Explanatory Abstractions

Lina Jansson and Juha Saatsi*

December 3, 2015

Abstract

We distinguish different senses in which an explanation can be more or less abstract, and we analyse the connection between explanations' abstractness and their explanatory power. According to our analysis abstract non-causal, mathematical explanations have much in common with counterfactual causal explanations. This opposes a clear recent trend to regard abstractions as explanatory in some sui generis way.

1 Introduction

> *On my view, explanation is epistemic, but with a solid metaphysical basis. A realist theory of explanation that links the determinative (or dependency) relations in the world with explanation gets at the intuitively acceptable idea that we explain something by showing what is responsible for it or what makes it as it is. This is what, in the end, explains explanation.* [Ruben, 1990, 233]

There is broad agreement that many explanations derive their explanatory power from information about *dependence*. Causal explanations are typically understood in terms of causal dependence, and many have suggested that various kinds of non-causal explanations could be similarly understood in terms of non-causal dependence of the explanandum on the explanans.[1] It is reasonable to hypothesize that all explanations involve information about what the explanandum (causally or non-causally) depends on. This hypothesis, to the extent it is tenable, could be part of a unified account of explanations, according to which all explanations—causal or otherwise—are explanatory due to providing information about dependences.

A consideration against this hypothesis comes from *explanatory abstractions*. Explanations vary in how abstract they are: some explanations appeal to relatively abstract features of reality, whereas others turn on much less abstract explanans.[2] We value explanatory abstraction, when appropriate: it filters out details that are in some way irrelevant to the explanation in question. But abstraction does not directly concern what

*This work is the result of an equal collaboration. We list the authors alphabetically.

[1] See e.g. Woodward [2003], Woodward and Hitchcock [2003a,b], Ylikoski and Kuorikoski [2010], Strevens [2008]. Ruben [1990] is an early proponent.

[2] Contrast an explanation that turns on abstract topological features of space—e.g. its orientability—to an explanation that turns on a specific concrete feature of an event's causal history.

the explanandum depends on. Rather, abstraction is about a kind of *independence* of the explanandum from the irrelevant details. And one may be tempted to think, along with various recent authors, that this information about *independence* is where the explanatoriness of many abstract explanations lies. This mode of explaining, by reference to what is irrelevant to the explanandum, can seem *sui generis*, different in kind from explanations in terms of dependences.

We will argue that explanatory abstractions do not point to a sui generis mode of explaining. They do not call us to revise the idea that explanatory power is a matter of providing information about what the explanandum depends on (and the more the better), and there is no need to abandon the hope for a unified account of explanatory power. In saying this we stand against a clear trend that urges the opposite: that some abstract explanations do not fit an account of explanatory power that is focused on (non-)causal dependence.[3] More specifically, we aim to (i) undercut prominent arguments exhibiting the trend against a unified account of explanatory power; (ii) to present such a unified account, capable of capturing influential exemplars of abstract explanations that have motivated the anti-unificationist trend; and (iii) to provide an argument for our unified account.[4]

Before we get to the nitty-gritty, let us illustrate explanatory abstraction by a simple example, and sketch some of the points that we will develop in the paper. The explanation of why Mother fails to divide her twenty-three strawberries equally among her three children (without cutting any strawberries) turns on the fact that twenty-three is not evenly divisible by three.[5] It would make for a worse explanation to cite the causal details or the causal laws involved in Mother's attempt, or the specific physical constitution of the strawberries. The explanation is *independent* of any such details; the explanatory mathematical facts hold irrespective of contingent causal laws and details. So where does the explanatory power associated with mathematics come from? It's not at all clear that mathematical facts capture dependence relations between numbers that are in any way analogous to contingent causal and nomological connections. It is tempting to think that the explanatory contribution of mathematics in this case cannot be captured in terms of dependences. Instead, one may think that mathematics' explanatory contribution is sui generis, that it somehow turns on mathematics' *independence* from the details of the physical constitution and the causal laws. (Later in the paper we will explicate and rebut different versions of this intuition.)

We see little reason to shift focus from dependence to independence in analysing abstract mathematical explanations of this kind. It is easy to overlook the wealth of information about dependence that we take for granted in this case. The explanation does not just tell us that Mother must fail to divide the strawberries evenly, as a matter of mathematical necessity. It also tells us what the failure depends on—the number of

[3]See e.g Lange [2013], Pincock [2007], Pincock [2015], Batterman and Rice [2014]. (We include Pincock [2015] here since he argues that none of the existing accounts with their focus on dependence can be extended to capture highly abstract explanations.

[4]By calling ours a 'unified' account we do not pretend to show that it captures *all* causal and non-causal explanations. Rather, we claim that it captures those prominent and influential exemplars of non-causal explanation that we discuss in this paper. Whether some other non-causal explanation fits our thinking would have to be looked at on a case-by-case basis.

[5]This example is from Lange [2013], who credits the inspiration for the case to Braine [1972].

strawberries—and exactly *how* it depends on this feature. For example, if Mother had two fewer strawberries or one strawberry more, then she could have succeeded in dividing them evenly among her three children. This change relating what-if-things-had-been-different information is exactly the kind associated with explanatory dependence.[6] We will argue, furthermore, that this information is doing all the explanatory work.

The importance of dependence information can be gauged by varying its amount while keeping other things fixed. Consider removing it *altogether*, first of all. That is, let us pretend that we can remove all information about dependence that relies on the background arithmetical knowledge, only retaining the information about the independence of the explanandum from causal laws and so on. If *all* we know is that as a matter of mathematical necessity Mother's attempt with twenty-three strawberries must fail, so that we are unable to answer any questions about under what changes the attempt would succeed, then it is unclear whether we have *any* explanation of Mother's failure.

Consider, now, adding just a *little bit* of information about dependence. That is, pretend that instead of being able to use our general understanding of arithmetic, we only get to know that Mother could have succeeded with twenty-four strawberries. Assume that is all the information we have, on top of knowing that twenty-three is not divisible by three. Here we may have an explanation, but it looks like a very thin one. Yet, in terms of independence from particular causal laws, processes, and microphysical conditions, it is on a par with the general explanation that appeals to our background arithmetical knowledge. If the explanation that makes use of our background arithmetical knowledge derives its explanatory power (partly) from its great independence from physical details and laws, then we would expect the modified explanation to be explanatory to a commensurate degree. Yet, we do not see this. When we vary the amount of information about dependence, we also seem to vary the degree of explanatoriness, in a way that suggests that it has little to do with independence information.

To carefully develop the line of thought sketched above, we first need to lay some conceptual groundwork. We will do this by distinguishing three different dimensions of abstraction (§2). We will then show how these distinctions apply to some prominent recent conceptions of explanatory abstractions (§3). Our key claim is that not all 'dimensions' of abstraction are responsible for explanatoriness, even if good explanations are typically abstract along several dimensions. Based on the conceptual distinctions between three different dimensions of abstraction we will present an alternative counterfactual account of the explanatoriness of abstract explanations (§4). This allows for a unified conception of explanation, in tune with the spirit (but not the letter) of the dependency oriented account developed by Woodward [2003] and Woodward and Hitchcock [2003a,b]. The virtues of this account are illustrated by revisiting a graph-theoretic explanation of Königsberg's bridges (§5) and further underlined in relation to other challenges faced by the less unified views of explanation (§6).

