

DIFFERENCE-MAKING AND DETERMINISTIC CHANCE^{*}

Harjit Bhogal

Forthcoming in *Philosophical Studies*

Abstract

Why do we value higher-level scientific explanations if, ultimately, the world is physical? An attractive answer is that physical explanations often cite facts that don't make a difference to the event in question. I claim that to properly develop this view we need to commit to a type of deterministic chance. And in doing so, we see the theoretical utility of deterministic chance, giving us reason to accept a package of views including deterministic chance.

There's a natural thought, expressed by, for example, by ?, p. 118 that determinism is incompatible with there being non-trivial objective chances (that is, with chances that have values greater than 0 and less than 1). The intuition is that if determinism is true, then every fact about the future is guaranteed by the current state of the world. And if some future fact — some coin toss landing heads, for example — is guaranteed to occur given the current state, then it's natural to think that it must currently have a chance of 1.

But, particularly over the past 10 years or so, many philosophers have argued in favor of deterministic chance. Those arguments typically work in one of two different ways. Either they argue that deterministic chance is consistent with the generally accepted platitudes about chance (at least when

^{*}Thanks to Rosa Cao, Laura Franklin-Hall, Ben Holguin, Barry Loewer, Tim Maudlin, Sam Scheffler, Brad Weslake, Mike Zhao and to audiences at the BSPS annual conference, the University of Maryland, and NYU Thesis Prep. Special thanks to Michael Strevens.

those platitudes are properly understood) so deterministic chance is in principle possible (see, for example, ? making this case, arguing contrary to ?). Or they argue that accepting deterministic chance has an important theoretical payoff — that it allows us to make sense of some important scientific or everyday phenomenon. And that gives us a reason to accept a package of views that includes committing to deterministic chance. For example, ? argues that deterministic chance allows us to make sense of our practice of asserting counterfactuals in light of concerns about counterfactual skepticism. ? argues that countenancing deterministic chance makes sense of our practices of experimental confirmation in science. And many people, notably ? and ?, argue that we need deterministic chance in order to make sense of statistical mechanics.

I'm going to make an argument of this second kind. I'll argue that accepting a kind of deterministic chance allows us to develop a plausible account of the value of higher-level, special science explanations — those explanations given by economics, or biology, or sociology, for example — in a physical world. In particular, I'll argue that the very attractive *difference-making* approach to the question of levels of explanation needs deterministic chance to be developed properly. (I'm going to assume determinism in everything to come.)

A couple of important points to note about this second kind of argument. Firstly, an argument of this second kind is open to the criticism that, whatever probability is useful in making sense of the higher-level explanations, it's not chance. Rather, it's some other type of objective probability that doesn't deserve the name 'chance'. I'm not going to argue much about the name 'chance' here. If, for example, someone agrees with ? that nothing in a deterministic world could possibly count as chance then they can construe the argument as an argument for an important type of objective probability in a deterministic world.

Secondly, arguments of this second kind are not conclusive. If I argue that my account, which incorporates deterministic chance, makes sense of the value of higher-level explanations then it is open to someone else to give an account that does the job without deterministic chance. But the attractiveness and plausibility of the account I give, or of similar accounts, does provide reason to

accept a package of views including deterministic chance. I'll talk about this objection more at the end of section 7.

One final qualification before we start. Because this topic connects the big literature about deterministic chance with the vast literature about levels of explanation a full review of the literature is impossible. But I will, along the way, point to how my account connects to previous ones. But I don't pretend to have conclusively refuted all the views I put aside, or even to have identified all the possible options.

In section 1 I'll motivate the question of levels of explanation and introduce the difference-making approach to the question. In section 2 I'll raise a problem for the difference-making approach as it stands, and suggest that the way to solve it is to give an account of graded difference-making. In sections 3, 4, and 5, I develop an account of explanatory goodness that captures the idea of graded difference-making. In section 6 I'll argue that this account allows us to give a good answer to questions about levels of explanation. And in section 7 I'll show how dealing with the question of levels of explanation in this way requires us to countenance a type of deterministic chance.

1 LEVELS OF EXPLANATION

We live, or so I assume, in a physical world. The basic constituents of the world are physical. Every phenomenon in the world is, ultimately, a physical phenomenon. However, there are some phenomena that we typically explain without explicit reference to physics. Instead we explain them in sociological, or economic, or biological terms and not in physical terms. On the face of it, this is puzzling. If every phenomenon is ultimately physical, why do we explain many phenomena in non-physical terms?

(In asking this question, and in what's to come, I'm thinking of explanation as a relation between facts that holds out there in the world. An explanation is the holding of this relation – it to be sharply distinguished from *acts of explanation* which are speech acts or mental states. So, in this sense, there

can be explanations, even if we don't, or can't know about them. As ? notes, this is the common usage of the term in the literature on explanation.)

One very natural thought is that pragmatic considerations provide the answer here. We just are not capable of enquiring into, for example, the economic systems that we are interested in by using physical vocabulary. The economic phenomena are so complicated when understood in physical terms that epistemic and computational limitations prevent us from, for example, doing the physics of exchange-rates.

There is clearly something to this. But it can't be the whole story – certain high-level explanations seem preferable to even the physical explanations we would give if the epistemic and computational limitations were relaxed. Consider, for example, the discovery of eighteenth century doctor John Arbuthnot that more males had been born than females in London in each of the last 82 years. There is a possible microphysical explanation of this fact, one which lays out separately all of the physical details which led to each male or female birth. But this does not seem like a good explanation.

Kitcher (2001, pp. 70-72) argues that 'even if we had this "explanation" to hand, and could assimilate the details, it would still not advance our understanding'. On the other hand, an evolutionary biological explanation does much better. The argument that there is an evolutionary tendency to produce a 1:1 sex ratio at the time of sexual maturity along with the fact that males are more prone to infant mortality than females, explains why there is evolutionary pressure for more males to be born than females. This evolutionary explanation is pitched at the right level, the microphysical explanation is not.

Of course, just this thought doesn't conclusively refute the pragmatic approach. But, I think, we do have reason to look for an account which makes sense of the way in which some higher-level explanations can have an objective, non-pragmatic, advantage over competing lower-level ones – an advantage that isn't merely to do with our weakness as reasoners, or enquirers, or communicators. That is what I'm going to look for in the paper to come.

