Strikingness: When Facts Call Out for Explanation

Harjit Bhogal

Abstract

Some facts are *striking* – they *call out for explanation*. They are facts that are particularly theoretically important to explain; facts that we think are very unlikely to be flukes or coincidences. For example, a sequence of 50 coin tosses landing heads calls out for explanation. Many other sequences of coin tosses do not. Similarly, the uniformity of temperature of the cosmic microwave background radiation calls out for explanation, while other distributions of temperature would not.

Strikingness plays a very important role in theory choice in both science and philosophy, but it's deeply underexplored. The aim of this paper is to exhibit the important epistemic role strikingness has in connection to theory choice; to explore possible accounts; and to defend a novel account where a fact being striking is to do with having an appearance of authorship.

In scientific practice, and in everyday life, some facts are *striking*. That is, they *call out for explanation* in a way that other facts do not. Imagine, for example, that I toss a normal-looking coin and it lands heads 50 times in a row. This calls out for explanation in a way that other sequences of 50 coin tosses would not. Or imagine that I, a very bad darts player, am trying to throw darts at the bullseye, but I keep missing and hitting other numbers. However, the numbers that I hit keep spelling out the birthdays of people in my family. That fact seems to call out for explanation. Or consider the discovery of eighteenth century doctor John Arbuthnot that in every year between 1629 and 1710 more males had been born in London than females. That is a striking fact that needs explanation.

I'll argue that this notion of strikingness, or calling out for explanation, (I'll use these terms synonymously) plays a very important role in our theory choice — very roughly, when a fact is striking we are less willing to accept theories that imply that it's just a fluke, or a coincidence. If we see a theory that implies that it's just a coincidence that more males than females were born in each of those years in London then that's a mark against the theory (not necessarily a conclusive one, but a mark nonetheless).[1] As we will see, strikingness is central to theory choice in many scientific and philosophical domains.

[1] The satisfying explanation of Arbuthnot's regularity, by the way, is that there is evolutionary pressure for there to be a one to one sex ratio at the time of sexual maturity, but infant mortality in humans is higher amongst males. So there is evolutionary pressure for more males to be born.

However, strikingness is surprisingly unexplored, given its importance. Although it comes up in various areas of philosophy there is little systematic investigation of the concept.[2]

My aim, then, is to take on this investigation. I'll clarify the notion of strikingness and exhibit its important epistemic role in connection to theory choice. In particular, I'll argue that it's important to our epistemic practice, in science and elsewhere, that some facts need explanation more than others. Then I'll consider and criticize some possible accounts, some that exist in the literature and some of my own devising. I'll then defend a novel account.

However, given the sparsity of the existing literature, the arguments here should be construed as more exploratory than conclusive. The hope is that I open up avenues of investigation, rather than close them down.

1 Characterizing the Notion

Let's start by giving an initial characterization of the concept of strikingness, before we go on to consider possible accounts. To do this, let's look in more detail at the coin toss case.

Imagine that I take a normal looking quarter and toss it 50 times. It lands heads every time. This fact is striking; it calls out for explanation.

When I say that it calls out for explanation I'm not making a psychological claim. I don't, for example, mean that I, or you, or the average person, would be surprised, or curious, or would have the disposition to investigate further. Though certainly if you observed the 50 coin tosses you would likely be surprised, curious, and perhaps disposed to investigate further.

Rather, I mean something *epistemic*. When a fact calls out for explanation it's a mark against certain theories — those theories which imply that there is no explanation which properly answers that call. That is, a striking fact is a *pro tanto* reason to reject such theories.

Or, to put it another way, a fact which calls out for explanation is a mark against theories which say that the fact is a fluke, or coincidence.

For example, our initial theory of the coin toss is just that it's a fair coin tossed normally. But this theory implies that it's just a coincidence that the coin landed heads 50 times in a row – that there is no explanation that properly answers the call for explanation. And as such, the observation of the 50 heads is a mark against that theory. We should gain credence in alternative theories — like it being

[2] Though there is some, especially the work of Dan Baras [fort.b, fort.a]. Roger White has also discussed these issues in a variety of places [e.g. 2005, 2007].

a double-sided coin, or there being some trickery in how the coin is tossed.[3]

Notice that observing a 'random-looking' sequence of 50 coin tosses – say HTHHHHTHTHTTT-THHHHTHTHTHHTHTTTTHHTHHTTTHHHHTHHTTH – does not provide reason to reject the fair coin theory. This sequence does not call out for explanation.

(And, to foreshadow some later discussion, notice that the 50 heads sequence is very unlikely given the theory that it's a fair coin tossed normally, *but so is this random-looking sequence*. In fact, *every* sequence of 50 coin tosses has the same probability given that theory — $1/2^{50}$. But not all of them call out for explanation. A good account of strikingness needs to be able to distinguish between the uniform sequence of heads and a 'random-looking' sequence of coin tosses.)

An important clarification: A fact being striking is not a mark against any theory that does not explain it. Standard plate tectonics doesn't do anything to explain why the coin tosses all landed heads, but that's not a mark against plate tectonics because it says nothing at all about the coin toss situation. Rather, a fact being striking is a mark against any theory that implies that there is no explanation which answers the call.

1.1 Which Explanations Answer the Call?

Things get a little complicated at this point, though, because only some types of explanation properly answer the call for explanation. Consider an explanation of why the first coin toss landed heads that cites the microphysical details of the coin toss and the surroundings — the velocity of the toss, the spin, the movement of the air particles and so on — as well as the basic physical laws to explain why it landed heads. Of course, in practice we couldn't access this explanation, given the limitations of our physical knowledge, but in principle it exists. Then conjoin this explanation with the analogous explanation for each of the 49 other coin tosses. This seems to be an explanation for why the 50 coin tosses landed heads.[4]

So, a defender of the view that it's a fair coin tossed normally might accept that the sequence of 50 heads calls out for explanation, but claim that their view can provide an explanation: this conjunctive explanation just described. But clearly, this isn't a good defense – the fact that the coin landed heads 50 times in a row is still a reason to reject the view that it's a fair coin tossed normally. The conjunctive explanation isn't the type of explanation we want here — it doesn't appropriately answer the call for explanation.

[3] Note, I'm using 'theory' in a permissive way. Your theory doesn't need to be anything complicated or scientifically rigorous — though it can be. By 'theory' I just mean your set of beliefs about the structure of the situation.

[4] Perhaps you think that this just doesn't count as an explanation. If you do I have no particular reason to argue since, as I will discuss very soon, I think there is a sense in which this explanation is very flawed, and the difference between a bad explanation and no explanation at all might be merely verbal.

Although this judgement is clear, the general question of how to characterize the set of explanations that do appropriately answer the call is more difficult. Consider the conjunctive explanation again: It tells us that the particular microphysical initial conditions of the coin tosses were such that they led to all the coin tosses landing heads. But, it's natural to ask, why did the initial conditions have that feature? Couldn't the initial conditions have very easily not had the feature that they led to 50 heads? The explanation would be better if, for example, it told us that the coin was double-headed and so *every* possible set of initial conditions for the 50 coin tosses would have led to 50 heads. Or it would be better if the explanation told us the the coin-tosser had great skill at manipulating the conditions so that they were very likely to produce the physical conditions that would lead to the coins landing heads. These types of explanations would appropriately answer the call for explanation.