[6] Most prominently in Woodward [2003].

2 Three dimensions of abstraction

As said, we will challenge the idea that explanatoriness turns on independence of the explanation from the particular physical laws, processes, or concrete physical structures. Yet, we think that abstract explanations typically do exhibit this kind of independence. The task of this section is to distinguish three conceptually different ways in which explanations can be abstract. They will all involve independence in one sense or another. However, we will later argue that only one kind of independence carries information about explanatory dependence, and, as a consequence, contributes to explanatory power.

Let us recall some simple examples of abstract, plausibly non-causal, mathematical explanations in order to bring out intuitions about different senses of explanatory abstraction.[7] The following examples have featured prominently in recent philosophical analyses: (1) Economical bees build hexagonal honeycombs because this is the most resource-efficient way to divide a Euclidean plane into regions of equal area with least total perimeter [Lyon and Colyvan, 2008]. (2) Plateau's laws hold for soap film geometry, because this geometry minimizes the system's energy by minimizing the surface area [Lyon, 2012, Pincock, 2015]. (3) Königsberg cannot be toured by crossing each and every bridge exactly once because of the bridges' relational configuration [Pincock, 2007]. (4) The strawberry case discussed above is another example, of course.

These are paradigmatic exemplars of what we call 'abstract explanations'. (We will focus on (3) and (4) in our main discussion, covering (1) and (2) in the Appendix.) Abstract explanations hinge on (relatively) abstract features of reality that are non-causal (or perhaps causal in a very broad sense of 'causal'). This abstractness is obviously partly a matter of relative independence of the explanans from the *actual physical structure* of the entities involved: strawberries, bridges, soap films, honeycombs, etc. This is the first intuitive aspect of explanatory abstraction that we wish to highlight. There is nothing special about the physical nature of strawberries, for example, that makes them indivisible in this way. The very same explanation essentially explains why twenty-three marbles (say) cannot be thus evenly divided. The strawberry explanation abstracts away from whatever concrete features need to be in place in order for there to be a set of twenty-three *individuals* that can be divided into mutually exclusive and collectively exhaustive subsets.

A second, more radical aspect of abstraction pertains to the fact that the explanations above are independent from the *actual laws of nature* that underlie the physical processes presupposed by the why-question: the processes of group division, bridge crossing, or comb building, for example. There is nothing special, as far as these explanations are concerned, about the actual physics that underlies the solidity of honeycomb walls, strawberries, or bridges. The same explanations would work just the same in more exotic, counter-legal possible worlds where the underlying physical laws are rather different (assuming the relevant why-questions still make sense). The explanation of Plateau's laws, for example, transcends the nomologically possible causes or grounds of the area-minimizing action of soap bubbles. This explanation is modally robust with

[7] It is a matter of an on-going debate whether these explanations are undeniably non-causal, as opposed to being causal in some very broad, non-paradigmatic sense (ditto regarding their status qua 'mathematical' explanations). Our concerns in this paper are independent of these debates.

respect to variation in the underlying physical laws, as long as the laws still give rise to area-minimizing action for some surfaces. To this extent the explanation abstracts away from the actual nomological facts pertaining to the systems in question.

So far we have identified a conceptual distinction between two senses of abstraction. There is also a third aspect of abstraction in the above explanations; one that is more easily missed. This has to do with the explanatory regularities involved, and the degree to which an explanatory generalisation is independent from the actual value of an explanans variable. For example, the fact that twenty-three is not evenly divisible by three is but an instance of the more general fact that no integer apart from whole multiples of three is evenly divisible by three. In as far as we judge that the explanatory work is really done by this more general arithmetical background knowledge, we can say that essentially the same explanation would answer the question of why Mother would have failed had she had twenty-two strawberries (or why she could have succeeded had she had twenty-one). The explanation turns on a regularity about numbers that is robust by virtue of allowing the same explanation to be given for a wide range of alternative numbers of strawberries. The wider this range is, the more abstract the explanation, in this third sense of abstraction.

The above three intuitive aspects of explanatory abstraction all clearly have something in common. Namely, each involves a way in which an explanation can effectively be independent from some feature of the system in question. This independence is naturally thought of as follows. An explanation can cover a range of possible systems. Explanatory abstraction is a matter of independence of the explanation from the details of 'realization'. In the above cases of abstract explanations the extent of unnecessary details omitted is striking. Consider the Königsberg case, for instance: the explanation is independent of all the physical and geometrical features of the bridges and the crossing processes, and of the underlying nomological facts pertaining to the bridges and their possible crossings, only relying on a very high-level global structural feature associated with a wide-ranging explanatory regularity. (We will make this more precise later in the paper.)

The extent of such independence is most striking in the case of abstract explanations, but more or less any explanation actually exhibits the three dimensions of abstraction to some degree. Consider the mundane case of gravitational pendulums, for instance. Why is the period of a given (more or less) 'ideal' pendulum 10 seconds? Explaining this causally, in terms of its length l and gravitational acceleration g, relies on a law-like regularity that supports the same explanation in a range of possible cases that vary in these two parameters. (The third dimension of abstraction, above.) Furthermore, the explanation can apply just the same regardless of whether the underlying nomological facts are rooted in a classical Newtonian world with gravitational force acting at a distance, or whether the pendulum occupies a general relativistic world with gravity being a manifestation of curved spacetime. (The second dimension above.) Finally, the why-question presupposes that the pendulum cord is inextensible, but the explanation is independent from whatever microstructural facts underlie this property. (The first dimension above.) We see the difference between this mundane causal explanation and 'abstract explanations' as one of degree, not of kind.

The notion of explanatory abstraction as a matter of independence is clearly a modal notion: it concerns the range of possible systems covered by the explanation. Different aspects of explanatory abstraction have to do with different dimensions of modal variation with respect to these possible systems. We can think of possible systems that vary from the actual explanandum only in the (nomologically) possible physical structures of the entities involved; possible systems that vary in the microphysical or dynamical laws they obey; possible systems that vary from the actual explanandum by varying a variable that explicitly features in the explanatory regularity employed. Although modal variation along these different dimensions is often interlinked and metaphysically intertwined in complex ways, the dimensions themselves are conceptually independent.[8] In the strawberry case, thinking about the explanatory regularity with respect to a system of twenty-three vs. twenty-five strawberries, one need not think about variation in the physical structures of these entities or the division processes. Similarly, in thinking about the explanation with respect to systems that vary in the physical structures of the entities to be divided, one need not think of variation in their number, or of counter-legal possibilities where the underlying physical laws are different. And vice versa.

Often abstract non-causal and mathematical explanations—like the ones above—exhibit a striking level of abstraction in *each* of the three senses. We believe this has confounded some recent commentators. If one thinks, prima facie, that abstraction in a particular sense is the source of explanatoriness, it is all too easy to find corroborating evidence when so many good explanations are highly abstract in that sense. Now that we have conceptually separated the different dimensions of abstraction, however, we can ask: If explanatory abstraction has something to do with explanatory power (as is commonly accepted), exactly which dimension(s) of abstraction can contribute to it?

