. (iii) is invalid.

(i)	$\models \neg\blacksquare\bot$	(**Coherence**)
(ii)	$\models \blacksquare\phi \rightarrow \neg\blacksquare\neg\phi$	(**D Axiom**)
(iii)	$\blacksquare\phi, \blacksquare\psi \models \blacksquare(\phi \vee \psi)$	(**Weakening**)

This concludes our discussion of the model.

5 Objections and Replies

Objection: If what subject matters are relevant varies according to context, how is the proposed approach any better than the contextualist variant of the standard model rejected in §2?

Reply: It is important to distinguish between what we might call 'heavyweight' and 'lightweight' versions of contextualism about normality talk. Heavyweight contextualism combines two ideas: first, that what property is expressed by normality talk can vary according to the context of utterance. And, second, that changes in context explain the apparent invalidity of certain inference rules

[20] Specifically, let $\mathcal{M}^* = \langle \mathcal{W}_{\mathcal{M}^*}, \mathcal{N}_{\mathcal{M}^*}, [\![\cdot]\!]_{\mathcal{M}^*} \rangle$ be defined such that $\mathcal{W}_{\mathcal{M}^*} = \mathcal{W}_{\mathcal{M}}$; $\mathcal{N}_{\mathcal{M}^*} = \mathcal{N}_{\mathcal{M}}$ and $[\![\cdot]\!]_{\mathcal{M}^*}$ is such that (i) $[\![\mathsf{A}]\!]_{\mathcal{M}^*} = \{w \in \mathcal{W}_{\mathcal{M}^*} | w(\mathsf{A}) = q\}$; and (ii) $\neg, \wedge, \vee$, and $\blacksquare$ are defined in the normal way for neighborhood semantics. The proof proceeds by induction.

[21] Given that Π is finite, this follows from the fact that, since $\bigcup \varnothing = \varnothing$, we have that for all $\pi \in \Pi$, $\varnothing \sqsubseteq \pi$.

[22] This follows from the fact that (i) $p \sqsubseteq \pi$ iff $(\mathcal{W} - p) \sqsubseteq \pi$ and (ii) $Max_w(\pi) \subseteq p$ iff $Max_w(\pi) \cap (\mathcal{W} - p) = \varnothing$.

	Small	Large
Black	125	0
White	5	20
Red	5	20
Blue	5	20
Green	5	20
Yellow	5	20

Figure 2: The contents of a bag of 250 balls.

(such as **Agglomeration**) which should, in fact, be classified as valid when the context is held fixed. Lightweight contextualism differs from heavyweight contextualism in subscribing to the former idea but not the latter.

Our objections, in §2, were directed exclusively at the second commitment of heavyweight contextualist theories. It strikes us as highly plausible that our talk about the modality of normality, like our talk about other modalities, would be context sensitive. What strikes us as implausible is that the context can shift in exactly the places and in exactly the ways required to explain apparent cases of **Agglomeration** failure.

Whereas the contextualist variant of the standard model is a type of heavyweight theory, the kind of contextualism espoused above is strictly lightweight. Our explanation of cases of **Agglomeration** failure is compatible with the context remaining unchanged across the evaluation of the premises and the conclusion. We assume only that what property is expressed by normality talk is partially dependent on context.

Objection: According to the proposal, the modality of normality is not closed under single-premise entailment (as observed in **Fact 4**). Even if **Agglomeration** fails, surely this makes it too weak?

Reply: We think that, in fact, this is precisely what we should want the framework to predict. To see why, imagine a bag containing 250 balls of six different colors and two different sizes. Suppose that the bag contains 125 black balls, all of which are small. Additionally, suppose that for each remaining color, it contains 25 balls that color, of which 5 are small and 20 large. **Figure 2** depicts the composition of the bag

In line with our observations above, we claim that it is coherent to hold that both (a) it would be abnormal not to draw a black ball from the bag and (b) it would not be abnormal not to draw a small ball from the bag. After all, the ratio of black balls to balls of any other color is 5:1. In contrast, the ratio of small balls to large balls is only 3:2. Together, however, these claims amount to a failure of single-premise closure. After all, drawing a black ball from the bag necessitates drawing a small ball.