The underlying intuition here is that the conjunctive explanation makes the 50 heads sequence appear *fragile* — it could very easily have not held — while the double-headed coin explanation makes the 50 heads sequence seem very *robust*. (This is in the spirit of Roger White's [2005] discussion of 'explanatory stability' and 'explanatory urgency'.) So, it seems like explanations that don't make the phenomenon in question appear robust don't appropriately answer the call for explanation.

This, I think, is broadly correct, and is enough for our purposes in the rest of the paper. But, developing this robustness thought into a full and precise account of what explanations appropriately answer the call for explanation is tricky. One way to see these difficulties is to note — following Bhogal [2020] — that the problem with explanations like the conjunctive explanation which do not appropriately answer the call for explanation is very closely analogous to the problem with lower-level explanations of higher-level phenomena. For example, imagine I try to explain why inflation rose by appealing to the complete microphysical details of the particles that make up the economic system and showing how they evolved into a situation where inflation rose. This microphysical explanation seems flawed for similar reasons to the conjunctive explanation of the 50 heads — the explanation is very fragile and doesn't makes the rise in inflation appear robust.

There is a vast literature on exactly what is wrong with these lower-level explanations of higher-level, special scientific, phenomena and how to cash out these intuitions about robustness. However the issues turn out to be rather difficult. It's not clear whether robustness is exactly the relevant factor here (see Weslake [2010], Wilson [1994]), or whether some related notion like difference-making (e.g Strevens [2008], Woodward [2003]) or proportionality (e.g. Yablo [1992]) is at work.

So let's take stock of where we are. Facts which are striking — which call out for explanation — are such that it's a mark against a theory if that theory implies that there isn't an appropriate explanation of that fact. Roughly speaking, the appropriate explanations are ones that makes the fact in question appear robust — like it couldn't have easily been false. But this is only rough; to be more precise we need to dig into the literature on levels of explanation — the badness of an inappropriate explanation

is similar to the badness of an explanation that is pitched at the wrong level. But we won't do any more of that digging here. Going forward I'm just going to call the appropriate types of explanation 'robust explanations', with the understanding that this terminology may be slightly misleading.

So we have our initial characterization of the notion of strikingness: Facts which are striking are such that it's a mark against a theory if that theory implies that there isn't a robust explanation of that fact.

2 Some Examples

Let's see some more examples of striking facts. This will help us get a better grip on the concept and will let us start to see the important epistemic role strikingness plays across a variety of domains.

Some examples I'll just mention, some I'll discuss briefly.

Coin Tosses I toss a normal-looking coin and it lands heads 50 times in a row.

Darts I, a very bad darts player, am trying to throw darts at the bullseye, but I keep missing and hitting other numbers. However, the numbers that I hit keep spelling out the birthdays of people in my family.

State Lotteries The same numbers are drawn in the lotteries of ten different states in the same week.

Mathematical Truths There is a striking correlation between the mathematical truths and our mathematical beliefs.

Moral Truths There is a striking correlation between the moral truths and our moral beliefs.

These last two striking facts are at the heart of arguments for anti-realist or reductionist positions about the mathematical and moral domains.

For example, Field [1989, p.26] claims that 'the correlation between mathematicians' belief states and the mathematical facts is so striking as to demand explanation; it is not the sort of fact that is comfortably taken as brute.' And, he argues, Platonist views of mathematics leave this correlation without a proper explanation, so such views should be rejected. The thought is that any good theory of mathematical facts must imply that the fact has an explanation.

What's more, an explanation of this correlation which makes it extremely fragile – one that suggests that we could very easily had completely unreliable mathematical beliefs is not the type of explanation that Field is looking for here. Field, and many others, would reject theories that say that the correlation doesn't have a *robust* explanation.

Higher and Lower-level Laws Higher-level, special scientific, laws harmonize with the lower-level laws.

For example, Malthus's law of population ecology (a law governing the growth of certain populations) somehow holds simultaneously with the fundamental laws of physics. It's striking that they fit together in this way. How do 'the physical objects constituting rabbits (etc.) "know" to move, projectibly, in ways consistent with Malthus's Law.'? [Callender and Cohen, 2010, p. 428]

Quantum Entanglement When two electrons are prepared in a certain way their behavior in properly calibrated spin measuring devices is strikingly anti-correlated. Whenever one is observed to have spin up in the x direction, the other is observed to have spin down.

These last few examples have been of striking correlations. It's clear, though, that there are lots of correlations that are not striking — that do not call out for explanation. Between 2000 and 2009 there was an extremely high correlation between the per capita cheese consumption in the USA and the number of people who died by becoming tangled in their bedsheets. In the same period the divorce rate in Maine was highly correlated with the per capita consumption of margarine. It's clear, I take it, that these correlations are not striking — we are perfectly happy to accept theories that imply that there is no robust explanation of these correlations. (See http://www.tylervigen.com/spurious-correlations for these and many other spurious correlations.)

The Continued Regularity of the World The world has been fairly uniform; the future has seemed to, broadly speaking, resemble the past.

It's often claimed that Humeanism — the view that there are no necessary connections between distinct existences — cannot properly explain this striking fact and so should be rejected [Armstrong, 1983, Strawson, 1989, Beebee, 2011].

Tides The tides come in and go out at (somewhat) regular intervals.

Microwave Background Radiation The temperature of the cosmic microwave background radiation (CMB) is (almost) the same wherever we look in the universe.

So now we have a variety of examples which start to show how strikingness plays a very important role in theory choice across a wide range of domains.

But it will be useful to discuss this last example in a bit more detail – let's briefly consider the scientific role of this striking fact. 'Standard' Big Bang theory — that is to say, the version of the theory that was standard in the 1960s and 1970s — faced a problem. (Or rather, a few problems, but I'll focus on one.) Wherever we look in the universe the temperature of the CMB is almost the same.[5]

This type of phenomenon isn't usually puzzling. Normally uniformity of temperature is what we expect because of the process of thermal equilibration — if we put a hot thing and a cold thing close to each other they will tend to equilibrate to the same temperature. But, given Standard Big

[5]The discussion of the physics here broadly follows Guth and Steinhardt [1989] and Maudlin [2007, Chapter 1].

Bang theory, that can't be what's going on in this case. This is because there are regions of the universe where the CMB is the same temperature but which are so far apart that they could never have interacted with each other (given that physical processes can't propagate faster than the speed of light). And so the process of thermal equilibration could not be the explanation for the uniformity here.

Importantly, though, this uniformity isn't inconsistent with Standard Big Bang theory — consistency just requires a choice of the right initial conditions. These initial conditions, if they hold, can explain the uniformity of the CMB [Earman and Mosterin, 1999, p. 19]. The problem is one of fine-tuning — the initial conditions need to be very specific. Guth and Steinhardt [1989, p. 36] note that the standard theory implies that 'To begin with at 10^{-43} s [seconds after the Big Bang], the matter must start in a peculiar state of extraordinary but not quite perfect uniformity …… The peculiarity of the initial state of matter required by the standard model is called the smoothness problem.'

Physicists find the explanation of the uniformity of the CMB that cites the very specific initial condition extremely unsatisfactory (see, for example, the quotes collected in McCoy [2015, p. 28]). This unsatisfactoriness seems to do with the fragility of the explanation. Guth and Steinhardt [1989, p. 34], for example, consider a competitor to Standard Big Bang theory – *Inflationary* Big Bang theory. They say that one of the major advantages that Inflationary Big Bang theory has over the standard theory is that 'from almost any set of initial conditions the universe evolves to precisely the situation that had to be postulated as the initial state in the standard model.' That is to say, the explanation that the inflationary version of Big Bang theory can give of the uniformity of the CMB is much more robust — basically any set of initial conditions would lead to the outcome we observe. Similarly, McCoy [2015, p. 31] argues that the best diagnosis of physicists dissatisfaction with the standard theory's explanation of the uniformity is that the explanation lacks robustness.