In its full generality, we cannot hope to comprehensively answer this question in a single paper. There is, however, a narrower question that we can answer. It is commonly agreed that explanation (and the associated notion of explanatory power) has both worldly and pragmatic/communicative aspects.[9] Here, we will largely set aside the communicative and pragmatic aspects. This is not because we think that there are no such desiderata associated with explanation.[10] What is at stake in the debate at hand, however, is whether abstract explanations require facts about independence to function as a source of explanatory power, instead of (or in addition to) facts about dependence. In particular, are facts about independence a worldly underpinning of explanatory power for the paradigmatic cases in the literature (mentioned above)? Our first aim is to undermine prominent recent arguments that answer 'yes' to the last question.

Our second aim is, more positively, to show how these abstract explanations can be accommodated in more unified terms, in the spirit of the counterfactual account developed by Woodward and Hitchcock [2003a,b] (the W-H account). According to the W-H account, explanations provide a particular kind of information about dependence.

[8]The dimensions are metaphysically intertwined because, for example, various possible changes in the physical laws are bound to imply changes in the physical structures.

[9]Some, for example van Fraassen [1977, 1980] and Faye [2014], take explanation to be solely pragmatic/communicative. This is not the case with most of the recent commentators on explanatory abstractions.

[10]Some of the dimensions of *depth* discussed by Woodward and Hitchcock [2003b] are best understood as pragmatic/communicative. So are some of the criteria listed by Ylikoski and Kuorikoski [2010].

This account associates explanatoriness essentially with our third dimension of abstraction. (Roughly, an explanatory generalisation that has a broader range under which we can vary the values of the variables provides more explanatory information about the dependence of the explanandum on these variables than a corresponding explanatory generalisation with a more restricted range of invariance.) This is in sharp contrast to recent claims according to which (some) abstract explanations explain either (i) by virtue of abstracting away from all the concrete physical features of the system in question (our first dimension above), or (ii) by virtue of abstracting away from the underlying laws of nature (our second dimension). The W-H account is well liked in connection with various causal explanations (even amongst the advocates of alternative analyses of abstract explanations), but there has been resistance to extending this counterfactual account to abstract explanations that are (plausibly) non-causal or mathematical. We will argue that this resistance is entirely unnecessary. Issues concerning explanatory power and its relation to abstraction are tractable in very natural terms in the spirit of the W-H account also for abstract explanations, after the key notions of this account are appropriately clarified and refined.

We will next critically review some recent claims regarding explanatory abstraction, questioning the intuitions that support them. After that we will more systematically discuss the third dimension of abstraction, which has a better claim to be associated with the source of explanatory power than either dimension one or two.

3 Explanatory power and independence

Various philosophers have recently argued (in different ways) for the disunity of explanatoriness by pointing to the sui generis role of abstraction as a source of explanatory power. We will focus on two authors who particularly clearly associate explanatory power with either the independence from concrete physical details (dimension one above), or the independence from the particular laws (dimension two above).

Pincock [2007, 257] brought the Königsberg bridge example into philosophical prominence by identifying it as an 'abstract explanation' that 'appeals primarily to the formal relational features of a physical system.' Relational features that are 'formal' are clearly meant to stand apart from causal relations. Furthermore, regarding the explanation of the impossibility of touring the town by crossing each and every bridge exactly once, Pincock [2007, 259] accounts:

> [A]n explanation for this is that at least one vertex [in the formal graph structure instantiated by the bridges] has an odd valence. Whenever such a physical system has at least one bank or island with an odd number of bridges from it, there will be no path that crosses every bridge exactly once and that returns to the starting point. If the situation were slightly different [so that] the valence of the vertices were to be all even, then there would be a path of the desired kind.

This is all surely correct. The question is how to capture this explanation in philosophical terms. Pincock [2007, 260] intimates that the explanation critically turns on abstraction:[11]

> The abstract explanation seems superior [to a microphysical explanation] because it gets at the root cause of why walking a certain path is impossible by focusing on the abstract structure of system. Even if the bridges were turned into gold, it would still have the structure of the same graph, and so the same abstract explanation would apply. By abstracting away from the microphysics, scientists can often give better explanations of the features of physical systems.

The notion of 'abstracting away from the microphysics' is naturally construed as being along the first dimension identified above: a matter of relative independence of the explanans from the physical structure of the entities involved. As Pincock [2011, 213] puts it: 'the explanatory power [in this case] is tied to the simple way in which the model abstracts from the irrelevant details of the target system.' The intuition is that the explanation comes from stripping away as irrelevant all the physical details pertaining to the bridges' make-up, length, location, angles, et cetera, thereby highlighting what is relevant, namely the formal 'mathematical structure found in the target system itself' Pincock [2011, 213].[12]

We can see Pincock's intuition, but we are unable to see a good reason for viewing this as a sui generis 'abstract explanation'. The explanation is undeniably highly abstract and plausibly non-causal. But the notion that good explanations only provide relevant information, leaving out unnecessary details, is common ground with dependence accounts of explanation. According to dependence accounts one ought to provide suitable information about what the explanandum depends on, and one also ought not to claim (or imply) that the explanation depends on something that it does not. The leaving out of irrelevant detail is insufficient to motivate the view that 'abstract explanations' are *sui generis* in the way Pincock regards them.

How about the notion that the explanatorily relevant features of Königsberg are 'formal-cum-mathematical'? Does this motivate a departure from familiar counterfactual accounts of explanation? We do not think so. We do not regard the explanans inherently formal or mathematical in any substantial sense. The explanation involves *applied* mathematics; it is not an intra-mathematical explanation. There is a clear sense in which the explanans is just concerned with how many bridges there are to/from each of the 'islands'. It is unclear why the 'valence' of an island—there being an even or odd number of bridges to/from it—is not a high-level *physical* feature of it. And similarly for the yet more abstract feature of there being *at least one* odd 'island' in the whole system of many 'islands'. Furthermore, there seems to be a clear and straightforward sense in which the explanandum at stake *depends on* these arguably high-level physical features: the explanation tells us how the explanandum would change if there were a different number of bridges (as Pincock himself notes; see the first quote above). Whether

[11] Pincock [2007] does not endeavour to develop a detailed account of the abstract explanation. This is given in Pincock [2015].

[12] Similarly Batterman [2010, 3] argues, in relation to other kinds of mathematical and non-causal explanations, that explanatory power is connected to a 'systematic throwing away of various causal and physical details'.

this explanatory dependence is causal or not is neither here nor there for the prospects of sticking to the core idea of the W-H account of explanation: explanatoriness is ultimately a matter of telling us what the explanandum *depends on*.[13] When it comes to identifying the worldly source of explanatory power, we have not yet been given a reason to think that in the Königsberg case explanatoriness derives (even partly) from the *independence* of the explanandum from some worldly features.