Another natural thought is that higher-level explanations are sometimes preferable to physical expla-

nations because higher-level explanations somehow carry more information than competing physical explanations.¹ However, as ? argues, if we are physicalists (which, again, for the purposes of this paper, we are) then it cannot be the case that there is information carried by higher-level explanations which is not present, so to speak, at the physical level. So, the informational account will not be able to favor a higher-level explanation over some, sufficiently informative, lower-level explanation. Again, there's more to say about this argument. In particular, we have to be careful about how to formulate this sufficiently informative lower-level explanation (especially in light of worries raised by ?, section 2). But the informational account seems to have difficulties in showing why we favor higher-level explanations *in a physical world*.

So, there's a hard problem of working out why the right 'level of explanation' for some explananda is not the physical level. In fact, ?, p. 3 has described a solution to the problem as 'a kind of Holy Grail in the philosophy of science, long sought but never found'. But this is just an instance of a more general problem: Why is the right level of explanation what it is? Why, for a particular explanandum, should we explain it at the economic level, or the biological level, or the sociological level?

Given the rejection of the idea that high-level explanations carry information that lower-level explanations do not it looks like we need some story on which high-level explanations are sometimes favorable to lower-level explanations *even given*, or perhaps *because*, they carry less information. In fact, this idea that higher-level explanations are better because they, in some sense, carry less information is fairly common. Perhaps the most popular version of this idea appeals to the notion of difference-making (see, in particular, ? and ?).²

The central idea of the difference-making approach is that explanations should cite all and only the things that make a difference to the truth of the explanandum. So, for example, the microphysical

¹? influentially develop a view of this kind. But this approach is also suggested by ?, ? and ?.

²There are other versions of this idea present in the literature. For example, ? gives an 'informational economy account' that favors less informative explanations, though, unlike the difference-making account, it doesn't come with a story about why less informative explanations should be favored. ?, section 3 also favors less informative explanations, but his story is based on Gricean norms about the pragmatics of conversation, and as such the account is probably better classified as a version of the pragmatic approach mentioned above. Though clearly, a fair consideration of these accounts would take much, much longer.

explanation of Arbuthnot's regularity — the fact that more males than females were born every year in London — that lays out separately all of the physical details which led to each male or female birth, contains *too much information*. In particular, it contains information that *does not make a difference* to the explanandum.

There are different conceptions of exactly what makes something a difference-maker, but the intuitive thought is that a fact A makes a difference to B when A is relevant to whether or not B holds. One natural way to develop this thought is by saying that for two facts, A and B, A makes a difference for B if were not-A to hold then not-B would hold. Arbuthnot's regularity counterfactually depends upon the fact that there are higher rates of infant mortality in males and so, on this conception, that fact counts as a difference-maker. The precise microphysical details of any particular birth, on the other hand, do not count as a difference-maker since if the details of any particular birth were different there would still be more males born than females. So we can start to see how the difference-making approach could favor higher-level explanations.

We shouldn't stick too closely to the conception of difference-making as counterfactual dependence — it is ultimately untenable for reasons closely related to the problems that face accounts of *causation* as counterfactual dependence (see ?, section 2.4). But it gives us the flavor of a difference-making approach.

The difference-making account is, I think, extremely promising. The idea that explanations should cite the things that are relevant to whether the explanandum holds is very intuitive — after all not just any information is explanatory, rather the information needs to be targeted in a certain way at the explanandum. And it gives the start of a plausible answer to why higher-level explanations are sometimes to be preferred over lower-level explanations. In the rest of the paper I'm going to focus on the difference-making account — arguing that we can develop it to make sense of why we sometimes favor higher-level explanations to lower-level ones, and in doing so we commit to a certain type of deterministic chance.³

³There are a vast number of accounts of the value of higher-level explanations that I am putting aside. For example, there are approaches that emphasize how certain high-level explanations are appropriate answers to certain 'why-questions'

2 DIFFERENCE-MAKING AND LOW-LEVEL DIFFERENCE-MAKERS

Again, the central idea of the difference-making approach is that explanations should cite all and only the things that make a difference to the truth of the explanandum. As we noted, this looks like a promising approach. However, the difference-making approach faces a problem. Put simply, special science explanations typically ignore lots of difference-makers — once we have given the special science explanation there are still additional physical facts that make a difference to the occurrence of the explanandum. This creates a problem for the difference-making approach.

The best way to see this is by focusing on an influential version of the difference-making approach given by ?. (Though ? faces a very similar problem.)

Strevens' account of explanation starts with a set of low-level physical facts, including physical laws, that entail the phenomenon under investigation.⁴ Roughly, the idea is to remove information from the premises piece by piece while still retaining the entailment of the phenomenon. For example, if the original premises included the facts that a particular object had mass of 10g and was moving at 2m/s then we might replace these facts with one stating that the object had momentum of over 15gm/s, if that fact is enough for the entailment to still go through. When the entailment has been abstracted as far as it can go then the premises that remain count as difference-makers and explain the phenomenon.

However, the demand that the explanans entails the explanandum means that the account has difficulties with explanations that are sufficiently high-level. Lots of higher-level explanations are such that the explanans does not entail the explanandum — there are ways that the explanans could be true that do not lead to the explanandum.

Imagine, for example, I have some peas and I explain why they are smooth and not wrinkled by

and not to others (e.g. ?, p.129-130, ?). There are accounts that focus on the generality of some law, or similar principle, that is part of the explanation (e.g. ?, section 2, ?). There are those that appeal to the institutional structure of science the corresponding division of scientific labor (e.g. ?, ?), and there are other approaches. I think there are reasons to favor the difference-making approach over these accounts, but clearly, these accounts cannot be properly investigated here.

⁴There has to be a restriction on the type of entailment here in order to avoid derivations — like that of the height of the flagpole from the length of the shadow — that are obviously not relevant to explanation.

noting that they were bred from parent plants that were homozygous in the allele associated with smoothness. This is a good classical genetics explanation. But the explanans does not necessitate, and so does not entail, the explanandum — there are cases, like those where mutations occur, where the peas are bred from parents that are homozygous in the relevant allele but the peas end up wrinkled and not smooth. The putative explanation ignores certain difference-makers, like the physical facts that determine that a mutation did not occur in this case.