So, the problem is that the uniformity of temperature can't be explained in an appropriate way given Standard Big Bang theory. And this is taken to be a powerful reason to reject the standard theory. So we have a striking fact – the uniformity of temperature of the CMB – but, given the standard theory, there is no robust explanation. And this is a pro tanto reason to reject the standard theory.

There's another important feature of the case that we should emphasize as well. As we noted, given the standard theory we need to fine-tune the initial conditions in order to explain the uniformity of temperature. But this isn't just true for the uniformity — in order to explain *any* particular distribution of the temperature of radiation throughout the universe the standard theory would need fine-tuned initial conditions.

This, in itself, is not a problem for the standard theory. The mere fact that the standard theory can't give a robust explanation of the distribution of temperature is not a problem with the theory. The

problem is that the standard theory can't give a robust explanation of the distribution of temperature, *and that the distribution is uniform*! This is clear from the way that physicists stress the importance of the uniformity as a problem for the standard theory. For example, Dodelson [2003] says that the uniformity of temperature is 'disturbing' in light of facts about the standard theory and asks 'Why then is the CMB so uniform? This is a profound problem…' (p.143) (Again, see McCoy [2015, p. 28] for a collection of similar quotes.) It is the uniformity, in particular, that is the striking fact, it seems to call for explanation more than other distributions of temperature.

To summarize: The uniformity of the temperature of the CMB is a striking fact — it calls out for explanation in a way that most distributions of temperature do not. But given Standard Big Bang theory there is no robust explanation. This is taken to be a powerful reason to reject the standard theory.

This case, along with the other we have seem, illustrates the role that strikingness plays in our epistemic practice, in science and elsewhere. That some facts call out for explanation and others do not is important for understanding our practices of theory choice.

2.1 Strikingness and Confirmation

One quick clarification about the role of strikingness in theory choice, before we move on to consider possible accounts.

The cases we discussed should make clear how strikingness has an important role in theory choice. In all of the cases, a theory saying that the striking fact doesn't have a robust explanation is a *pro tanto* reason to reject that theory.

But, the role of striking facts is very different from other facts that provide reasons to reject your theory. In particular, there are cases where you observe a fact that seems to be inconsistent with the theory that you accept — where your theory makes a false prediction. Of course, such a fact is a pro tanto reason to reject your theory.

It's important to be clear, though, that such cases are not striking facts, at least in the sense of strikingness under investigation here. Imagine, for example, your theory of the stock market predicts that some stock will have a high price but it in fact has a low price. That's a pro tanto reason to reject your theory. But the fact that the stock has a low price isn't striking in the relevant sense – it doesn't call out for explanation more than other facts. One way to see this is that the low price of the stock isn't a mark against *any* theory that says that it doesn't have a robust explanation. We may well be perfectly happy to accept some theory of the stock market that says that the low price is just a fluke. Rather, the low price of the stock is a mere counterexample to your theory, not a fact that especially demands explanation.

Going forward we should be mindful of the distinction between facts that call out for a robust explanation, and other facts, which may well be surprising or puzzling, but just because our favorite theory appears to get them wrong.

3 Some Accounts of Strikingness

In this section I'm going to consider, and criticize, some possible views of strikingness – some in more detail than others. I'll consider some existing views of strikingness but, since there are few of these I'll also consider some possible alternative approaches of my own devising.[6]

Criticism of these approaches is part of the motivation and argument for the alternative approach I develop in the next section.

3.1 The Alternative Hypothesis View

Perhaps most common reaction to cases like **Coin Tosses**, is to say that when the coin lands heads heads 50 times in a row natural alternative hypotheses are suggested — alternative, that is, to our initial theory that it's a fair coin tossed normally. Perhaps it's a double sided coin, for example, or perhaps I have great skill in tossing the coin and controlling how it lands. But when the sequence of coin tosses is HTHHHHTHTHTTTTHHHHTHTHTHHTHTTTTTHHTHHTTTHHH-HTHHTTH no such natural alternative hypothesis is suggested. (Baras [fort.a, pp. 1409-1410] surveys authors who have accepted this kind of thought.)

Of course, there are hypotheses that would make sense of this random-looking sequence. Perhaps I have great control over how the coin lands, and I just happen to like that sequence so much that I replicate it with my coin. But such hypotheses seem very strange and unlikely.

This, you might think, is the difference between the striking sequence of 50 heads and the non-striking, random-looking sequence — there is a natural alternative hypothesis suggested in the first case but not the second.

These kinds of considerations motivate a theory of strikingness that has been attributed to Paul Horwich [1982, p.100-104].[7] Call this the *alternative hypothesis view*: A fact is striking if and only

[6] One view that I won't discuss is Grimm's [2008] account of the 'need for explanation'. Grimm's discussion is very rich and enlightening, but it doesn't address our issue. In particular, it focuses on a psychological sense of puzzlement rather than the epistemic sense of strikingness that we are focused on. And relatedly, his account doesn't aim to distinguish between the 50 heads sequence and the random-looking sequence of coin tosses – it implies that both need explanation.

[7] Though Horwich's account was an account of surprise, not of strikingness, but others (e.g. White [2005], Bostrom [2002]) have taken it up as an account of strikingness. And, as Baras [fort.a, p. 1409] notes, the view that has often been attributed to Horwich may not be his actual view, but that doesn't matter for our ends.

if (i) the fact is unlikely given your theory and (ii) there is some alternative theory that you think is not wildly improbable that makes the theory likely.

Both the random-looking sequence of coin tosses and the 50 heads sequence are unlikely given your current theory that it's a fair coin tossed normally, but only the 50 heads sequence has a not-wildly-improbable alternative theory, for example, the theory that it's a double-sided coin.

Here's a problem with this view: Some facts are striking even when there is no alternative hypothesis that you take to be not-wildly-improbable.

For example, imagine a variant of the 50 heads coin case, where I toss the coin and it lands heads 50 times in a row. But I know that this isn't a double sided coin and that I'm not trying to control the coin toss so that it lands heads — in fact, the way I tossed the coin was somewhat different every time. And also, let's assume, that I'm an atheist so take it to be wildly improbable that God is sending me a message via coin tosses or anything like that. In fact, every hypothesis I can think of that would make sense of the 50 heads I take to be wildly improbable. Nevertheless, the 50 heads sequence is striking.

A slightly different type of example: Imagine a variant of **Microwave Background Radiation** where we have just observed the uniformity of temperature of the CMB. But, we can not yet formulate any sensible-looking scientific hypotheses that would make the uniformity likely — formulating such scientific hypotheses is not an easy task. So there is no hypothesis that we take to be not-wildly-improbable that makes sense of the uniformity. Clearly, though the uniformity is still striking.

Here's a possible response to that objection: Perhaps sometimes there is no *specific* hypothesis that makes the striking fact likely, but what about the 'catch-all' hypothesis — the hypothesis that there is some hypothesis, whatever it is, that makes the fact in question likely. Can't we defend the alternative hypothesis view by taking this hypothesis into account?

However, if we allow such catch-all hypotheses then the alternative hypothesis view reduces to this: a fact is striking if and only if (i) it is improbable given your current theory and (ii) it's not wildly improbable that there is some theory — whatever it is — that makes the fact likely.