Let us now move on to consider Lange's [2013] analysis of the explanatory power of the cited abstract explanations (as 'distinctly mathematical').[14] Lange [2013, 486–488] argues that his analysis of these explanations reveals 'a fundamental difference' between causal explanation and abstract non-causal explanation, and its significance 'lies in what it reveals about the kinds of scientific explanations there are'. Abstract explanations, Lange argues, explain not by supplying information about the world's network of causal relations but by 'showing how the fact to be explained was inevitable to a stronger degree than could result from the causal powers [actually] bestowed by the possession of various properties.' This is a nod towards Wesley Salmon's 'modal conception' of scientific explanations, according to which such explanations 'do their jobs by showing that what did happen had to happen' [Salmon, 1985, 293]. As Lange [2013, 505] puts it:

> [They explain] not by describing the world's actual causal structure, but rather by showing how the explanandum arises from the framework that any possible causal structure must inhabit, where the 'possible' causal structures extend well beyond those that are logically consistent with all of the actual natural laws there happen to be.

In this way Lange relates explanatoriness of abstract explanations to information about the *independence* of the explanandum from some contingent features of the system in question. In contrast to Pincock, the critical sense of independence for Lange is in the direction of the second dimension of abstraction identified above: abstraction away from the actual laws underlying the features of the system presupposed in the context of the relevant why-question. In relation to the Königsberg case, for instance, Lange [2013, 505-506] writes:

> [The] explanation of the repeated failure to cross the Königsberg bridges shows that it cannot be done (where this impossibility is stronger than physical impossibility) [...] The explanans consists not only of various mathematically necessary facts, but also [...] of various contingent facts presupposed by the why question that the explanandum answers, such as that the arrangement of bridges and islands is fixed. The distinctively mathematical

[13]After heavily emphasising the idiosyncratic features of abstract explanations Pincock [2015] suggests that it may be possible to provide a unified account in terms of dependence. The idea is that in addition to causal dependence there may be a dependency relation captured by the relation of a concrete entity instantiating an abstract entity. We find it difficult to see how a concrete entity is an instance of an abstract entity in analogy to the type-token distinction—Pincock's explication of the dependence relation. More importantly, Pincock's [2015] notion of dependence is not in the spirit of the W-H account, to which Pincock objects for a somewhat technical reason. We will address this objection In section §4.

[14]The term 'distinctly mathematical' is a term of art in Lange [2013]. It is aiming to capture many of the cases that Pincock [2015] would call 'highly abstract'.

explanation shows it to be necessary (in a way that no particular force law is) that, under these contingent conditions, the bridges are not crossed.

Lange's account thus emphasizes the way in which the actual physical laws do not come into play in explanations such as this. Undoubtedly Lange is right to note that the sense of abstraction at play goes beyond the first dimension concerning the irrelevance of nomologically possible physical realizations of the contingent features presupposed in the context of the why-question. The same explanation would work, in the same way and to same extent, in a far removed possible world with alien properties and laws—as long as there is a system with 'traversible bridges' for which the why-question makes sense.

But, having said that, why should we regard *this* sense of abstraction as a source of explanatoriness in connection with abstract explanations? The answer to this question is surprisingly difficult to find in Lange [2013].[15] There is little to directly motivate this, beyond the commonplace idea that good explanations do not mention irrelevant details, which in this case include all causal laws. Yet again, the idea that good explanations only provide relevant information, leaving out unnecessary details, is common ground with any account according to which we ought to provide information about dependence, *and* we ought not to claim (or imply) that the explanandum depends on something that it does not. Thus, the independence from the actual laws seems insufficient in itself to motivate the view that abstract 'mathematical' explanations are *sui generis* in the way Lange regards them.

Both Lange and Pincock are contrasting their views to causal accounts of explanation. It is natural, therefore, to take them to provide a competing account of the worldly source of explanatory power.[16] So far we have undermined broad motivations for thinking that abstract explanations require a source of explanatory power that is not centred on dependence. Our view is that two of the three dimensions of abstraction—even though undeniably exhibited by abstract explanations—do not actually provide a source of explanatoriness.[17] The problem is that these dimensions of abstraction do not pro-

[15]Lange's motivations are best understood in the broader context of his related views on other kinds of non-causal explanations (e.g. Lange [2015]), and on the modal metaphysics of laws of nature in general [Lange, 2009]. Like other philosophical views, Lange's account of abstract explanations can gain indirect support from being a coherent part of a 'bigger picture'. Here we do not wish to assess the pros and cons of this bigger picture; we will focus, rather, on what can be said to directly motivate (or otherwise) Lange's association of explanatoriness with the second dimension of abstraction in connection with the paradigmatic examples of abstract explanations.

[16]Causal accounts typically also require relevance constraints (cf. Strevens [2008]).

[17]This is *not* to say that they can never be relevant to *any* aspect of having a good explanation. To only provide relevant information (and to not include irrelevant information) is a shared commitment of any account of explanation. No matter what the source of explanatory power is taken to be. This goes some way towards explaining the intuition that dimension one and two provide a separate source of explanatory power. For example, if we assume that we begin with a mistaken belief that the island taken as a starting point was relevant to explaining the failure to complete a round-tour of Königsberg, then it would have been informative and enlightening to find out that the starting island is irrelevant. However, a dependence account can easily capture this intuition. The failure to complete a round-tour of Königsberg does not depend on the starting island. A good explanation cannot cite misinformation about the dependences. Note that the intuition that this kind of information about independence is a source of explanatory power disappears once we are not focusing on a case of correcting a mistaken belief about dependences. We are

vide information about explanatory dependence. In the rest of the paper we will discuss this notion of explanatory dependence and how it connects to the third dimension of abstraction.

But before we move on, let us address a natural worry. It may seem that dependence and independence are conceptually so closely connected that we cannot cleanly separate the two. In particular, any information about independence is also information about dependence: 'E is independent from A' means that 'E does *not* depend on A', and if we assume that an explanandum E depends on something, then 'E is independent from A' entails (and gives us the information) that 'E depends on something else than A'. It is thus the case, the worry goes, that information about independence along every dimension of abstraction can provide information about dependence.

In response, we simply note that it is *not* true that information about independence ipso facto provides information about dependence that is *explanatory*—that is, dependence information in the W-H sense. It is undeniably true that one provides what-if-things-had-been-different information, broadly speaking, by showing that even if the laws of nature were different Mother could not divide twenty-three strawberries equally among three children. (What if things had been different with respect to laws of nature? Then Mother would have failed just the same!) But this is *not* the sort of what-if-things-had-been-different information that counts as explanatory in the counterfactual framework that we favour, according to which only what we call *change relating* counterfactual information counts as explanatory.[18] To understand Mother's failure, we need to be able to say under what conditions Mother could have succeeded. We now move on to discuss this framework.

4 Abstraction in a counterfactual framework

The starting idea of the W-H account is that explanation 'is a matter of exhibiting systematic patterns of counterfactual dependence' (Woodward [2003, 192]). To develop this into a theory of explanation, we need to say more precisely exactly *which* patterns of counterfactual dependence matter for explanation. On the W-H account this is done by focusing on modal information that allows one to answer questions about how the explanandum would have been *different* under a special type of change in the explanans [Woodward, 2003, 192]. The changes in the explanans that are relevant are those that result from interventions on the explanans variable(s) (with respect to the explanandum variable). The notion of intervention is technical; it plays the role of ruling out changes in the explanans variable that are brought about by (i) changes in the explanandum variable, and (ii) changes to some other ('common cause') variable that changes the explanandum variable independently of the explanans variable.