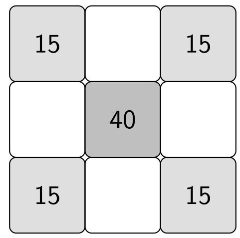

Figure 3: An unusually shaped dartboard.

Our proposal is able to accommodate these judgments. Suppose that the only relevant subject matters are the subject matters: `What color of ball is drawn?` and `What size of ball is drawn?`. The state of affairs of drawing a black ball will be uniquely associated with the former. In contrast, the state of affairs of drawing an large ball will be uniquely associated with the latter. Suppose that it is certain that some ball or other will be drawn from the bag. Then, wherever $1.5 \leq t \leq 5$, drawing a black ball will be necessitated by all maximally normal states in the former but drawing an large ball will be compatible with some maximally state of affairs in the latter.

Objection: Even if normality is not closed under single-premise entailment, isn't giving up **Weakening** (as observed in **Fact 5**) unmotivated?

Reply: We don't think so. Despite superficial appeal, on reflection the principle also admits of counter-examples. Imagine a square board divided into nine equal-sized regions, as depicted in **Figure 3**. A dart landing in the centre square scores forty points, while a dart landing in any corner square scores 15 points. A dart landing in any other square scores nothing. Imagine an expert dart player throws a dart at the board so that its probability of hitting any given square is proportional to the points value of the square.

The squares on the board can be grouped into rows, columns and numerous other ways. In different contexts, different ways of grouping the squares may be salient. Nevertheless, we claim that it is possible to get in a state of mind in which each of the following three judgments are appealing. First, it would be abnormal for the dart not to land in the central row. After all, the chance of the dart landing in the central row is $\frac{1}{3}$ greater than that of it landing in the top row and, similarly, $\frac{1}{3}$ greater than that of it landing in the bottom row. Second, it would be abnormal for the dart not to land in the central column. After all, the chance of the dart landing in the central column is likewise $\frac{1}{3}$ greater than that of it landing in the left column and, similarly, $\frac{1}{3}$ greater than it landing in the right column. Third, it would be abnormal for the dart not to land in one of the corner squares. After all, the chance of the dart landing in one of the corner squares is $\frac{1}{2}$ greater than that of it not landing in one of the corner squares.

On our lightweight contextualist view, the explanation for these judgements is simple. There is some context in which all of the three sentences used to express them are true.Yet, together, they amount to a counter-example to **Weakening** (and, *a fortiori*, also to single premise closure). After all, for the dart to land in one of the corner squares just is for it not to land in either the central row or the central column.

6 Normality and Justification

Normality has played an increasingly prominent role in recent work in epistemology. This work has tended, whether tacitly or explicitly, to assume a version of the Standard Model. Accordingly, it has generally taken **Agglomeration** for granted. In this section, we consider the implications of our account for epistemology. We will suggest that, in at least one key respect, epistemic theorizing in terms of normality is better-placed under our proposal than it would be under the Standard Model.

A number of authors have proposed a close connection between normality and epistemic justification (Smith (2010, 2013, 2016, 2018, 2021); Goodman (2013); Goodman & Salow (2018)).

Justification An agent is justified in believing that ϕ obtains iff given things are the way her evidence represents them to be, $\neg\phi$ would be abnormal.

Justification does not appeal to what would be abnormal, *simpliciter*. Rather, it is framed in terms of conditional normality. What an agent is justified in believing is a matter of whether one state of affairs (corresponding to the content of her belief) would be abnormal conditional on a second (corresponding to how her evidence represents things to be) obtaining.[23]

[23] Another, partially overlapping, strand of work proposes a connection between normality and what can be known (Greco (2014); Goodman & Salow (2018, 2021); Beddor & Pavese (2018)) (we are grateful to [redacted] for helpful discussion on this point). On this picture, an agent is in a position to know that ϕ obtains on the basis of E iff given things are as they are represented to be by E, $\neg\phi$ would be *sufficiently* abnormal. Within the standard model, the idea is (roughly) that ϕ is sufficiently normal given ψ at w only if ϕ is true at all ψ-worlds which are not (much) less normal than w. An obvious worry is that failures of **Agglomeration** for $\blacksquare$ will motivate failures of multi-premise closure for knowledge on this picture.