But, this view comes very close to simply giving the Bayesian conditions for when your current theory is substantially disconfirmed. To see this consider what factors are relevant to whether a fact disconfirms your current theory. Firstly, if the probability of the fact given your current theory is low then that tends to favor disconfirmation of your current theory. Secondly, if there is an alternative theory that is probable then that tends to favor disconfirmation. And thirdly, if the probability of the fact is high given the alternative theory then that favors disconfirmation.

As we can see, these conditions are very close to the conditions given by the alternative hypothesis view when the catch all hypothesis is allowed. (The obvious difference being that the Bayesian

conditions should be understood as graded while the conditions given by the alternative hypothesis are naturally understood as thresholds. On the Bayesian approach there is no threshold for the probability of alternative theories that is required for disconfirmation, for example. Though if we were to adapt the alternative hypothesis view to yield a graded notion of strikingness, the natural way to do this would be to make the conditions identical to the Bayesian conditions on disconfirmation.)

When we allow the catch-all hypothesis to be one of the relevant alternative hypotheses, the view effectively reduces to a view that says that a fact is striking when it substantially disconfirms your current theory.[8] The problem with this view is that it is extremely uninformative. We want to know why the 50 heads sequence of coin tosses is a mark against your theory that it's a fair coin tossed normally while the random looking sequence is not. Saying that the 50 heads sequence is striking because it disconfirms your initial theory doesn't help at all. (Baras [fort.a, section 5] makes a similar argument against this disconfirmation view of strikingness.)

The first version of the alternative hypothesis view faces counterexamples; the second, where we allow the catch-all hypothesis, looks almost trivial and uninformative. Either way, the alternative hypothesis view isn't satisfactory.

3.2 Simplicity

Another natural reaction to **Coin Tosses** is that the 50 heads sequence is *simple* or easily computable, in a way that the random-looking sequence is not. Perhaps we could try to develop this into an account of strikingness — the thought being that striking facts are patterns that are simple, in some sense of simple to be specified.

I don't think this is a particularly promising approach. Here's the central worry: Consider, again, **Darts**: I, a very bad darts player, am trying to throw darts at the bullseye, but I keep missing and hitting other numbers. However, the numbers that I hit keep spelling out the birthdays of people in my family.

The problem is that the sequence of numbers that I hit is not particularly simple, but it is striking.

Perhaps the defender of the simplicity view could respond that we have not properly individuated the striking fact. The striking pattern isn't the particular sequence of numbers that I hit, but rather it's the fact that there is this matching between those numbers and my families birthdays. But is this pattern particularly simple? The matching between the numbers I hit and the birthdays of my cousin, my brother and my grandmother isn't clearly simpler than the matching between the numbers I hit and

[8] Baras [fort.a, p. 1409] suggests that Horwich's theory has often been misinterpreted and that this disconfirmation view of strikingness is more plausibly his actual view.

the birthdays of three random people who I have no relationship to. But this second matching is clearly not striking.

Perhaps the defender of the simplicity view could try to deny this last claim. Undoubtedly there are ways that they could nuance their view so there is a lot more that could be said here. But the general point is that some facts, like the matching between the numbers and the birthdays of my family members, seem to be striking because they have a certain *significance* to us — and the prospects for reducing this notion of significance to some story about simplicity don't seem promising.

3.3 Typicality

Here is a different type of reaction to the coin case. The random-looking sequence of coin tosses – HTHHHHTHTHTTTTHHHHTHTHTHHTHTTTTTHHTHHTTTHHHHTHHTTH – is the type of thing we expect to see. Of course, that particular sequence is low-probability, but it's a *normal*, *typical* sequence. And so it doesn't call out for explanation.

On the other hand, the 50 heads sequence is atypical – it has features that are very different from those that we expect to see. And so it does call out for explanation.

The challenge here, of course, is to give an account of this notion of typicality. Elga [2004] and Williams [2008] develop a notion of typicality in a very different context, but it might be of use here.

The idea is that there are certain features of sequences of events that have a very high probability (given your current theory of the situation) — call these *normal features*. A sequence is atypical if it lacks a normal feature. So, for example, it's high probability, given your current theory of coin tosses, that in a long sequence of coin tosses there will be about as many heads as tails. And that the HTH subsequence will appear about as often THT [Elga, 2004, p. 74]. An atypical sequence will lack such a feature.

Given this account of typicality we could claim that a sequence is striking if it is atypical.[9] (Perhaps we could extend this idea to facts which are not understood as sequences. Some additional problems arise if we do this but I'm not going to go into that here since, I think, this account has major problems anyway.)

The obvious problem with such an account of strikingness is that every sequence seems to be atypical in some way. Take the random-looking sequence of coin tosses: HTHHHHTHTHTTTTHH-HHTHTHTHHTHTTTTTHHTHHTTTHHHHTHHTTH. It's very-high probability that sequences of 50 coin tosses will not have the property of constituting that exact sequence of heads and

[9] Put this way, the account is in the spirit of Schlesinger's [1987] account of when facts are surprising.

tails. So, it seems like it's a normal feature of sequences of 50 coin tosses that they don't land HTHHHHTHTHTTTTHHHHTHTHTHHTHTTTTTHHTHHTTTHHHHTHHTTH. And so that random-looking sequence, is, in fact, atypical. Clearly this reasoning can be extended to every other sequence. This type of issue is discussed by both Elga and Williams.

Elga suggest that we need some constraints on which properties count as normal features of a sequence. If *any* high-probability feature of a sequence counts as normal, then all (long enough) sequences are going to be atypical. Rather, Elga suggests, the relevant properties are ones which are *simple* — that is, ones which are simply definable in terms of some language in which it is natural to describe the relevant sequences. The thought is that properties like 'there being about as many heads as tails' will be simple to define, and so relevant to the typicality of a sequence, while properties like the one described in the last paragraph will not.

But this appeal to simplicity faces similar issues to those discussed in the last section. Some sequences seem striking because they have a certain significance to us. And the prospects for reducing this notion of significance to some simply defined property which most sequences do not have don't seem particularly promising.

Perhaps we could forget about simplicity and try to come up with some other story about which properties make sequences atypical and which do not. This seems like a reasonable project. But it is, in effect, the project of trying to identity the properties that few sequences have, and in virtue of which those sequences call out for explanation. Put like this, it should be clear that this is just a slightly different framing of our project about strikingness — it not really an advance from where we started.

Consequently, I don't think the appeal to typicality helps us with our project.

3.4 Deviations from Normality

Perhaps striking facts are ones which deviate from normality. There is clearly a sense in which it's natural to say that these anomalous facts call out for explanation. But in our sense of calling out for explanation, where a fact calling out for explanation makes us suspicious of any theory which says that it is just a coincidence, then it's clear that often uniformity calls out for explanation, not just deviations from that uniformity.

3.5 (Un)Expectability

Yudell and Wong [2015] give a 'normative account' of the need for explanation. Their view is that each person has 'maps' (which are somewhat similar to theories) of parts of the world. A fact needs

explanation when it doesn't fit what is to be expected given the map. They have lots of useful and intricate discussion of the nature of such maps, but we can put that aside to focus on the main idea.

Their view is an instance of an approach where striking facts are those that are 'unexpected'. The main concern with this approach is that is doesn't properly draw the distinction between the 50 heads sequence and the random-looking sequence of coin tosses. Both sequences are unexpected – they both have low probability, for example, given our prior theory of the situation, but only one calls out for explanation.