We can easily multiply cases. Consider, for example, the explanation of a particular person getting the job from the fact that they performed best at the interview. Or the explanation of the formation of a volcanic arc from the fact that one tectonic plate subducts under another leading to cracks in the crust of the upper plate. Or the explanation that dollar declined in value because the central bank decided to cut interest rates (thus leading to fewer investors putting their money in US bank accounts). In all of these explanations the explanans doesn't necessitate the explanandum.⁵ We need to say more then, I think, to see how Strevens' approach can be squared with these good special science explanations.

Notice that this isn't just an issue for Strevens — it affects difference-making accounts more generally. Take any explanation where the explanans doesn't physically necessitate the explanandum. Then there are physical facts that make a difference to whether we are in one of the cases where the explanans does lead to the explanandum, or one of the cases where it does not. And so, those physical facts make a difference to the explanandum.

Any explanation, then, where the explanans doesn't physically necessitate the explanandum leaves out physical facts that are difference-makers. But, almost every explanation we actually see in the special sciences does not have the explanans physically necessitate the explanandum.⁶ It seems like good explanations often ignore lower-level difference-makers.

⁵ Strevens does go on to add further nuances to his account, allowing that in some situations explanations need not entail the explanandum — for example, he develops a theory of probabilistic explanation. But none of these seem to help in the cases under consideration.

⁶ We would have to change the form of this argument very slightly if the focus of the discussion was Woodward's (?) account of *proportionality* since his view takes *variables* to be the relata of explanations. But this is a simple fix.

However, the core idea of the difference-making approach still seems compelling – we want explanations that leave out that which is irrelevant to the explanandum, and it does seem like microphysical explanations of higher-level facts include things that don't make a difference. How can we keep this while making sense of the way in which good special science explanations leave out difference-makers?

Here's a very natural suggestion: Just relax the condition that explanations need to cite *all* difference-makers.⁷

That sounds good, but it's not enough to stop there. The key question is: If we don't need to cite all difference-makers which ones do we need to cite? The most obvious answer is that we should cite the facts that make *enough* of a difference. So the defender of the difference-making approach might say 'Sure, low-level facts make a difference – low-level physical facts do make a difference to phenomena like the peas being smooth, for example, but they don't make *much* of difference, so it is reasonable to ignore them in giving our explanations.' Or alternatively, they might say, 'Sure, low-level facts *sometimes* make a difference, but they don't make a difference *most of the time*, so it is reasonable to ignore them in giving our explanations.'

Such thoughts seem intuitive, but on reflection is hard to know what to make of them. Let's start with the second thought. It is not true that the physical facts only sometimes make a difference to the smoothness of the peas. The facts that determine whether there is a mutation always make a difference — sometimes they lead to the peas being smooth and sometimes they lead to the peas being wrinkled, but, either way, it makes a difference that the physical facts are one way rather than another.

Now take the first thought. This also seems misguided, at least on its most obvious interpretation. It seems false that the physical facts that determine whether there is a mutation don't make much of a difference to the smoothness of the peas. They make the difference between the peas being smooth

⁷Another way to put the suggestion: Add a *ceteris paribus* clause. So, for example we could say that *ceteris paribus* the fact that the peas were bred from parent plants that were homozygous in the allele associated with smoothness explains their smoothness. The idea being that we push some of the difference-makers — like the fact that there was no mutation — into the *ceteris paribus* clause. This is just another way of saying that we don't need to explicitly cite all the difference-makers.

and them not being smooth, surely this is a substantial difference.

So, is there no reasonable way of developing this idea that explanations should cite things which make enough of a difference? Should we just give up on difference-making providing the answer to the question about levels? I think the answer to both questions is no. We can still develop the idea that explanations should cite things which make enough of a difference and defend the spirit of the difference-making approach.

To do this, we need a graded conception of difference-making — developing this, as we have just seen, is tricky — and some sense of what makes for *enough* of a difference. What this is going to require is a kind of deterministic chance — a measure over physically possible worlds. In the rest of the paper I'm going to show how accepting this kind of deterministic chance allows us to give a good account of the value of higher-level, special science, explanations — and account that is a development of the difference-making approach.

But before I actually give the account, there is a little bit of groundwork to lay.

3 EXPLANATORY GOODNESS AND EXPLANATORY CORRECTNESS

Let's start by distinguishing between *explanatory goodness* and *explanatory correctness*. An account of explanatory correctness is just an account of when something counts as an explanation and when it is not. An account of explanatory goodness is an account of when explanations are better or worse.

A difference-making component to explanation could be built into an account of explanatory correctness, as ? does, or added as an additional virtue or good-making feature of an explanation, as ? does. Since I am developing a sense in which difference-making can be graded, it's natural for me to take the latter strategy, since goodness can be naturally graded while correctness cannot.

So, I'm going to develop my account as an account of dimensions of explanatory goodness — ways in which explanations can be better or worse. We will see how these dimensions of explanatory goodness capture difference-making intuitions and answer the question about levels of explanation.

But in developing the account we are going to need the notion of explanatory correctness too. We won't need any one specific account, just the generic notion of explanatory correctness will be enough. But for definiteness, and ease of presentation, I'm going to fix on a simple causal account of explanation, given by ?. The account says that A explains E's occurrence if and only if A gives information about the network of causal relations that leads to E.

We should stop here to note one important constraint that was perhaps implicit in Lewis's account but it will be useful to make explicit: Very unnatural or disjunctive properties cannot be part of the explanans of correct explanations. For example, imagine we tried to explain the decline in the value of the dollar by appealing to the deterministic fundamental laws of nature, and the fact *the world was in A_1 or A_2 ...or A_n at its initial time*, where A_1 through A_n are all and only the possible initial conditions that would, along with the deterministic laws, entail the decline in the dollar. This, I take it, is not acceptable as an explanation, because of its disjunctive character. Neither would an explanation that introduces a single property, F, that replicates the disjunction by applying to all and only cases where A_1 through A_n hold.