This worry is unwarranted, however. Crucially, to ensure factivity, the standard for being sufficiently abnormal must be assumed to be contingent and set in such a way that, necessarily, no sufficiently abnormal state of affairs obtains (Goodman & Salow (2018); Beddor & Pavese (2018)) . Accordingly, at each world, there can be no inconsistent set of states of affairs Γ such that, for each $\phi_i \in \Gamma$, $\neg\phi_i$ would be sufficiently abnormal (given ψ). Yet, the kinds of cases which motivated failures of **Agglomeration** all involved an inconsistent set of states of affairs, such that each state of affairs in the set has the $\blacksquare$ property (or, put another, such that for each state of affairs in the set, it would be abnormal for that state not to obtain). Accordingly, there is reason to think that the kinds of failures of **Agglomeration** which arise for $\blacksquare$ need not arise for the corresponding modality invoked by normality-theoretic accounts of knowledge.

The category of conditional modalities is familiar from reasoning about deontic modality (Hansson (1969); Van Fraassen (1972)). Just as we can ask what would be permissible/impermissible conditional on some particular state of affairs obtaining, we can likewise ask what would be normal/abnormal conditional on a particular state of affair obtaining (Smith (2007)).

The framework introduced above does not permit us to formulate claims about conditional normality. The language introduced in **Definition 1** is limited to claims about what would be normal or abnormal, *simpliciter*. To remedy this, we could always move to a richer language containing every sentence of $\mathcal{L}$ as well as a binary operator such that, if ϕ, ψ are sentences of the enriched language, then $\blacksquare^{\phi}\psi$ is, too. Intuitively, $\blacksquare^{\phi}\psi$ is interpreted as saying that, given ϕ obtains, $\neg\psi$ would be abnormal.

What is the logic of conditional normality? An appealing null hypothesis is that it is at least as strong as the logic of unconditional normality (Boutilier (1994c,a,b); Smith (2007)). That is, if some argument involving only sentences of in $\mathcal{L}$ is valid, then, for any ϕ, the result of substituting $\blacksquare^{\phi}$ for every instance of $\blacksquare$ will likewise be a valid argument.[24] This null hypothesis has the advantage of tying the relatively unfamiliar modality of conditional normality to the comparatively familiar modality of unconditional modality. We can reasonably hope to learn a lot about the features of the former by investigating the features of the latter.

Yet, at least under this null hypothesis, **Justification** will generate some surprising results, given the standard model. It is widely assumed that justification does not agglomerate. An agent may be justified in believing each of a set of claims, given some evidence, and yet fail to be justified in believing their conjunction (Kyburg (1961); Makinson (1965)). Indeed, failures of agglomeration for justification can be motivated by cases of exactly the kind which motivated failures of agglomeration for normality above.

For any person born in the US in 2023, we would be justified in believing that they will not die before reaching the age of 30. Yet, equally, we would not be justified in believing that no-one born in the US in 2023 will die before 30. Indeed, it seems plausible that we are justified in believing the opposite, that at least some person born in 2023 will die before 30.

The proponent of **Justification** who accepts the standard model cannot accommodate these judgments. As long as they maintain that the logic of conditional normality is at least as strong as that of unconditional modality, they are committed to accepting that justification agglomerates. As a result, they face a dilemma. Either they must deny that an agent is justified in believing, of any person born in the US in 2023, that that person will live to 30, or else they must accept that we are justified in believing that no-one born in the US in 2023 will

[24] We may wish to leave open that the logic of conditional normality will be strictly stronger. For example, it is plausible that $\blacksquare^{\phi}\phi$ should be a theorem, for any ϕ in the enriched language.

die before reaching 30. Neither alternative looks particularly tenable.[25]

On our account, in contrast, the proponent of **Justification** faces no such tension. By allowing for failures of **Agglomeration** for normality in precisely the kind of cases which lead to agglomeration failure for justification, the account is able to accommodate the kinds of judgments reported above. As such, in addition to its independent motivation, the approach to normality developed above appears a better fit for the theoretical role attributed to normality in theorizing about justification.[26]

7 Conclusion

The modality of normality is not an objective modality. It does not have a normal logic and, as a result, cannot be characterized as a restriction of metaphysical modality. This conclusion is at odds with the prevailing approach to theorizing about normality, what we have called the Standard Model.