This approach has a number of attractive features. First, it is natural to take relations (whether these are nomological, causal, or mathematical) that are explanatory to be ca-

not tempted to regard as explanatory the information that failure to complete the tour is independent of the fact that the bridges are made of stone and not wood. We never mistakenly believed that the failure to complete the tour depended on these features.

[18] In the terminology of Woodward and Hitchcock [2003a,b] an explanatory generalisation must be invariant under at least one *testing* intervention.

pable of providing some sort of modal information. Second, by requiring more specifically that the modal information relates to different possible states of the explanandum—how the actual explanandum would have been different had the explanans been different—the account captures the natural idea that explanatory information is information about worldly dependences. In the W-H account the critical modal notion of dependence gets carefully cashed out in causal terms through the notion of a (testing) intervention. This allows one to rule out backtracking counterfactuals—for example, 'had the period of the pendulum been different, then the length of the pendulum would have been different'—as providing the right dependences. The focus on what-if-things-had-been-different questions that concern *changes* in the target explanandum also provides a clear contrast to mere subsumption under modal regularities in the spirit of the DN account of explanation. Third, this focus on interventions squares well with much of our experimental practice; we often manipulate systems in order to discover information about their explanatory causal structures. Fourth, the account is, as Woodward and Hitchcock [2003a,b] discuss at length, very well suited to capture not only what it takes to have an explanation, but also what makes an explanation better or worse. Having an explanation is a matter of having the right information about the dependences, and *the more the better*.

The W-H account separates an explanans into two parts: a specification of an invariant explanatory generalisation and a specification of the actual values of the variables in the explanatory generalisation.[19]

For a simple illustration, let us look at the explanation of the period of a simple gravitational pendulum. Let us assume that we want to explain the fact that the period T takes the value t_1 (our explanandum M) by using the simple pendulum law. In terms of the W-H account, we have an explanatory generalisation $T \approx 2\pi\sqrt{\frac{l}{g}}$. The explanans consists of this generalisation and in a specification of the actual values of l and g. In order to have an explanation on the W-H account, the simple pendulum law has to (at least approximately) correctly give the actual value of the period as t_1 under an intervention that fixes the values of l and g to the actual values. Moreover, the simple pendulum law must capture (in at least one case) how the period would *change* under at least one (and ideally more) interventions that change(s) the length or the gravitational acceleration. It thus captures the *dependence* of the explanandum on these variables.

Now we are in a position to make the contrast between the different dimensions of abstraction more precise. Let us take the first dimension of abstraction first. In the context of the simple pendulum explanation, we can ask whether the explanandum $T = t_1$ is independent of particular background conditions which are in some sense part of the system but not represented as variables in the explanatory generalisation $T \approx 2\pi\sqrt{\frac{l}{g}}$. We can think of the microphysical features that make the pendulum cord inextensible, for example, or the mass of the bob, and so on. The degree of *independence* of the explanation from such background conditions was the focus of the first dimension of abstraction. (In contrast to such 'internal' background conditions there are other background conditions that are external to the system and clearly irrelevant, such as, the bob being made in Japan, the exchange rate of the US dollar to the Malaysian ringgit, the

[19] See Woodward [2003, 203] for a detailed account.

colour of the shirt of the person setting the bob in motion, etc.)

In the second dimension of abstraction, we considered the degree of independence of the explanation from the actual laws of nature. In the context of the pendulum example, we can ask whether we can vary the actual laws of nature and still expect the simple pendulum explanation to apply. We can alter many of the laws of, for example, electromagnetism, without affecting the explanation, provided that those alterations do not run afoul of the assumptions presupposed in the context of the why-question, for example, the cord being nearly inextensible. As mentioned earlier, we can also alter many of the fundamental aspects of gravitational acceleration, such as whether its nomological basis is an action at a distance effect or a manifestation of curved space-time.

Finally, in the third dimension of abstraction, we consider the range of conditions directly relevant to changes in the variables in the explanatory generalisation, $T \approx 2\pi\sqrt{\frac{l}{g}}$ such that the simple pendulum explanation of the period holds. For example, we could decrease g by, say, moving the pendulum to a higher altitude or to the moon, and we could change the length of the pendulum, and yet, the explanation of the period would work in just the same way. On the W-H account, abstractness in this dimension is just a matter of the invariance of the explanatory generalisation, 'measured' by the range of alternative values of variables l and g for which the generalisation holds and is change relating.

So far we have simply described the W-H account. Before we can apply these ideas to the paradigmatic examples of abstract explanations, we need to extend and refine some of them. In particular, the W-H account was developed as an account of *causal* explanation, while the abstract explanations we are interested in are plausibly non-causal. As we move beyond causal explanation, we can no longer appeal to the exact understanding of the relevant class of explanatory counterfactuals that Woodward and Hitchcock use. They delineate the class of relevant counterfactuals in terms of the causal notion of intervention: explanatory counterfactuals describe the effect of a testing intervention. These counterfactuals are underwritten by contingent explanatory generalisations, and the effects of surgical testing interventions are naturally understood as concerning causal dependences. In the non-causal cases this notion of causal dependence becomes inapplicable or unnatural, because the notion of causal intervention is inapplicable, or because the counterfactuals are not underwritten by contingent laws of nature (but, rather, by logic or mathematics, or metaphysical truths).[20] Is it still possible in these cases to characterise explanatory dependences in counterfactual terms?

The answer is yes; we have analogous explanatory dependences in the non-causal situations involving Mother's strawberries, Königsberg's bridges, etc.[21] We have a good

[20] In Mother's case, for example, changing the number of strawberries may count as an intervention. But its 'effect'—rendering the set divisible-by-three, say—is not causally related to this intervention, but is rather more intimately connected to it. In the Königsberg case, it is not clear to us exactly which changes of bridge configuration even count as interventions, to begin with.

[21] [Woodward, 2003, 221] notes passingly that it seems natural to extend the account to non-causal explanations, but he does not develop it further.

> [T]he common element in many forms of explanation, both causal and noncausal, is that they must answer what-if-things-had-been-different questions. ...When a theory or derivation answers a what-if-things-had-been-different question but we cannot interpret this as

grasp of what it means to change the explanans, and how the explanandum depends on these changes (but not vice versa), regardless of whether or not the notion of causal dependence is applicable. This is enough to apply the central idea that the relevant counterfactuals are those that (typically asymmetrically) relate changes in the explanans to changes in the explanandum. For example, in the case of Mother we know how changes in the number of strawberries changes the system's divisibility-by-three, and consequently the possibility of Mother's success or failure. We thus grasp how Mother's success or failure explanatorily depends on the number of strawberries that she has (and the number of children), but not the other way around. This dependence is appropriately directed. As we indicated earlier, the notion of intervention plays the role of blocking changes to the explanans that go through changes to the explanandum (or through a common cause variable, etc.). Similarly, changes to the number of strawberries clearly do not go through changes to the system's divisibility-by-three, or through changes to Mother's success or failure. However, changes to Mother's success or failure do go through changes to her number of strawberries (or children). From now on, let us call the explanatory counterfactuals simply *change relating counterfactuals* (adapting Woodward's use of the term) to keep in mind that they are a strict subset of all counterfactuals and a broader class than interventionist counterfactuals.[22]