Note that ruling out these unnatural or disjunctive properties is something that everyone who wants to understand high-level explanations has to do. Some properties, the natural or projectable ones, can be used in other explanations, while the unnatural, disjunctive, or gerrymandered properties cannot. I'm not going to give an account of high-level naturalness here though, that's a big project that I take on elsewhere.

Whether or not you like this causal account of explanatory correctness doesn't really matter because, as we will see more clearly later, the account of explanatory goodness is not committed to it. (It is committed to the claim that unnatural properties can't explain though.) But there is a dialectical reason for using this account: It's a very minimal account — it sets a very low bar for something to count as an explanation. Consequently, using this account of correctness makes very clear that the all the interesting work for identifying the level of explanation is done by the account of explanatory goodness.

4 ROBUSTNESS

Again, the aim is to develop an account driven by the thought that explanations should cite the things that make *enough* of a difference to the explanandum, and then to use this to answer the question about levels of explanation. And to do that we need to start by developing a graded conception of difference-making. As we saw in section 2, the obvious ways to do this run into problems.

But there is, I think, a way to make sense of the idea that some facts can make more of a difference and others less. To see this, consider a different example. Imagine that you drop an ice cube in warm water and the ice cube melts. You can explain the ice melting by citing the fact that it was dropped into warm water. Alternatively, you could (at least in principle) explain the melting of the ice cube by citing the precise microstate, call it *M*, of the ice-cube water system and showing how the physical laws lead from that to the melting. Something seems problematic about this second explanation though — it appears to be pitched at the wrong level, just as, for example, the microphysical explanation of Arbuthnot's regularity is pitched at the wrong level.

When such cases are discussed in the literature on statistical mechanics it is heavily stressed that most, in fact, nearly all, of the possible microstates of the ice-cube water system would lead to the ice melting when it is dropped in the water. Though not *all* microstates — there are some possible microstates that would lead to the ice cube *growing* when it is dropped in warm water, but those microstates are incredibly rare and unlikely.

The fact that the overwhelming majority of the microstates lead to the melting, it is often thought, is a reason why it is reasonable to appeal to the higher-level explanation that just cites the ice being dropped in the water rather than having to appeal to the precise microstate (see, for example, ?, p. 1117, ?, pp. 150-151, (?, pp. 309-318)). I think this is right — the fact that most of the microstates lead to the melting tells us about why it is bad to give the microphysical explanation that cites *M*.

This is because there is a sense in which the particular microstate *M* which holds doesn't make much of a difference to the melting of the ice cube, because *if almost any other microstate held the ice cube would still have melted* after being dropped in the water. To put it another way, *M*'s holding is not

required or *close to required* for the melting of the ice cube. If M had not held then the melting would almost certainly have still occurred.

We can generalize these ideas into a measure of graded difference-making. In particular, I propose a dimension of explanatory goodness called ROBUSTNESS. The central idea of ROBUSTNESS is that an explanation of B from A is better if more (in the sense of a higher proportion) of the ways that B could hold are such that, in that situation, A explains B.⁸

We can translate this thought into the language of possible worlds to give the official formulation of ROBUSTNESS: An explanation of B from A is better if that explanation holds in more, that is, a higher-proportion, of the physically possible worlds where B holds.⁹

So, we can see that the microphysical explanation of the ice melting scores badly on ROBUSTNESS since there are many physically possible worlds where the ice melts but in very few of them is the melting explained by the holding of the particular microstate M. In fact, in only very few of the physically possible worlds where the ice melts does M hold. In the overwhelming majority of physically possible worlds where the ice melts this microphysical explanation does not apply; for most (in fact, nearly all) of the ways that the melting could hold, M does not explain the melting.

The higher-level explanation which just cites the fact that the ice cube was dropped in warm water to explain the melting scores much better on ROBUSTNESS — lots of the worlds where the ice melts are such that the explanation is that the ice was dropped in warm water.

Similarly, take the microphysical explanation of the decline in the dollar — one that cites the movement of all the particles that make up the economic system. This explanation scores badly on ROBUSTNESS because the explanans of that microphysical explanation holds in very few worlds, so the explanation must also hold in few worlds. In the overwhelming majority of physically possible worlds where the dollar declines in value this microphysical explanation does not apply. The economic ex-

⁸ROBUSTNESS (and PRECISION which will be discussed in the next section) were stated and used, though not fully developed or defended (and not connected to issues of difference-making) in ?, pp. 12-13. The discussion of the rest of this section, and the next, develops the discussion there. (Or, perhaps it's more accurate to say that the discussion there follows the discussion here.)

⁹This dimension is similar to the account of 'explanatory depth' given by ?.

planation does much better on ROBUSTNESS — lots of the worlds where the dollar declines in value are such that the explanation is that interest rates declined. More generally, ROBUSTNESS will tend to disfavor explanations that are very specific, because those explanations can only hold in few worlds.¹⁰

Again, we can see the intuitive force of ROBUSTNESS by seeing how it captures a graded sense of difference-making. Consider an explanation that scores well on ROBUSTNESS. If A explains B and this explanation scores very highly on ROBUSTNESS then A is close to required for B. This is because there are few situations where B occurs without A — in most of the physically possible worlds where B occurs it is explained by A, so there are only few physically possible worlds where B occurs and A doesn't.

To put it another way, if A does not hold then it will be rare that B holds since in nearly all of the cases where B holds it is explained by A. A, then, makes a substantial difference to whether B occurs.

On the other hand, if A explains B and this scores badly on ROBUSTNESS then A is not close to required for the occurrence of B. If the explanation scores badly on ROBUSTNESS, typically, this is going to mean that A holds in few of the worlds where the B holds — there are many physically possible worlds where B occurs without A.¹¹

It is rare for A to be part of the cause that leads to B, most of the time when B occurs, it is caused by something else. In this sense, A does not make much of a difference to B. The microphysical explanation of the ice melting is a good example of this — it is rare for the specific microstate M to be part of the cause of the melting. M is not required, or close to required, for the melting — in this sense it doesn't make much of a difference to the melting.

¹⁰Of course, the good economic explanation here doesn't score maximally, or close to maximally on ROBUSTNESS, since there are other economic reasons that the dollar could decline, other than a fall in interest rates. But still, it's clear that it scores much better on ROBUSTNESS than the microphysical explanation. More generally, good explanations that we see in the special sciences will typically not score close to maximally on ROBUSTNESS, but will score much better than explanations that are pitched at too low a level.