Wong and Yudell's strategy for dealing with similar cases is to argue that certain unlikely events do not need an explanation because they are not described in terms that fit with the way in which our 'map' of the situation is formulated. For example, they say that 'The fact that Winnie won the lottery does not need an explanation because the description 'Winnie won the lottery' is not composed of the elements of any particular map by means of which we can understand why a particular person (rather than another person) won the lottery' (p.2879). But this strategy doesn't apply to the coin toss case because both sequences are described merely in terms of the coin landing heads or the coin landing tails, both of which, presumably, are elements of our map of coin tosses. And so this strategy doesn't solve the problem.

3.6 The Deflationary Approach

So far, I've been talking as if strikingness is a unified phenomenon which has a substantial role in play in our theory choice — a fact being striking provides a distinct reason to reject theories that don't give a robust explanation for that fact. But perhaps that's not the case. Perhaps strikingness is just a label we put on facts where other epistemic principles lead us to believe that the fact has a robust explanation. If this is the case, then strikingness isn't unified — we could come to believe that facts have a robust explanation for all sorts of reasons in different cases. And the strikingness of a fact wouldn't provide any reasons for belief distinct from the ones that we already have. (This view is suggested by Baras [fort.b] and Baras (ms.).)

So, let's, very briefly, consider some other epistemic principles and whether they might give us reason to believe that some facts have a robust explanation. *Bayesianism*, on its own, doesn't seem to do the job. Neither the claim that rational credences should be probability functions nor that they should be updated via conditionalization seems to do much, on its own, to suggest that some facts have robust explanations.

Appealing to *Inference of the Best Explanation* (IBE) initially seems a bit more promising. IBE tells us that we have reason to believe the best explanation of phenomena that we observe. And if robustness makes an explanation better then it seems we have reason to favor theories which give robust

explanations of phenomena. However, this reason is totally indiscriminate. It doesn't distinguish between intuitively striking facts – where we do have a distinct reason to think that there is a robust explanation – and non striking facts – where we are perfectly happy to say there is no robust explanation.

An appeal to *Induction* might help in some cases — perhaps we sometimes have reason to think that a fact has a robust explanation because similar facts in the past have had a robust explanation. It's plausible that this reasoning does apply in some cases. But I'm doubtful that it applies particularly widely. Take, for example, the coin toss case. Although I have a lot of experience with coin tosses I've never seen a sequence of 50 heads. And all the shorter sequences of heads that I have seen have been coincidences. So induction doesn't seem to be relevant to me wanting a robust explanation of the 50 heads sequence.

In fact, it seems to me rather rare that we have a substantial history of observing robust explanations to rely upon. Imagine, for example, that one day I see a lot of people on various parts of campus wearing matching t-shirts. I come to think that there must be a robust explanation of this. And it's true that I've seen similar things in the past and believe that there were robust explanations in those cases too. But that's not because I've observed, so to speak, the robust explanation in the past — I've never talked to any of the people about why they were wearing those t-shirts. Rather, I've just assumed that those past striking facts must have a robust explanation. Clearly, then, to use this track record as inductive evidence that the current instance of people wearing matching t shirts has a robust explanation is unacceptable bootstrapping. Many cases, I suspect, are like this — where the apparent inductive evidence that similar cases in the past had robust explanations in fact itself relies on an appeal to strikingness.

Perhaps instead we should appeal to *Simplicity*. The idea is that theories which imply that a striking fact doesn't have a robust explanation will be overly complicated and so that's a mark against those theories. But this, again, doesn't seem to properly distinguish between striking and non-striking cases. For example, the theory which says that there is no robust explanation of the 50 heads is no more complicated than the theory that says that there is no robust explanation of the random looking sequence of coin tosses.

Clearly, all of these doubts about how strikingness can be reduced to specific epistemic principles don't refute the deflationary view. Perhaps if we balance all our epistemic reasons just right then we can make sense of every case of strikingness. But these doubts do make it harder to see how the deflationary view would work.

And there's another consideration that is evidence against the deflationary view. Consider **Microwave Background Radiation** again. The relatively brief discussion of this case in section 2 allows us to

correctly perceive that that the uniformity of temperature of the radiation is striking. We don't need to know a lot of extra detail about the case to make the judgement that the fact calls our for explanation. We don't, for example, need to have evidence about prior cosmological theories so that we can perform inductions; and we don't have to make judgments about how the relative simplicity of the standard theory compared to alternative theories. In fact, we don't need to think very hard at all to come to the conclusion that the fact calls out for explanation. (Though of course, we would have to think very hard in order to work out what the implications of this are — whether we should reject the standard theory and accept some other theory or do something else.) If strikingness was just a label for when balancing all the other epistemic principles leads us to believe that there is a robust explanation then we couldn't quickly form such clear judgements in complicated cases where we have little information.

I'm doubtful that the deflationary view works then. But there clearly are some cases where our reason to think that some fact has a robust explanation does derive from other epistemic principles. For example, consider some intuitively non-striking fact. But then a reliable scientist tells you that the fact does have a robust explanation. That's a reason to disbelieve any theory that doesn't give a robust explanation of the fact.[10]

Given such cases we face a terminological choice. We could continue to use the terms 'striking' and 'calls out for explanation' to refer to any facts where we have reason to be doubtful of theories that don't give a robust explanation of those facts. Or we could reserve those terms for facts where those reasons aren't fully explained in terms of other epistemic principles. So cases where we believe that a fact has a robust explanation on purely testimonial or inductive grounds, for example, would not count as striking in this sense.

It's this second sense of 'striking' which fits better with the intuitive notion of strikingness, and with the cases I've discussed so far, so let's use the terms in this sense going forward.

4 The Authorship View

The last section has shown us that lots of the most natural approaches to strikingness face major problems.

In this section, then, I'm going to suggest a new approach – one that is rather different to those discussed so far. The view I develop will be a rather bold one. So bold, in fact, that some people might be inclined to reject it out of hand. (I will discuss that kind of reaction later.) But, I will claim, the view does have plausibility.

[10] Thanks to [removed for blind review] for discussion of this case.

The starting point, though, is that there is something right about the alternative hypothesis view. The thought that the difference between the 50 heads sequence and the random-looking sequence of coin tosses is something to do with the alternative ways that the 50 heads could have come about is extremely compelling. When I toss the coin and it lands heads every time, it's natural to suspect that it's a double sided coin, or something similar. The alternative hypotheses that are suggested are ones which give the 50 heads sequence the feel of a magic trick, or something similar.

A similar thing is true of **Darts** — the case where the numbers I hit on the dart board spell out the birthdays of people in my family. The case has the feel of a trick, or a conspiracy. In fact, I think the same is true for the other examples of strikingness we considered — they have the appearance of a kind of conspiracy occurring to make the striking fact hold.

Take, for example, **Higher and Lower-level Laws** — the fact that higher-level, special scientific, laws harmonize with the lower-level laws. It's common to call this a 'conspiracy' problem — there appears to be some 'metaphysically extravagant conspiracy among fundamental particles' to move in just exactly the ways that make the laws of, say, economics or population ecology, turn out to be true [Callender and Cohen, 2010, p. 428]. And it is this apparent conspiracy which makes the fact so deeply puzzling.

Or take **Quantum Entanglement**. Similarly, there is this appearance of conspiracy here, of the fundamental particles conspiring to make it the case that spin behavior of the two particles is strikingly anti-correlated. In fact, various possible explanations of the behavior of the particles have been called 'conspiratorial' (see Myrvold et al. [2019, section 3.2.2]).