Since the relevant counterfactuals are selected by focusing on *changes* to some object or system, the modal information that we are interested in is tied to the system or object in question. They are what Woodward calls "same object" counterfactuals":

> Suppose that we wish to explain the behavior of some object or system o. As the standard view is usually understood, it claims that generalizations of form "All As are Bs" are explanatory of the behavior of o if they support counterfactuals of the following form: ... If some object o^*, different from o and that does not possess property A, were to be an A, then it would be a B.
>
> Call such counterfactuals "other object" counterfactuals: they describe what the behavior of objects other than o would be under the counterfactual circumstances in which they are A. By contrast, according to the view I have been defending, to count as invariant and hence explanatory with respect to o, a generalization must support "same object" counterfactuals that describe how the very object o would behave under an intervention. [Woodward, 2003, 281]

At first glance this may seem to suggest that all of the explanations that we have in mind must be about specific, *particular* systems or objects in order for the distinction—central to the W-H account—between "same object" (SO) and "other object" (OO) counterfactuals to apply. This raises a worry that according to Pincock [2015] rules out the

an answer to a question about what would happen under an intervention, we may have a noncausal explanation of some sort.

[22] For more discussion on how to decouple the counterfactual aspect and the causal aspect of Woodward's account, see for example Saatsi and Pexton [2013], Saatsi [forthcoming], and Rice [2015]. For a suggestion of how to recover directionality without causal notions see for example Jansson [forthcoming].

application of the W-H account to abstract explanations, since some highly abstract explanations do not seem to provide information about any particular object at all (much less information about changes to any particular object).

In our view, one should not think of the contrast between SO counterfactuals and OO counterfactuals in the way Pincock does. First, note that even in the case of causal explanations, most scientific explanations do not have a particular, individual object as the explanatory target. The explanandum is typically generic (and general even when couched in language such as 'the period of a simple pendulum'). The importance of the distinction between SO and OO counterfactuals lies in the fact that it focuses our attention on the right kind of *change relating* counterfactuals. When we make use of the simple pendulum law to explain the period of a simple pendulum, the right counterfactuals to have in mind are those that ask how the period of a generic kind of dynamical system, viz. simple pendulum, is affected by changing, say, its length.[23] Counterfactuals concerning objects other than simple pendulums do not come into play, even when they are reasonable and well defined. (It may be true that had a hammer been suspended around an appropriate pivot with the head down, then we could have used the simple pendulum law to explain its period, but such OO counterfactuals are not required in order to explain the period of a simple pendulum.)

Although the W-H account does not apply directly to the case of Königsberg's bridges, the central notion of change relating (SO) counterfactuals can be carried over to this case. This is the work of the next section.

5 Königsberg – Encore!

We will now revisit the Königsberg case to illustrate the above counterfactual account and its virtues. So far we have mainly criticised Pincock's and Lange's motivations for thinking that there is a sharp distinction between causal explanations and abstract non-causal explanations, maintaining that one can instead approach both types of explanations in the fundamentally same spirit. We now push for a stronger claim, arguing that the core claim of the counterfactual account (as we understand it), that explanatoriness is associated exclusively with the third dimension of abstraction, can be *tested* against the alternative viewpoints. This will provide a clear reason to prefer the counterfactual account.

Let us go back to the 18th c. Königsberg, and ask Q_K: Why is it impossible to make a round-tour of Königsberg crossing each of its seven bridges exactly once? An intuitive explanation-sketch response goes as follows. Clearly each visit of a landmass ('island') requires the use of two bridges: one in, and one out. Else you get stuck. Therefore, in order for a network of bridges to allow for a round tour—to be 'tourable'—each island must have a number of bridges to/from it that is some multiple of two. On the other hand, if there is one (or more) island(s) with an odd number of bridges to/from it, the system is not tourable. 18th-century Königsberg had four such troublesome junctions, rendering a round-tour of Königsberg impossible.

[23] We can then go further to apply this to explain the period of a particular pendulum that is of the right kind, e.g. that in my grandfather clock.

Euler initiated a famed graph-theoretic explanation that makes the above sketch precise. This is standard material in graph-theory textbooks, illustrating the explanatory use of mathematical notions (such as *connected graph*, its *vertices*, and their *degrees*).[24] But before we get to the graph-theoretic explanation in its full generality, it is worth attempting to answer Q_K *without* graph theory. In the counterfactual framework explanations must involve an invariant, change-relating generalisation that supports counterfactuals indicating an explanatory dependence of the explanandum on the explanans. A generalisation that thus underwrites an answer to Q_K could be markedly less abstract and less general than the graph-theoretic explanation, without thereby being unexplanatory.

As a matter of fact it is easy to find a simple invariant generalisation that furnishes a non-mathematical answer to Q_K. Let us focus our attention on the kind of bridge system that connects *exactly* four islands, with *at most* two bridges between any two islands. Euler's Königsberg (represented below) is one of these systems.

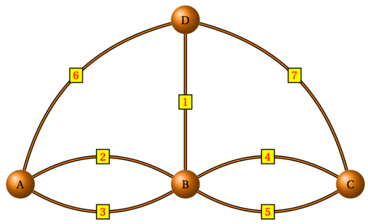

Figure 1: Königsberg's bridge system.

There are 395 such bridge systems altogether. We can classify them as follows. Call a bridge system *'even'* iff each island has an even number of bridges to/from it. In this particular case, this means that each island must have 2, 4, or 6 bridges to/from it. If a bridge system is not even, call it *'odd'*. Considering the specific type of bridge system exhibited by Königsberg, we ask: why is it not tourable? In answering this question we naturally look for an explanatory generalisation capable of providing suitable modal information by supporting appropriate, change relating counterfactuals. Focusing our attention, for now, only on the bridge systems consisting of four connected islands with at most two bridges between any two islands, a fitting generalisation is not hard to find: of all these bridge systems, all and only the *even* ones are tourable.[25] This is a *true* generalisation about this set of 395 different types of bridge systems. It is also a generalisation that can afford us with a degree of explanatory purchase on Q_K. That is, by reference to this generalisation we can begin to answer Q_K simply by noting that Königsberg's bridge system is not tourable, because it is not *even*; it would be tourable, if it were *odd*. Königsberg's tourability—the feature that is our explanandum—*depends on* this high-level physical property of the system.

[24]It is unsurprising that many authors refer to this explanation as a paradigmatic case of a (distinctly) mathematical explanation of an empirical fact. In addition to the authors already covered, see e.g. Lyon [2012].

[25]This fact about these bridge systems can be in principle established by a variety of means, e.g. by attempting to draw every such system without lifting your pen, or by playing with a comprehensive collection of miniature models of such systems. Getting epistemic access, or representing that fact, in principle need not involve mathematics. (cf. Saatsi [2011])

More formally, we can define the following binary variables X and Y:

$$X = \textit{even/odd} \text{ system}$$
$$x = 1 : \textit{even}$$
$$x = 0 : \textit{odd}$$

$$Y = \text{(non-)tourable system}$$
$$x = 1 : \text{tourable}$$
$$x = 0 : \text{(non-)tourable}$$

The simple explanatory generalization at stake then states that for bridge systems of this kind—four islands, maximum of two bridges between any pair—it holds that

$$X = Y$$

This equation should be read left to right, as indicating an asymmetric dependence of Y-variable on the X-variable. The tourability (or otherwise) of a bridge system depends on it being *even* (or *odd*). Being *even* (or *odd*) does *not* depend on tourability; it only depends on the number of bridges. All in all, the explanation naturally fits the counterfactual framework presented above (§4).