¹¹This is only true 'typically' because it is possible that the explanans holds in a similar (or wider) range of worlds as the explanandum, but only explains the explanandum in a few of these worlds. In this case, the explanation would score low on ROBUSTNESS. Even in such cases, though, I think it's correct to say that the explanans is not required or close to required for the explanandum, in the sense of 'required' that is natural in discussions of explanation.

However, given the extremely minimal account of explanatory correctness we are working with, these cases will be very rare so, for the most part, I will ignore them in what's to come. (Though see footnote 14 for a closely related point about the next dimension of explanatory goodness.)

Accepting ROBUSTNESS as a dimension of explanatory goodness is a way of cashing out the difference-making intuition — it implies that explanations are better, at least along one dimension, if they cite facts that make more of a difference.

So we have a graded notion of difference-making. This is part of what we needed to fix up the difference-making account and so to answer the question of levels of explanation. But this isn't all we needed. Our aim was to develop the idea that explanations should cite the things that make enough of a difference. So, now we need some grip on what is 'enough'. After that we will come back to look at the connection between this account and deterministic chance.

5 PRECISION

One option is to just have a threshold for ROBUSTNESS — we demand that explanations should score highly enough on ROBUSTNESS and so make enough of a difference. But lots of clearly bad explanations meet this criteria.

To see this note that ROBUSTNESS tends to favor less specific explanations over more specific ones. Consequently, lots of explanations can score well on ROBUSTNESS in virtue of being very unspecific. If A is part of the causal nexus that leads to B and holds in many of the worlds where B holds then the explanation of A from B will score well on ROBUSTNESS. Consider such an explanation of Arbuthnot's regularity — the fact that more males than females were born in London every year for 82 years in the eighteenth century. It is part of the causal nexus that led up to this fact that there were humans in London at the start of the eighteenth century. And in most of the possible cases where Arbuthnot's regularity holds there were humans in London at the start of the eighteenth century. So the explanation of Arbuthnot's regularity that just says that there were humans in London at the start of the eighteenth century scores well on ROBUSTNESS.

In one sense this is the right result — the fact that there were humans in London at the relevant time clearly makes a difference to the truth of Arbuthnot's regularity. But on its own this isn't a good explanation — it is too unspecific; it misses out on almost everything of importance.

What exactly is wrong with very unspecific explanations like these? Intuitively, the issue is that just saying that there were humans in London at the start of the eighteenth century doesn't do much to lead us to expect that more males would be born than females each year. Although it is part of the causal history of the births we can infer almost nothing about whether those births will be male or female from the fact that there were humans in London.

The idea that explanations should make the phenomena in question expectable is a classic idea in the philosophy of explanation, notably expressed by ? and ?. Hempel's account, and many subsequent accounts, like for example, Strevens' account discussed in section 2, take this idea of expectability to the extreme by saying that an explanans must *necessitate* the explanandum, so the explanandum is *guaranteed* by the explanans.¹² And it does seem like a virtue of an explanation if the explanans guarantees the explanandum. But even if an explanation does not have this feature, it is good if it approximates it — it is good, that is, if the explanans makes the explanandum more expectable. Very unspecific explanations do badly on this count.

In order to deal with this issue, I propose that we should accept another dimension of explanatory goodness: PRECISION. PRECISION will tell us what is bad with these very unspecific explanations. And also, it will help capture our main idea that explanations should cite things that make enough of a difference. Difference-making is measured by ROBUSTNESS; what makes for enough of a difference is given by PRECISION — explanations should score well enough on ROBUSTNESS whilst still scoring reasonably well on PRECISION. (As we will see, this isn't so easy, because in many cases there is a trade-off between PRECISION and ROBUSTNESS.)

The formulation of PRECISION is very similar to that of ROBUSTNESS, but the guiding idea is very different. The central idea of PRECISION is that an explanation of B from A is better if more (in the sense of a higher proportion) of the ways that A could hold are such that, in that situation, A explains B.

We can translate this thought into the language of possible worlds to give the official formulation of

¹²Or at least, the simple version of Strevens' account does this, ignoring the nuances mentioned in footnote 5.

PRECISION: An explanation of B from A is better if that explanation holds in more, that is, a higher-proportion, of the physically possible worlds where A holds. (To compare, ROBUSTNESS said that an explanation of B from A is better if that explanation holds in more, that is, a higher-proportion, of the physically possible worlds where *B holds*.)

The evolutionary explanation of Arbuthnot's regularity which cites the greater infant mortality of males and the pressure towards a one-one sex ratio at the time of sexual maturity scores well on PRECISION because most of the physically possible worlds where there is greater infant mortality of males are such that Arbuthnot's regularity holds — that is, more males are born than females each year — and this is explained by the greater infant mortality of males.

On the other hand, the explanation of Arbuthnot's regularity which just cites the fact that there were humans in London at the start of the eighteenth century does worse on PRECISION. This is because there are many physically possible worlds where the explanans holds, that is, where there were humans in London at the start of the eighteenth century but where Arbuthnot's regularity does not hold, and so where the explanans doesn't explain the explanandum. For example, there are many physically possible worlds where females have greater infant mortality, and so in those cases Arbuthnot's regularity does not hold.

Or consider the ice cube case again: The explanation of the melting of the ice cube that just cites the fact that that you dropped it in warm water scores very highly (but not maximally) on PRECISION because in nearly every possible world where you do drop the ice cube in the water the dropping leads to it melting. But the explanation that cites the precise microstate M and the deterministic laws will score maximally on PRECISION¹³. Notice that PRECISION tends to favor more specific explanations over less specific ones.

One useful way to think about PRECISION is as implying a kind of graded necessitation. If A explains B and that scores highly on PRECISION that means that in most of the physically possible worlds where A holds, A explains B. And that implies that in most of the physically possible worlds where A holds,

¹³Again, assuming determinism.

B holds. So, when the PRECISION of this explanation is high then A ‘comes close’ to necessitating B within the range of physically possible worlds.