In fact, I think, in all the cases of strikingness there is something like this feeling of conspiracy. And, I mean conspiracy in the literal sense – some feeling that there is some agent behind things. To put it another way, striking facts have an *appearance of authorship*.

In cases like **Coin Tosses** and **Darts** this seems like a reasonable suggestion. In those cases a conspiracy seems like a live and reasonable hypothesis — it really might be the case that someone is playing a trick on you. But it's a stranger suggestion in cases like **Higher and Lower-level Laws** and **Microwave Background Radiation** where, at least if we are not theists, we don't take there to be an author controlling things — where the conspiracy hypothesis is one that we take to be unlikely.

However, I want to claim that even in these cases there is still an appearance of authorship here — there is a sense in which the authorship hypothesis is compelling, even though we might judge it to be unlikely to be true. Looking at the literature on inference to the best explanation helps us see what this sense could be, because, in that literature, there is already a generally accepted distinction between the credence we have in an explanation and another way in which the explanation can be compelling.

Lipton [2004] gives an influential, and widely accepted, account of how inference to the best explanation works. The idea is that we consider a variety of potential explanations of the phenomenon in question and consider which one would be best. But by 'best' here, we don't mean the explanation which is likeliest to be true, because that would trivialize the IBE procedure. Rather, we mean the explanation which is 'loveliest'. And, in particular, we look at the potential explanations of the phenomenon and we consider what would be the loveliest *if it were true*. That is, we pick the explanation that 'would, if correct, be the most explanatory or provide the most understanding: the 'loveliest' explanation.' (p. 59) Where the likeliest explanation is the one that has the highest probability, the loveliest explanation is the one that would provide the most understanding.

Importantly, an explanation can be lovely even if it is unlikely — we can have reason to believe that some potential explanation is not true, but that if it were true it would be a lovely explanation of the phenomenon. (Though if we accept IBE the loveliness of the explanation does contribute to it's likeliness).

For example, Lipton notes that: 'Newtonian mechanics is one of the loveliest explanations in science and, at one time, it was also very likely. More recently, with the advent of special relativity and the new data that support it, Newtonian mechanics has become less likely, but it remains as lovely an explanation of the old data as it ever was.'

In cases like **Higher and Lower-level Laws** and **Microwave Background Radiation** the authorship hypothesis might not be a likely explanation (depending on if we are theists or not), but still, it is relatively lovely. Given this we can say that a fact has the appearance of authorship when the authorship hypothesis is a somewhat lovely potential explanation (though there's no requirement that the authorship hypothesis has to be the loveliest). A fact can have the appearance of authorship in this sense even if the authorship hypothesis is unlikely. 'Appearance of authorship', then, should be understood along the lines of the notion of loveliness that is familiar from our practice of IBE.

Of course, there is a huge project about what makes for loveliness. That's not a project that I can take on now. But, to the extent that we have a grip on how to do inferences to the best explanation we have a grip on loveliness.

So that's the suggestion: Striking facts are ones that have this appearance of authorship.

This is a very bold and simple hypothesis. I'm going to try to show that this simple view has plausibility.

To be clear, this view is about when facts call out for explanation – it's when they have the appearance of authorship – it does not say that when we have a fact which calls out for explanation we should give it an authorship-based explanation. Rather, when a fact calls out for explanation we should give it a robust explanation, perhaps that will be authorship-based, perhaps not.

How does the authorship view apply to the cases we have considered so far? Consider **Microwave Background Radiation** again. The thought is that the uniformity of temperature gives some appearance of authorship — uniformity, especially without any obvious cause, gives the impression that there is someone producing the uniformity. An agent controlling things might more naturally choose uniformity over some other random distribution of temperature, and hence the authorship hypothesis is one that gives us understanding and is lovely.[11] And similarly for uniformities more generally. Further, striking correlations, like the (anti-)correlation of the spin behavior of the two particles in **Quantum Entanglement** also have a similar appearance of authorship. And clearly, cases like **Darts** where the patterns observed have some significance to us very much suggest an author. On the other hand, I don't think anyone would judge that the 'spurious correlations' we considered above — for example, the fact that between 2000 and 2009 divorce rate in Maine was highly correlated with the per capita consumption of margarine — suggest that there is some author behind them.

The authorship view also has some substantial advantages over the previous approaches we mentioned. Some versions of the alternative hypothesis view are close to trivial — this certainly isn't the case for the authorship view. Other versions fail to make sense of cases where there we have a low credence in all the alternative hypotheses. This isn't a problem for the authorship view because there is no requirement that we have a high credence in the authorship hypothesis. Further, the view has an advantage over the views based on simplicity and typicality — in that it can clearly make sense of the cases where facts appear to be striking because they have a certain significance to us. And, further, I think the view avoids many of the other counterexamples that faced other approaches.

4.1 Agency-Detection and Salience

However, at this point, it's reasonable to be rather unconvinced. Perhaps there is something to this authorship idea; maybe reflection on the examples of striking facts that have come to your mind while reading the paper does reveal something like an appearance of authorship in those cases, but still, it's very natural to be suspicious. In particular, here's perhaps the most major concern driving that suspicion: Just what does the appearance of authorship have to with the need for explanation? How are those two things in any way connected? Why is it the authorship hypothesis that matters for strikingness?

This is a very good question. In this section I'll discuss a very important consideration relevant to answering it. This will involve a bit of a detour through psychology. We will come back to epistemic issues soon.

[11] This is in the spirit of Roberts [2012] and his discussion of things an agent might find 'choiceworthy'.

There is a huge amount of work, in psychology, cognitive science, and related fields, about the ways in which humans see agency and authorship everywhere. Humans have a very strong inclination to understand phenomena in terms of agency (see Bloom [2007] for a brief survey). The most classic example comes from Heider and Simmel [1944] 'who made a simple movie in which geometrical figures – circles, squares, triangles – moved in certain systematic ways, designed, based on the psychologists' intuitions, to tell a tale. When shown this movie, people instinctively describe the figures as if they were specific people (bullies, victims, heroes) who have goals and desires' [Bloom, 2007, p. 149] Similarly, Guthrie [1995] surveys lots of cases of humans (including scientists and philosophers!) attributing agency to a vast range of features of the natural world. As Bloom notes, citing Csibra et al. [2003] and Scholl and Tremoulet [2000], these kinds of attributions of agency occur in babies too. And a large literature (for example, Boyer [2007], Atran [2002], see Barrett [2000] for a survey) appeals to this human pull towards agency-based explanations in order to explain the widespread existence of religious thinking among humans.

A common hypothesis is that we have a distinct 'module' for detecting agency — for example, Atran [2002, p. 57] claims that 'the concept of the supernatural agent is culturally derived from innate cognitive schema, "mental modules," for the recognition and interpretation of agents, such as people and animals' — and that this module exists due to the huge evolutionary importance of noticing agents in our environment. Consequently, we 'automatically' form these hypotheses about agents. 'If we, or our ancestors, were to find that the scratching at the door or the howling in the air was not the stalker that seems "automatically" to first come to mind but only the play of branches and leaves in the wind, we would suffer nothing. But if there were a stalker, we would be prepared and likely to suffer less than if we weren't' [Atran, 2002, p. 57].

So, these agent-based explanations of phenomena, and importantly, the phenomena that trigger the appearance of these explanations are, for evolutionary reasons, extremely salient to us. And their salience is immediate, and automatic. Given that, here's the somewhat speculative bit: The salience of these agent-based explanations will affect what we feel needs explanation.