The explanatory generalisation, while narrow, is explanatory nevertheless, even if minimally so. Undoubtedly, the explanation provided immediately raises further questions. Why exactly is this generalisation true, for example? What if we start adding or subtracting 'islands'? What if we relax the restriction that there are at most two bridges between any pair of islands? These are obvious further questions, and this clearly renders the explanation shallow and somewhat contrived, especially in comparison to a full-blown graph-theoretic account. But none of this diminishes the philosophical significance of this explanation. For however minimal and shallow the toy explanation is, we deem it to have *some* explanatory power nevertheless, and we maintain that this is due to the explanation providing modal information of the right sort.[26]

The shallow explanation above clearly has a degree of abstractness along the three dimensions introduced in §2. Indeed, the explanation is *highly* abstract with respect to the different specific material realizations of the bridges (the first dimension), and also with respect to the underlying laws of physics (the second dimension). As a matter of fact, the explanation is as abstract along these dimensions as the full-blown graph-theoretic explanation! Yet the explanation is shallow. This clearly speaks against the idea

[26]Is minimal explanation of this sort ever completely *satisfactory*? That depends on the context of the why-question. Imagine Königsberg without Euler—or any other mathematician for that matter. The King of Königsberg, annoyed by his failure to do an round-tour of the city, wishes to understand the situation. "This city must be made thus tourable, at whatever cost!", he commands, putting his best people to work. "Well, actually, the less it costs the better—and we certainly cannot afford more than two bridges between any two islands!" The King's minions get to work, and painstakingly demonstrate, by non-mathematical means, that for any financially feasible set-up of bridges it holds that it is tourable if and only if it is *even*. Equipped with this knowledge, they can explain the situation to the King: "What ever you do with the bridges, make sure the whole system is *even*, and you will be able to tour it, since tourability depends on this feature alone (at least for the kinds of bridge systems we can afford)."

the (minimal) explanatoriness in question springs from abstraction along these lines, as Pincock and Lange would have it.

With regard to the third dimension of abstractness, the explanatory generalisation does allow us to answer some change relating what-if-things-had-been-different questions. We have counterfactual information of the right, SO, kind. We can answer counterfactual questions about how the tourability of a generic kind of structural system, viz. bridge system, is changed by changes to the oddness or evenness of the system (at least as long as we stay within the constraints of four connected islands and a maximum of two bridges between any two islands). This is what makes it explanatory, even if shallowly so. Yet, its degree of abstractness along this dimension is very limited. Although the explanatory generalisation never delivers the *wrong* answer, the generalisation is simply silent on what happens in cases of more than four islands or more than two bridges between some islands. Thus, the explanation using the generalisation breaks down in these cases.[27]

In order to understand precisely how the shallow explanation compares in its abstractness to the deeper graph-theoretic explanation, we need to pay attention to the fact that the variable X concerns a *determinable* property of the system: it being *even* or *odd*. This determinable property is determined by the bridges' configuration. There are various ways for a bridge system to be *even*. One way is for each island to have (say) 4 bridges to/from it. This is still a determinable property of the system, determined by a specific way of having 4 bridges to/from each island. (Cf. Figure 2)

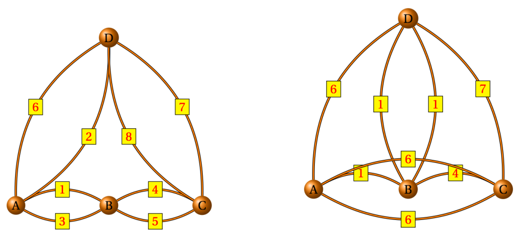

Figure 2: Two determinate 4-4-4-4 configurations.

Here is another way for a bridge system to be *even*: one of the islands has 8 bridges to/from it, and the other three islands have 2, 2, and 4 bridges to/from them, respectively. While our generalisation supported some explanatory what-if-things-had-been-different questions, it is simply silent on the system's tourability if we change X from *odd* to *even*

[27] Here two differences between our third dimension of abstraction and (even the extended) notion of invariance in Woodward and Hitchcock [2003a,b] is important. First, we take the important question to be whether or not the variations destroy the explanation (not merely the generalisation). The generalisation itself does not break down. It just does not apply. However, the generalisation not applying destroys the explanation. Although we do not share their accounts in general, we take this focus on the whole explanation from Potochnik [2010] and Weslake [2010]. Second, the range of variations are understood as the range of cases in which one of the variables in the explanans can be changed without destroying the explanation. While this is also the idea of the W-H account, for the discussion in Woodward and Hitchcock [2003a,b] it is often natural to interpret the range of invariance as the range of changes in the *variable values* for which the generalisations continue to hold. This can come apart from the range of cases of changing the value of the variable under which the generalisation (or in our case, the explanation) continues to hold. Our case above illustrates this.

by allowing one island to have as many as 8 bridges to/from it. Similarly, it is silent on what happens if we change X from *even* to *odd* by including 7 bridges to/from some island. To answer what-if-things-had-been-different questions corresponding to these cases, we need a broader generalisation, $X = Y$, that applies to systems as rich in bridges as these.

Differences in the determinate configurations do not explicitly feature as a variable in the explanatory generalisation $X = Y$. Indeed, on the face of it the explanatory generalisation looks the same regardless of the range of determinate what-if-things-had-been-different questions it is taken to support. Differences in the range of determinate realizations of evenness/oddness do get into play through the restriction in application. When we consider *changes* to the variable (*even*, *odd*) these changes have to go through changes in the bridge system (there is no way to change the evenness of the system that does not go through changing the specific configurations of the bridges). If we are justified in taking the generalisation $X = Y$ to apply to systems that have (say) a maximum of three bridges between any two islands, then our explanation covers a wider range of conditions directly relevant to changing the value of the explanans variable X. In particular, we can now also consider what-if-things-had-been-different situations with 7, 8 or 9 bridges to/from some island(s).

In this way we can straightforwardly compare different Königsberg explanations with respect to their degree of abstraction in the third dimension. The explanation where we restrict the generalisation to a maximum of three bridges between any two islands applies to all the ways of varying the variable in the explanatory generalisation that the explanation restricted to a maximum of two bridges covers—and then some! The full-blown graph-theoretic explanation is, of course, maximally abstract along this third dimension of abstractness. The explanatory generalisation now covers any (connected) bridge system of arbitrary many islands and bridges. The graph-theoretic explanation has considerable depth in contrast to the shallow, non-mathematical explanation. This is solely due to increased abstraction along the third dimension. The explanatory dependence of the shallow explanation is subsumed under a more general explanatory dependence, enabling us to answer a much wider range of change relating *what-if-things-had-been-different* questions. The shallow explanation does not contain irrelevant information about the nature of the bridges, their material constitution or their length, say, or about the underlying nomological features, such that the increase in explanatory depth could be due to abstracting away from such information. Indeed, as already noted, the explanans of these two explanations are already maximally abstract along the first and second dimensions, and the considerable increase in the explanatory power should be attributed solely to the third dimension of abstraction.