And so now we can see how PRECISION captures the idea that explanations should lead us to expect the explanandum. If the PRECISION of an explanation is high then the explanans comes close to necessitating the explanandum, in the way described above, and so the explanans should lead us to expect the explanandum.¹⁴

So, I claim, PRECISION is a second dimension of explanatory goodness. Explanations, I claim, should cite things that make as much of a difference as possible while still making the explanandum expectable. That is, they should score as well as possible on ROBUSTNESS while still scoring well on PRECISION.

These are the two dimensions of explanatory goodness. But how do they work together to get the plausible results regarding levels of explanation?

6 SPECIFICITY OF EXPLANATIONS

If we put PRECISION and ROBUSTNESS together we can see how we can value less specific, higher-level, explanations over more specific, lower-level explanations.

But in appealing to the combination of PRECISION and ROBUSTNESS I’m not going to claim that there is a single unique way to weight these dimensions against each other. It is plausible that different scientific fields have different preferences for PRECISION and ROBUSTNESS. The typical physicist,

¹⁴ In fact, PRECISION is a particularly attractive measure of this sense of expectability – more attractive, for example, than the probability of the explanandum conditional on the explanans. The expectability intuition is that the *explanation* should render the explanandum expectable. If B has a high probability conditional on A this might be true for reasons that are totally independent of any explanation of A from B. But, if the PRECISION of an explanation of B from A is high then this means that given A we should expect B *in virtue* of the explanatory connection between A and B.

Relatedly, the account avoids certain Hempelian problems for this reason. Take, for example, the classic case of an attempting to explain the high of a flagpole by appealing to the position of the sun and the length of the shadow that flagpole casts. Because PRECISION builds in the notion of explanatory correctness it gets the right result here. In none of the physically possible worlds where the sun is in that position and the shadow is that length do those facts explain the height of the flagpole. So, this putative explanation scores zero on PRECISION.

I suspect, values `PRECISION` highly; the typical sociologist values `ROBUSTNESS` more. The typical biologist perhaps somewhere in between. I don't think any of these preferences are mistaken.

Relatedly, identifying a unique balance of `PRECISION` and `ROBUSTNESS` would lead us to identify a single correct level of explanation for each explanandum — one explanation that 'best balances' the two dimensions. But this would be the wrong result. An account of levels of explanation should not, for example, say that the psychological level is the right level at which to explain the phenomenon of loss aversion, for example, and the neuroscientific level is not. Scientific practice is pluralist about the level of explanation here, and so we should be too.¹⁵

Rather, I'm just going to rule out very extreme preferences for `PRECISION` over `ROBUSTNESS`, and vice versa. These extreme preferences exhibit themselves in cases where people would accept a big loss in `PRECISION` to get a tiny gain in `ROBUSTNESS` or vice versa. Of course, it's vague what counts as an 'extreme' preference. But hopefully our discussion going forward will only appeal to relatively clear cases.¹⁶

(There's a question about what this 'ruling out' comes to. Am I saying that those extreme preferences are irrational? I am not. Rather, the idea is that people typically do not have such extreme preferences, so in attempting to make sense of what levels of explanation scientific practice typically favors in what cases it is legitimate to ignore those preferences. If someone has a very extreme preference for `PRECISION`, say, and subsequently favored much more detailed explanations than we normally do, I would not claim that they are mistaken, rather they are just not in keeping with the explanatory standards of the community.)

Given this, we are now in a position to see how `PRECISION` and `ROBUSTNESS` tell us how specific an explanation should be. The first step is to notice that the ideal case, where `PRECISION` and `ROBUSTNESS`

¹⁵In this way the view is reminiscent of? in that there are two dimensions that scientists weight differently.

¹⁶A brief methodological aside: Ultimately, we are trying to understand and account for a feature of scientific practice — we are trying to make sense of the way certain levels of explanation are taken in scientific practice to be acceptable for a particular explanandum and some are not. This is not a particularly sharp phenomenon — there are no bright lines to be seen in the practice between explanations that clearly acceptable and those that are not. Disagreement and vagueness abound. So, we should not expect, and not desire, our philosophical account to be totally free of vagueness and to draw bright lines.

are maximal, is one where the explanans and the explanandum ‘overlap’ in the range of physically possible worlds – that is, they hold in exactly the same range of physically possible worlds. To see this look at Figure 1.

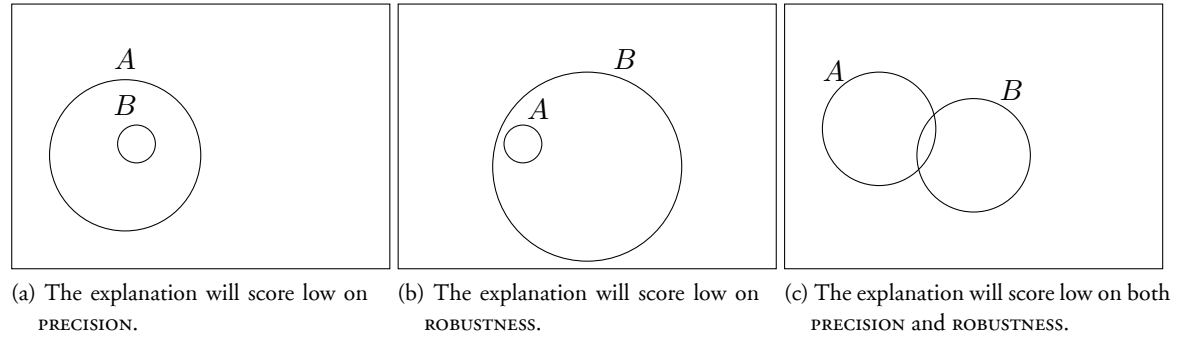


Figure 1: PRECISION and ROBUSTNESS together imply proportionality

In the diagrams the boxes represent the space of physically possible worlds and the circles represent the range of worlds where the propositions A and B hold. In 1(a) A holds in a much wider range of worlds than B, and so an explanation of B from A is guaranteed to score low on PRECISION because in most of the worlds where A holds, A does not explain B.

1(b) is the opposite case. B holds in a much wider range of worlds than A and so an explanation of B from A is guaranteed to score low on ROBUSTNESS because in most of the worlds where B holds it is not explained by A. The microphysical explanation of the decline in the dollar is a case of this kind.