In particular, when we observe a fact that has an appearance of authorship we are immediately and automatically drawn to some agent-based explanation of the fact. This has implications for how we think about the way these facts need explanation.

Consider a case where there is a reasonably strong appearance of authorship: Imagine I'm in bed late at night and I hear something from the living room that sounds like a voice. This sound, and the agent-based explanation of this sound are, unsurprisingly, very salient to me. In this case I have a few (epistemic) options. (i) I could come to believe the agent-based explanation — that there is a person in my living room. (ii) I could come to believe some other explanation of the sound — perhaps my partner's phone is still connected to the bluetooth speaker in the living room and she accidentally

hit play on the podcast she was listening to earlier. (iii) Or I could be agnostic – there are no other signs of there being someone in the living room, and so it seems unlikely that someone is there. But it's hard to see what the explanation of the sound was. In this case, I would be on high attention, looking for what the potential explanation of the sound is.

What isn't a psychologically plausible option is to just ignore this sound and not have it affect your reasoning. It's not plausible that I would, or could, treat this sound in just the same was as lots of other sounds that come from my house; it would be extremely difficult for me to just dismiss this sound as a strange fluke, or coincidence. (Until, of course, some time passes and my attention drifts elsewhere.)

Similar reasoning holds with the types of cases we have been considering in the paper so far. Consider **Darts**, or **Microwave Background Radiation**. In such cases the phenomenon in question and the appearance of agency or authorship are very salient. And we have a few epistemic options.

(i) We could come to believe the agent-based explanation — that in **Darts** some agent is playing a trick on me and controlling what numbers the darts hit, and in **Microwave Background Radiation** some, perhaps divine, agent made it such that the temperature of the radiation is uniform. People taking this option about a variety of natural phenomena is, in effect, the genesis of religious thinking, according to Boyer [2007], Atran [2002], Guthrie [1995] and others.

(ii) We could accept some other explanation of the phenomenon. For example, as we noted earlier, modern physicists typically take the uniformity of the temperate of the microwave background radiation to be reason to accept *inflationary* big bang theories which can give a robust explanation of the uniformity.

(iii) We could be agnostic. I take it that this was the state of physics before the development of inflationary theories where there was attention to the issue of the uniformity of temperature active work looking for what the explanation could be.

Similarly, in these cases, it's psychologically implausible for these facts — the uniformity of the temperature, or the numbers hit on the dartboard spelling out the birthdays of your family — to not affect your reasoning. When we have this very salient phenomenon then it's psychologically difficult to treat it in the same way as other facts that you are happy dismiss as not needing explanation. And this is for same reason as it's psychologically implausible to dismiss the human-like sounds you heard late at night. Of course, when you hear the sounds there's a concern that your physical safety is at risk, which is not present in **Darts** or **Microwave Background Radiation**. But, as our discussion of the psychological literature suggested, the processes of agency detection which are so sensitive because of situations where our safety is at risk apply much more widely.

In these cases, then, and other cases where there is an appearance of authorship, the huge salience of

agency explanation affects our feeling of what needs to be robustly explained and what we are happy to take as flukes. This consideration helps to answer the questions that we started the section with — it makes less puzzling that there could be a connection between the feeling of authorship and the need for explanation.

Of course, this discussion is somewhat speculative. But the way in which we are automatically pulled towards agency-based explanations and the salience of the facts that suggest these explanations is not speculative — we've pointed towards some of the literature on these facts. And if Boyer [2007], Atran [2002], and others are right that this feature of our psychology is what's ultimately responsible for the religious thinking that's hugely influential across a very wide variety of cultures, then it's not surprising that it might have some effects on our scientific and everyday reasoning about explanation.

4.2 Two Worries

Here are two very natural reactions that a reader might have at this point.

Firstly, earlier I was clear that strikingness is an *epistemic* phenomenon. If a theory implies that a striking fact doesn't have a robust explanation then that's a problem for the theory. But the discussion of the last section was all about psychology. What does any of that have to do with *epistemology*? How does the discussion of the appearance of authorship and agency detection give us any evidence about what's likely to be true, or about what we should believe?

Secondly, isn't this view just crazy? Isn't it just extremely strange to think that facts about the appearance of authorship have implications for theory choice? Isn't it implausible that scientists favor the inflationary version of Big Bang theory because of the way in which the uniformity of the CMB appears authored to them?

Let's take the second concern first. Start by noting that the view doesn't imply that physicists have to have any explicit beliefs about authorship, and it certainly doesn't imply that when a physicist is arguing against standard Big Bang theory they need to discuss the appearance of authorship. Rather, when physicists argue against standard Big Bang theory they say, as we noted in section 2, that, for example, the uniformity of the CMB is 'disturbing' and 'a profound problem' [Dodelson, 2003, p.143] or is 'remarkable' [Ryden, 2017, section 10.2] and 'puzzling' [Guth, 1981] or, most often, they simply assume without discussion that the uniformity must have a robust explanation.[12] It is these feelings of discomfort — of the phenomenon being disturbing and puzzling — that are the real scientific considerations that lead to physicists rejecting standard Big Bang theory. The discussion of the appearance of authorship allows us to make sense of this discomfort.

[12] Again, see McCoy [2015, p. 28] for further quotes of this kind.

But, even so, perhaps it seems crazy that the appearance of authorship plays any role at all. If the diagnosis of these scientific judgments is ultimately to do with the uniformity of the CMB having an appearance of authorship then isn't that just reason to think that scientists are being irrational and are making a mistake?

And, this connects up to the first concern — perhaps it is psychologically plausible that the appearance of authorship affects what scientists, and the rest of us, feel needs explanation, but what does this have to do with what we should believe?

One natural response to these considerations is to take the discussion of this paper, and the authorship view in particular, as debunking judgements of strikingness. Sure, there are these compelling intuitions of strikingness. But once we see that they derive from the same source as our automatically forming the hypothesis that there is a stalker there where we hear scratching at the door or the howling in the air then we should understand that those intuitions are unreliable and should be put aside. It's just a mistake for scientists to reason in this way, though an understandable one, given the psychological salience of agent-based explanations.

If this is your reaction, then that's perfectly reasonable — I'm happy with the argument of this paper being taken as a debunking of strikingness intuitions.

But I'm inclined to say something different. I'm inclined to say that is really is rational for scientists to reason in the way that they did when they rejected standard Big Bang theory – this important and widely accepted part of the practice of cosmology is not just a mistake. And more generally, it's too revisionary of science to say that these judgements of strikingness should be rejected — the uniform distribution of temperature really does call out for explanation more than some random-looking distribution!

So again, the question arises — what does all the discussion of psychology and the appearance of authorship have to do with epistemology? This is a very forceful worry.

In fact, I think similar worries can be raised for other considerations that are relevant to theory choice, like simplicity. It's natural to be concerned that at least some considerations of simplicity seem more psychologically important than epistemically important. What does simplicity have to do with what's likely to be true? Aren't (at least some) considerations of simplicity more about us than the world outside us?[13]

[13] I don't mean to ignore the large literature that there is on simplicity as a theoretical virtue – especially focused on questions surrounding *Occam's Razor* and the so-called *curve-fitting problem* – that does aim to address questions like these. A huge amount of progress has been made about various ways we might formulate various aspects of simplicity and what assumptions we would need to justify their use in theory choice — especially in the way that methods from statistics have been brought to bear on these questions. (Sober [2015], esp. Ch. 2, is a very good overview of this work.) I certainly can't do justice to this literature here. But, even ignoring potential worries with these approaches, they don't seem

So I don't think this issue is distinctive to strikingness. Rather gets at a more general point about the role theoretical virtues in science.