In failing to satisfy **Agglomeration**, the modality of normality resembles other putatively non-objective modalities, such as deontic and epistemic modality. However, at least on the account developed in the latter half of this paper, it also preserves a number features of the objective modalities which its epistemic and deontic counterparts are often assumed to lack. In this way, normality occupies an interesting, medial place in the space of modalities. Despite not being a restriction of metaphysical modality, it remains a property of states of affairs (rather than being sensitive to modes of presentation) and exhibits a number of common logical features. While the formal properties of modalities of this type have been well-studied specific examples have been rare, making the modality of normality worthy of further consideration.

[25] Smith (2016, §4.3) is sensitive to this tension and opts for the latter horn of the dilemma.

[26] Despite concerns raised in footnote 24, it is worth considering the predictions of our account under the hypothesis that, for arbitrary ϕ, the logic of $\blacksquare^{\phi}$ is no stronger than the logic of ϕ (i.e., that if an argument involving only sentences of $\mathcal{L}$ is invalid, then the result of substituting $\blacksquare^{\phi}$ for $\blacksquare$ will be invalid). While we suggested that there are good reasons to deny single-premise closure of $\blacksquare$, single-premise closure for justification appears more plausible. Accordingly, if single-premise closure also fails for $\blacksquare^{\phi}$, there may be reason to adopt the following revision of the relationship between normality and justification.

Justification*   An agent is justified in believing ϕ iff there is some ψ such that (i) given that things are the waye her evidence represents them to be, $\neg\psi$ would be abnormal; and (ii) ψ necessitates ϕ.

Observe, importantly, that the move from **Justification** to **Justification*** will not help proponents of the Standard Model to avoid prediction agglomeration for justification (Indeed, under the Standard Model, the two are equivalent). We are grateful to [redacted] for encouraging us to think about this issue.

References

Asher, Nicholas, & Morreau, Michael. 1995. What Some Generic Sentences Mean. *Pages 300–338 of: The Generic Book*. University of Chicago Press Chicago.

Asher, Nicholas, & Pelletier, Francis Jeffry. 1997. Generics and Defaults. *Pages 1125–1177 of: Handbook of Logic and Language*. Elsevier.

Asher, Nicholas, & Pelletier, Francis Jeffry. 2012. More Truths about Generic Truth$_*$. *Pages 312–333 of:* Mari, Alda, Beyssade, Claire, & Del Prete, Fabio (eds), *Genericity*. Oxford University Press.

Bacon, Andrew, & Zeng, Jin. forthcoming. A Theory of Necessities. *Journal of Philosophical Logic*.

Beddor, Bob, & Pavese, Carlotta. 2018. Modal Virtue Epistemology. *Philosophy and Phenomenological Research*, **101**(1), 61–79.

Boorse, Christopher. 1977. Health as a Theoretical Concept. *Philosophy of Science*, **44**(4), 542–573.

Booth, Richard, Meyer, Thomas, & Varzinczak, Ivan. 2012. PTL: A Propositional Typicality Logic. *Pages 107–119 of:* Farinas del Cerro, Luis, Herzig, Andreas, & Mengin, Jerome (eds), *Logics in Artificial Intelligence*. Springer.

Boutilier, Craig. 1994a. Conditional Logics of Normality: A Modal Approach. *Artificial Intelligence*, **68**(1), 87–154.

Boutilier, Craig. 1994b. Conditional Logics of Normality as Modal Systems. 7.

Boutilier, Craig. 1994c. Unifying Default Reasoning and Belief Revision in a Modal Framework. *Artificial Intelligence*, **68**(1), 33–85.

Carter, Sam. 2019. Higher Order Ignorance Inside the Margins. *Philosophical Studies*, **176**(7), 1789–1806.

Chellas, Brian F. 1980. *Modal Logic: An Introduction*. Cambridge University Press.

Delgrande, James P. 1987. A First-Order Conditional Logic for Prototypical Properties. *Artificial Intelligence*, **33**(1), 105–130.