When PRECISION and ROBUSTNESS are both maximal A and B perfectly overlap. Deviations from this overlap reduce explanatory goodness. This, on its own, tells us a lot about the level of specificity an explanation should have – the specificity of the explanans should match the specificity of the explanandum.

(This idea of ‘matching’ or ‘proportionality’ means that the account is in the spirit of Yablo’s (?) discussion of causation. However, my account doesn’t put in the claim about matching by hand – rather it follows from independently motivated thoughts about explanations making the explanandum expectable, and citing things that make a substantial difference. And, more importantly, PRECISION

and ROBUSTNESS allow us a measure of how things can be more or less proportional, thus allowing the account to be a graded difference-making account.)

In most realistic cases, though, we won't have perfect matching between explanans and explanandum. Let's consider a more realistic case that we mentioned earlier: I have some peas, and I can explain why they are smooth and not wrinkled by noting that they were bred from parent plants that were homozygous in the allele R that is associated with smoothness. This is a good explanation, but it does not score maximally on PRECISION – there are cases where the peas were bred from parents that were homozygous in the allele associated with smoothness but where the peas do not end up smooth. For example, there is always the possibility of mutations leading to the peas being wrinkled. Similarly, disease or dehydration, or a variety of other factors could lead to the wrinkling. Nevertheless PRECISION is high because such cases are rare – in most of the physically possible worlds where the peas were bred from parents that were homozygous in the allele associated with smoothness that fact explains the peas' smoothness.

We could give an alternative lower-level explanation that is designed to do better on PRECISION. For example we could give an explanation that cites the specific molecular details of the inheritance. This molecular explanation would have an advantage on PRECISION over the genetic explanation because it would rule out the possibility of certain mutations. But, since the genetic explanation already scores well on PRECISION this advantage will be relatively small. The molecular explanation, on the other hand, would score *far* worse on ROBUSTNESS than the genetic explanation – the huge amount of added detail needed to move from a classical genetic explanation to a full molecular account leads to a substantial loss in ROBUSTNESS. Unless we have very extreme preferences for PRECISION over ROBUSTNESS, then, my account is going to favor the genetic explanation over the molecular explanation.

Or, to put it another way, the molecular explanation includes lots of facts that don't make much of a difference to smoothness of the peas — the peas could easily still be smooth even if the precise molecular details were different — and these facts don't add enough to the expectability of the

explanandum to be worth including in a good explanation.

Another example: Consider the case of the dollar declining in value that we discussed above. The microphysical explanation of the decline in the dollar – the one that cites the movement of all the particles that make up the economic system does a little better than the economic explanation on PRECISION but far worse on ROBUSTNESS. That is to say, the explanation cites facts that don't make much of a difference, and this isn't compensated for by a big gain in PRECISION.

So we can start to see, then, how my account can pick out the right degree of specificity. The ideal case is one where the specificity of the explanans matches that of the explanandum. But in most cases we don't get this perfect matching and we have to weigh PRECISION and ROBUSTNESS against each other.

7 BACK TO DETERMINISTIC CHANCE

So we have an account of explanatory goodness that develops the notion of difference-making we discussed in section 2. And this account gets us the results we want about the cases that motivated us — we can see, for example, why we would favor a genetic explanation of the smoothness of the peas to a molecular one, even though the genetic explanation does miss out on certain facts that make a difference. Those facts that are missed out don't make enough of a difference, as measured by the interaction of ROBUSTNESS and PRECISION.

There's a lot more to say about this account, and the implications it has for levels of explanation. But we need to bring the issue back to deterministic chance. How does this account connect to deterministic chance? Well, as we noted in section 2, it's vital to have a *graded* notion of difference-making — it is only by having such a notion that we can make sense of why good higher-level explanations leave out lower-level facts that seem to make a difference. I argued that an appropriate notion of graded difference-making can be had by appealing to ROBUSTNESS and PRECISION. But both ROBUSTNESS and PRECISION are defined in terms of proportions of worlds. The ROBUSTNESS of

an explanation of B from A, for example, is given by the proportion of physically possible B-worlds where A explains B.

To make sense of these claims about the proportion of worlds we need a type of objective probability. Following §, section 4 I'm going to call this *counterfactual probability*. We can introduce counterfactual probability as follows:

$S \sqsubset_x \rightarrow T$ is true if and only if x of the physically possible S -worlds are worlds where T

x here is the *proportion* of physically possible S -worlds where T is true. That is, x is given by counting the number of physically possible $T \& S$ -worlds and dividing by the number of physically possible S -worlds. We say x is the counterfactual probability of T given S .

We can formulate PRECISION and ROBUSTNESS in these terms. Here is how to characterize the PRECISION of an explanation of B from A:

$A \sqsubset_x \rightarrow A \text{ explains } B$

x here is the counterfactual probability of A explains B given A, and gives us the PRECISION of an explanation of B from A.

For ROBUSTNESS consider:

$B \sqsubset_x \rightarrow A \text{ explains } B$

x here is the counterfactual probability of A explains B given B, and gives us the ROBUSTNESS of an explanation of B from A.

A natural concern to have about this notion of counterfactual probability, however, is that it's somewhat unclear how to define proportions of worlds when there is an infinite number of relevant worlds. To get the counterfactual probability of T given S I said to count the number of physically possible $T \& S$ -worlds and divide by the number of physically possible S -worlds. But the notion of counting isn't clear in the infinite case — the generalization of the notion of counting to the infinite case is (extremely) non-unique. So, we need to say something about how to count infinite numbers of

worlds — more specifically, we need to claim that there is some privileged measure over physically possible worlds.

How do we argue that there is a privileged measure over such worlds? Well, I think, by doing exactly what we have been doing in the paper so far — that is, showing the theoretical payoffs of taking there be such a measure. If accepting some measure over physically possible worlds is a powerful addition to our theory — if it helps us understand important phenomena — then that's a reason to accept such a measure. And in the paper so far I have, in effect, argued that accepting such a measure allows us to understand our practices of favoring higher-level explanations over lower ones and, more generally, in making sense of the value of the higher-level sciences.