It's hard to know how to square the way in which these theoretical virtues — simplicity, and strikingness (and perhaps others, like commonsensicalness [Emery, 2017]) — seem to be largely driven by facts about our psychology with their role in theory choice in science, and in everyday life. But the option I'm most inclined towards is that the role of things like strikingness and simplicity in our theory choice is evidence for certain views of epistemic normativity — views where facts about psychology play a role in grounding facts about what we ought to believe. Work in metaethics has suggested a vast range of views of this kind. For example, various different kinds of relativism fit the bill, as does subjectivism, and certain kinds of constructivism, and views where the normatively relevant properties are response-dependent secondary qualities. There are many other possibilities as well.

Further, if we accept such a view of epistemic normativity then the view might start to seem a bit less crazy — given such a view it is perhaps somewhat natural to think that these extremely deep features of our psychology, to do with agent-detection, do have a role to play in epistemic normativity.

Of course, fully developing such a view of epistemic normativity is not a task for this paper. And accepting such a view is a very big commitment. So again, I think that a perfectly reasonable alternative option is to take the authorship view as debunking strikingness intuitions, although that view comes with real costs.

5 Conclusion

I've argued that strikingness has a very important role to play in theory choice — we need to understand strikingness if we are to understand our scientific practice — and that it's not easy to see how to see how to understand the role of strikingness in terms of the other epistemic principles that we accept. What we need, then, is an approach to strikingness which understands it as a novel theoretical virtue. The view that I suggested, that striking facts are those that have the appearance of authorship is a view of that kind – one that I think fits well with facts about our psychology and captures our intuitions about strikingness well. But strikingness is a deeply underexplored phenomenon, and that

to apply to some aspects of simplicity that are relevant for theory choice. For example, Baker [2016] describes the case of Mendel's experiments with growing hybrid varieties of peas by cross-pollinating varieties with different traits. 'Across his experiments with seven different such traits, the ratio of dominant trait to recessive trait averaged 2.98:1. On this basis, Mendel hypothesized that the true ratio is 3:1. This "rounding" was made prior to the formulation of any explanatory model, hence it cannot have been driven by any theory-specific consideration.' It seems perfectly reasonable for Mendel to do this, but we don't have an explanation for what this type of simplicity consideration has to do with what's likely to be true.

true even in light of this discussion — there is still a huge amount of investigation to do. Hopefully, the discussion here opens up pathways for that investigation.

References

David Armstrong. *What is a Law of Nature?* CUP, 1983.

Scott Atran. *In gods we trust: The evolutionary landscape of religion*. Oxford University Press, 2002.

Alan Baker. Simplicity. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016.

Dan Baras. Why Do Certain States of Affairs Call Out for Explanation? A Critique of Two Horwichian Accounts. *Philosophia*, pages 1–15, fort.a.

Dan Baras. A Strike Against a Striking Principle. *Philosophical Studies*, pages 1–14, fort.b.

Justin L Barrett. Exploring the natural foundations of religion. *Trends in cognitive sciences*, 4(1): 29–34, 2000.

Helen Beebee. Necessary Connections and the Problem of Induction. *Nous*, 45(3):504–527, 2011.

Harjit Bhogal. Coincidences and the grain of explanation. *Philosophy and Phenomenological Research*, 100(3):677–694, 2020. doi: 10.1111/phpr.12570.

Paul Bloom. Religion is natural. *Developmental science*, 10(1):147–151, 2007.

Nick Bostrom. *Anthropic Bias: Observation Selection Effects in Science and Philosophy*. Routledge, 2002.

Pascal Boyer. *Religion explained: The evolutionary origins of religious thought*. Basic books, 2007.

Craig Callender and Jonathan Cohen. Special Sciences, Conspiracy and the Better Best System Account of Lawhood. *Erkenntnis*, 2010.

Gergely Csibra, Szilvia Bíró, Orsolya Koós, and György Gergely. One-year-old infants use teleological representations of actions productively. *Cognitive Science*, 27(1):111–133, 2003.

Scott Dodelson. *Modern cosmology*. Elsevier, 2003.

John Earman and Jesus Mosterin. A critical look at inflationary cosmology. *Philosophy of Science*, 66 (1):1–49, 1999.

Adam Elga. Infinitesimal chances and the laws of nature. *Lewisian themes: the philosophy of David K. Lewis*, 2004.

Nina Emery. Against radical quantum ontologies. *Philosophy and Phenomenological Research*, 95(3): 564–591, 2017.

Hartry Field. *Realism, Mathematics and Modality*. Realism, Mathematics and Modality. Blackwell, Oxford, 1989.

Stephen R. Grimm. Explanatory inquiry and the need for explanation. *British Journal for the Philosophy of Science*, 59(3):481–497, 2008.

Alan Guth and Paul Steinhardt. The inflationary universe. In P. Davies, editor, *The New Physics*. 1989.

Alan H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23:347–356, Jan 1981.

Stewart Guthrie. *Faces in the clouds: A new theory of religion*. Oxford University Press on Demand, 1995.

Fritz Heider and Marianne Simmel. An experimental study of apparent behavior. *The American journal of psychology*, 57(2):243–259, 1944.

Paul Horwich. *Probability and Evidence*. Cambridge University Press, 1982.

Peter Lipton. *Inference to the Best Explanation*, volume 102. Routledge/Taylor and Francis Group, 2004.

Tim Maudlin. *The Metaphysics within Physics*. OUP, 2007.

Casey D McCoy. Does inflation solve the hot big bang model's fine-tuning problems? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 51: 23–36, 2015.

Wayne Myrvold, Marco Genovese, and Abner Shimony. Bell's theorem. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2019 edition, 2019.

John T. Roberts. Fine-tuning and the infrared bull's-eye. *Philosophical Studies*, 160(2):287–303, 2012. doi: 10.1007/s11098-011-9719-0.

Barbara Ryden. *Introduction to cosmology*. Cambridge University Press, 2017.

George N. Schlesinger. Miracles and probabilities. *Noûs*, 21(2):219–232, 1987.

Brian J Scholl and Patrice D Tremoulet. Perceptual causality and animacy. *Trends in cognitive sciences*, 4(8):299–309, 2000.

Elliott Sober. *Ockham's Razors: A User's Manual*. Cambridge University Press, 2015.

Galen Strawson. *The Secret Connexion: Causation, Realism, and David Hume*. Oxford University Press, 1989.

Michael Strevens. *Depth: An Account of Scientific Explanation*. Harvard University Press, 2008.

Brad Weslake. Explanatory depth. *Philosophy of Science*, 77(2):273–294, 2010.

Roger White. Explanation as a guide to induction. *Philosophers' Imprint*, 5:1–29, 2005.

Roger White. Does origins of life research rest on a mistake? *Noûs*, 41(3):453–477, 2007.

Robert Williams. Chances, counterfactuals, and similarity. *Philosophy and Phenomenological Research*, 77(2):385–420, 2008.

Robert A Wilson. Causal Depth, Theoretical Appropriateness, and Individualism in Psychology. *Philosophy of Science*, 61(1):55–75, 1994.

James Woodward. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, 2003.

Stephen Yablo. Mental Causation. *The Philosophical Review*, 101(2):245–280, 1992.

Zanja Yudell and Wai-Hung Wong. A normative account of the need for explanation. *Synthese*, 192 (9):2863–2885, 2015.