Dretske, Fred I. 1981. Knowledge and the Flow of Information. *Noûs*, **19**(2), 299–306.

Dretske, Fred I. 1988. *Explaining Behavior: Reasons in a World of Causes*. Vol. 100. MIT Press.

Eckardt, R. 1999. Normal Objects, Normal Worlds and the Meaning of Generic Sentences. *Journal of Semantics*, **16**(3), 237–278.

Fetzer, James H. 1982. Probabilistic Explanations. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, **1982**, 194–207.

Fetzer, James H. 1983. Probability and Objectivity in Deterministic and Indeterministic Situations. *Synthese*, **57**(3), 367–86.

Giere, Ronald. 1973. Objective Single-Case Probabilities and the Foundations of Statistics. *Pages 467–483 of: Studies in Logic and the Foundations of Mathematics*, vol. 74. Elsevier.

Gillies, Donald. 2000. Varieties of Propensity. *British Journal for the Philosophy of Science*, **51**(4), 807–835.

Gillies, Donald. 2016. The Propensity Interpretation. *In:* Hajek, Alan, & Hitchcock, Christopher (eds), *The Oxford Handbook of Probability and Philosophy*. Oxford: OUP.
Goldman, Alvin I. 1986. *Epistemology and Cognition*. Vol. 98. Harvard University Press.
Goodman, Jeremy. 2013. Inexact Knowledge without Improbable Knowing. *Inquiry: An Interdisciplinary Journal of Philosophy*, **56**(1), 30–53.
Goodman, Jeremy, & Salow, Bernhard. 2018. Taking a Chance on KK. *Philosophical Studies*, **175**(1), 183–196.
Goodman, Jeremy, & Salow, Bernhard. 2021. Knowledge from Probability. *Electronic Proceedings in Theoretical Computer Science*, **335**(June), 171–186.
Greco, Daniel. 2014. Could KK Be OK? *Journal of Philosophy*, **111**(4), 169–197.
Hale, Bob, & Leech, Jessica. 2017. Relative Necessity Reformulated. *Journal of Philosophical Logic*, **46**(1), 1–26.
Hansson, Bengt. 1969. An Analysis of Some Deontic Logics. *Noûs*, **3**(4), 373–398.
Ismael, Jenann. 2017. An Empiricist's Guide to Objective Modality. *Pages 109–125 of:* Slater, Matthew, & Yudell, Zanja (eds), *Metaphysics and the Philosophy of Science: New Essays*. Oxford University Press.
Kyburg, Henry E. 1961. *Probability and the Logic of Rational Belief*. Wesleyan University Press.
Leplin, Jarrett. 2009. *A Theory of Epistemic Justification*. Springer.
Lewis, David. 1988a. Relevant Implication. *Theoria*, **54**(3), 161–174.
Lewis, David. 1988b. Statements Partly about Observation. *Philosophical Papers*, **17**(1), 1–31.
Lewis, David K. 1975. Adverbs of Quantification. *Pages 178–188 of:* Keenan, Edward L. (ed), *Formal Semantics of Natural Language*. Cambridge University Press.
Loets, Annina J. 2022. Choice Points for a Theory of Normality. *Mind*, **131**(521), 159–191.
Makinson, David C. 1965. The Paradox of the Preface. *Analysis*, **25**(6), 205.
Millikan, Ruth. 1989. Biosemantics. *Pages 281–297 of:* McLaughlin, Brian P., & Beckerman, Ansgar (eds), *Journal of Philosophy*, vol. 86. Oxford University Press.
Millikan, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Vol. 14. MIT Press.
Pietroski, Paul, & Rey, Georges. 1995. When Other Things Aren't Equal: Saving Ceteris Paribus Laws from Vacuity. *British Journal for the Philosophy of Science*, **46**(1), 81–110.
Plebani, Matteo, & Spolaore, Giuseppe. 2021. Subject Matter: A Modest Proposal. *Philosophical Quarterly*, **71**(3), 605–622.
Popper, Karl. 1990. A World of Propensities - Two New Views on Causality. *Pages 24–329 of:* Popper, Karl (ed), *A World of Propensities*. Thoemmes.
Popper, Karl R. 1957. The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory. *Pages 65–70 of:* Körner, Stephan (ed), *Observation and Interpretation*. Butterworths.