As we noted in section 1 it looks like some higher-level explanations – for example, the higher-level explanation of Arbuthnot's regularity – have an objective, non-pragmatic advantage over competing lower-level explanations. My account makes sense of this by appealing to such an objective measure over worlds and an objective sense of counterfactual probability which we can use to define ROBUSTNESS and PRECISION.

(Again, this is somewhat in the spirit of ?, who introduces a notion of counterfactual probability and claims that it is connected to Jackson and Pettit's (?) claim that higher-level explanations provide a distinctive type of information. Though I give an account of counterfactual probability, where he did not, and I give a precise account of explanatory goodness based on this type of probability.)

And as I noted in the introduction, this fits with a variety of other ways of arguing for deterministic chance, for example, Emery arguing that deterministic chance allows us to make sense of our practice of asserting counterfactuals and Ismael arguing that it's needed for experimental confirmation in science.

So, I think, we do have reason to think that there is such a privileged measure, and thus that there is an objective notion of counterfactual probability. But which measure is the privileged one? Well, the natural answer is that it's the standard statistical mechanical measure (or some extension of this measure, like the so-called 'Mentaculus' of ? and ?). The standard SM measure is precisely the way

that scientists express the idea that there are more ways that the ice cube could melt rather than grow when dropped in warm water. (And, as a reviewer notes, the use of such a measure in statistical mechanics is a further reason to think that there is a privileged measure.)

But regardless of the details of the measure, I think we have reasons to think that a privileged measure exists. And consequently, that there is a corresponding objective notion of counterfactual probability. Seeing how such a notion of probability helps us properly understand the value of higher-level explanations is a powerful reason to accept it.

Of course, though, as we noted in the introduction, someone might accept that my approach to levels of explanation involves accepting counterfactual probability, but deny its attractiveness. Perhaps they would suggest there are more attractive approaches to the question of levels of explanation that do not commit to anything like deterministic chance.

This is a perfectly reasonable concern, and since I certainly can't survey and criticize all these other approaches here I can't take what I've said to be anywhere near conclusive. But, I think there are some general reasons to think that an appeal to something like counterfactual probability is needed when addressing the question of levels of explanation. Put very generally, the issue is that when we give a higher-level explanation of a phenomenon we ignore lots of things, in particular, we ignore all the relevant lower-level facts. This wouldn't be problematic if once we have the higher-level explanation the lower-level facts become irrelevant — that is, if they don't add anything to the higher-level facts we have. (This is what ? calls 'conditional irrelevance'.) If this were the case it would be easy to justify our ignoring the lower-level facts. But this doesn't typically hold — lower-level facts are still relevant. Even when we know that the ice cube was dropped in warm water, there are still lower-level facts that are relevant to whether it melts. And so, we need a reason to ignore those lower-level facts, like the exact microstate of the ice cube. This, I think, is the central issue in the literature on levels of explanation as it stands. Can we give a, preferably non-pragmatic, reason why we should ignore these relevant lower-level facts in our explanations? But, it's somewhat hard to see how this could be done without appealing to ideas like *if almost any other microstate held the ice cube would still have*

melted. Perhaps that idea doesn't need to be developed in exactly the way I did. But it does seem to involve a commitment to measures over ways the ice cube could have been — that is, a measure over worlds. In fact, I think the approaches of ? and ?, among others, are based on this kind of idea, and also suggest a commitment to a measure over worlds.

So, I think we have a good reason to think that satisfying answers to the question about levels of explanation will involve commitment to objective deterministic probability, even if the account is not precisely the one that I gave. Though again, I can't claim to be conclusive. Perhaps other approaches are possible. Certainly if we want to deny a claim made at the start — that the lots of higher-level explanations are superior to lower-level ones, even when we can properly grasp the lower-level explanations — then pragmatic approaches, that focus on the difficulty to humans of theorizing in lower-level terms, are always possible. (In fact, as a reviewer notes, if we think that the advantage the higher-level explanations have is pragmatic then it may be possible to accept a large amount of my account but construe it differently. For example, by taking the measure over possible worlds as non-objective and PRECISION and ROBUSTNESS as representative of pragmatic advantages that explanations can have.) But regardless, I think the argument here gives us reason to be pushed towards accepting deterministic chance, even if it's not conclusive.

One final point: As I noted early in the paper someone might agree with much of what I've said, but think that the probability I argue for doesn't count as *chance*. Again, I don't want to rest much on the word chance. If people want to use a restricted notion of chance that rules out counterfactual probability as being chance I will not argue much — we are still left with an argument for an important type of objective deterministic probability.

But, nevertheless, there are some considerations that suggest it is reasonable to call the type of probability I've argued for 'chance'. Most importantly, it's plausible that counterfactual probability is credence-guiding. Consider the counterfactual probability of the ice cube melting given that it was in warm water. This counterfactual probability is high. So, if counterfactual probabilities guide credence then we should have a high-credence in the ice cube melting when it is in warm water.

And this is exactly what we see in the practice of statistical mechanics — the fact that most (in fact, nearly all) of the physically possible ways that the ice-cube and the water could be lead to the ice cube melting in warm water is taken to imply that when we see an ice cube in warm water we should expect it to melt.

Further, counterfactual probabilities seem to play other roles associated with chance. Consider, for example, the commonly discussed connection between between chance and possibility (??) – intuitively that when an event has a non-zero chance there is a relevant physically possible world where it occurs. This comes easily for counterfactual probability, since counterfactual probability is just a measure over relevant physically possibility worlds.

(Precisely formulating these connections of deterministic chance to credence and possibility is tricky – see, for example, ? – but the intuitive connections are there.)

More generally, the close connection between counterfactual probabilities and the probabilities of statistical mechanics, which many take to be reasonably classified as chances, give us reason to suspect that counterfactual probability is properly classified as chance. But, again, not much rests on this. If you disagree with such a classification just think of counterfactual probability as an objective deterministic probability that plays an important role in explanation.

8 CONCLUSION

I've argued that if we accept deterministic chance then we can develop the difference-making approach in order to give an attractive answer to the question of levels of explanation. And, more generally, it looks like an approach that appeals to deterministic chance is going to be needed to make sense of why we ignore lower-level facts that do make a difference to the event in question. This gives us reason to accept a package of views including deterministic chance.

References