Popper, Karl R. 1959. The Propensity Interpretation of Probability. *British Journal for the Philosophy of Science*, **10**(37), 25–42.
Roberts, Alexander. 2020. Relative Necessity and Propositional Quantification. *Journal of Philosophical Logic*, **49**(4), 703–726.
Rosen, Gideon. 2006. The Limits of Contingency. *Pages 13–39 of:* MacBride, Fraser (ed), *Identity and Modality*. Oxford University Press.
Schurz, Gerhard. 2001a. Pietroski and Rey on Ceteris Paribus Laws. *British Journal for the Philosophy of Science*, **52**(2), 359Ð370.
Schurz, Gerhard. 2001b. What Is 'Normal'? An Evolution-Theoretic Foundation for Normic Laws and Their Relation to Statistical Normality. *Philosophy of Science*, **68**(4), 476–497.
Schurz, Gerhard. 2002. Ceteris Paribus Laws: Classification and Deconstruction. *Erkenntnis*, **57**(3), 351Ð372.
Schurz, Gerhard. 2004. Normic Laws, Nonmonotonic Reasoning, and the Unity of Science. *Pages 181–211 of:* Rahman, S. (ed), *Logic, Epistemology, and the Unity of Science*. Kluwer Academic Publishers.
Segerberg, Krister. 1971. *An Essay in Classical Modal Logic*. Uppsala, Filosofiska Föreningen Och Filosofiska Institutionen Vid Uppsala Universitet.
Smith, Martin. 2007. Ceteris Paribus Conditionals and Comparative Normalcy. *Journal of Philosophical Logic*, **36**(1), 97–121.
Smith, Martin. 2010. What Else Justification Could Be. *Noûs*, **44**(1), 10–31.
Smith, Martin. 2013. Entitlement and Evidence. *Australasian Journal of Philosophy*, **91**(4), 735–753.
Smith, Martin. 2016. *Between Probability and Certainty: What Justifies Belief*. Oxford University Press UK.
Smith, Martin. 2017. Why Throwing 92 Heads in a Row Is Not Surprising. *Philosophers' Imprint*, **17**.
Smith, Martin. 2018. The Logic of Epistemic Justification. *Synthese*, **195**(9), 3857–3875.
Smith, Martin. 2021. The Hardest Paradox for Closure. *Erkenntnis*, 1–26.
Spohn, Wolfgang. 2002. Laws, Ceteris Paribus Conditions, and the Dynamics of Belief. *Erkenntnis*, **57**(3), 373–394.
Stalnaker, Robert C. 1999. *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford University Press UK.
Stampe, Dennis W. 1977. Towards a Causal Theory of Linguistic Representation. *Midwest Studies in Philosophy*, **2**(1), 42–63.
Strohminger, Margot, & Yli-Vakkuri, Juhani. 2019. Knowledge of Objective Modality. *Philosophical Studies*, **176**(5), 1155–1175.
Thompson, Michael. 2008. *Life and Action: Elementary Structures of Practice and Practical Thought*. Harvard University Press.
Van Fraassen, Bas C. 1972. The Logic of Conditional Obligation. *Journal of Philosophical Logic*, **1**(3/4), 417.
Wachbroit, Robert. 1994. Normality as a Biological Concept. *Philosophy of Science*, **61**(4), 579–591.
Williamson, Timothy. 2016. Modal Science. *Canadian Journal of Philosophy*, **46**(4-5), 453–492.

Williamson, Timothy. 2017. Modality as a Subject for Science. *Res Philosophica*, **94**(3), 415–436.
Yablo, Stephen. 2014. *Aboutness.* Princeton University Press.
Yalcin, Seth. 2011. Nonfactualism About Epistemic Modality. *In:* Egan, Andy, & Weatherson, B. (eds), *Epistemic Modality.* Oxford University Press.
Yalcin, Seth. 2016. Modalities of Normality. *Pages 230–255 of:* Charlow, Nate, & Chrisman, Matthew (eds), *Deontic Modality.* Oxford: Oxford University Press.