6 Conclusion

We have argued that paradigmatic abstract (plausibly) non-causal explanations can be naturally accommodated with an account that associates explanatoriness with suitable information about dependence. This improves the prospects of subsuming many explanations under a unified counterfactual framework, opposing the current trend that emphasizes the explanatory value of abstraction as a sui generis source of explanatoriness. We found the motivations for this trend questionable (§2), leaving room for a more unified account.

More unified theories are in general often better, ceteris paribus, but we are not just expressing this kind of a prima facie preference for unification. Rather, we motivate the unified account by the following considerations. First, we argued that the unified account better captures intuitions about 'explanatory depth'. Part of the force behind the W-H framework comes from the fact that it provides a natural starting point for capturing intuitions about explanatory goodness (Woodward and Hitchcock [2003a,b] and Ylikoski and Kuorikoski [2010]). We argued, in the same spirit, that our viewpoint can be tested by varying dependence information—the hypothesized source of explanatoriness—keeping other things fixed (§5). We maintain that our intuitions about radically varying explanatory power naturally correspond to radically varying amounts of dependence information, in a way that is difficult to accommodate from the alternative viewpoints.

Secondly, the unified account has a further virtue worth flagging. This has to do with *dissolving* difficult questions facing the more disjunctive accounts that regard abstract explanations fundamentally different from causal explanations in their explanatoriness. Noting that causal explanations also typically abstract away from a huge amount of physical detail raises a question about Pincock's point of view, for example: What explains the distinct qualitative difference between causal explanations that exhibit a degree of abstraction along each of the three dimensions, on the one hand, and the *sui generis* 'abstract explanations', on the other? Is there a 'threshold' of abstraction above which the counterfactual conception fails, despite capturing abstract causal explanations so well?

A similar question can be raised for Lange's account. Many causal explanations also abstract away from a huge amount of underlying nomological detail.[28] Thus, all causal explanations that incorporate a degree of abstraction along the second dimension of abstraction also show how the explanandum is, to a corresponding degree, 'necessary to a stronger degree of necessity' by virtue of showing the irrelevance of some of the actual laws involved. So why is it that such modal information becomes explanatory *in a sui generis way* in connection with abstract non-causal explanations? Or is it the case that such modal information always contributes to the explanatory power, but it contributes in a different way when an explanation abstracts away from (almost) *all* nomological information, as in the case of 'mathematical' explanations?[29] Whence the difference? Why do abstract non-causal explanations explain so differently—in Salmon's 'modal' mode—from causal explanations that incorporate a degree of similar abstraction?

Such questions have not been considered, never mind answered, by Pincock and

[28] Recall the discussion in §2 of the gravitational pendulum explanation, which does not depend on the actual nomological structure underlying 'gravitational force'. (A world with action-at-a-distance Newtonian gravitational force has a different law governing that force, from general relativistic world with no such force.) Also, a causal equilibrium explanation of why a ball ends up at the bottom of a bowl can abstract away from the specifics of the actual gravitational force law, and even from the specific proportionality of force and acceleration in Newton's second law. And so on.

[29] *Almost* all, since these explanations still contain information associated with the various contingent facts presupposed by the why question.

Lange. A more unified account has the virtue of allowing us to side step these issues entirely. According to this account the paradigmatic abstract explanations, despite their non-causal character, are explanatory for the fundamentally same reason as causal explanations are. In both cases the explanatory power springs from counterfactual information of the same sort, and the paradigmatic exemplars of abstract explanations need not be regarded as *sui generis*.

Appendix

At the start of section §2 we cited four exemplars of abstract, plausibly non-causal, mathematical explanations that have featured prominently in the philosophical literature. We have covered two of these in depth in the main text; here we briefly cover the remaining two.

Economical bees. Why do bees build hexagonal honeycombs? Answer: because it is the most resource-efficient way to build them. That is, there is an evolutionary fitness advantage associated with the neural architecture responsible for hexagonal honeycombs. This has to do with the geometrical fact that hexagonal grid uses the least total perimeter in dividing a planar region into regions of equal area. Lange [2013, 500] notes that "this explanation works by describing the relevant features of the selection pressures that have historically been felt by honeybees, so it is an ordinary, causal explanation, not distinctively mathematical." (It is perhaps worth noting further that this causal explanation is quite abstract in dimensions one and two: the causal facts described are rather independent of the actual microphysical structures of honeycombs, as well as the microphysical causal laws.)

Lange [2013, 500] goes on to specify a different, related explanandum that has a non-causal ('distinctively mathematical') abstract explanation:

> But suppose we narrow the explanandum to the fact that in any scheme to divide their combs into regions of equal area, honeybees would use at least the amount of wax they would use in dividing their combs into hexagons of equal area (assuming combs to be planar regions and the dividing walls to be of negligible thickness). This fact has a distinctively mathematical explanation: it is just an instance of the Honeycomb Conjecture.

That is, the question is: why does the hexagonal grid *minimize* the total amount of material needed to cover the perimeter? According to Lange this is explained as an 'immediate application' of the mathematical fact stated by the Honeycomb Conjecture, showing the explanandum to be mathematically necessary—independent of any dynamical and causal laws involved (e.g. by governing the materials that bees use and how they use it, or shaping the actual evolutionary trajectory).

We agree that there is an abstract, plausibly non-causal explanation here, but our analysis of the (minimally) explanatory application of the Honeycomb Conjecture is different from Lange's, as follows. The amount of wax needed is clearly directly proportional to the volume to be covered. (This just follows from the meaning of 'amount' in this context.) The volume to be covered is clearly directly proportional to the total perimeter of the comb-dividing walls. (We can assume this in the context of the question as presented.) It is a geometrical fact about honeycomb-like systems that they minimize the total perimeter (and thus the volume required for construction) if and only if they divide the planar regions into regular hexagons. (The Honeycomb Conjecture represents this fact about physical space.) The amount of wax needed for construction depends on the geometry of such systems. The dependence is plausibly not causal. Having the property of minimizing the amount of wax needed for construction, in particular, *depends on* the hexagonal structure. We explain by providing information about this dependence.

The explanatory generalisation involved is very simple. Let's fix a range of possible honeycomb structures by specifying the average comb area, as well as the total area covered by the whole honeycomb. Then it is the case that a (regular) hexagon honeycomb requires some specific amount of wax, V, and it is also the case that any alternative honeycomb requires more wax than this. This explanatory generalisation answers a range of what-if-things-had-been-different questions regarding the amount of wax required, by effectively relating two explanans variable values (*hexagonal* vs. *non-hexagonal*) to a binary explanandum variable (*minimizing* vs. *non-minimizing*). This generalisation allows us to answer a huge range of what-if-things-had-been-different questions, since the explanans variable *non-hexagonal* concerns a determinable property that corresponds to an infinite number of determinate non-hexagonal shapes. There is thus a sense in which the explanation is highly abstract in the third dimension of abstraction. The explanation nevertheless feels quite minimal, because there is no interesting pattern of dependence (unlike in the Königsberg case).

Similar abstract explanations are very easy to come by. Why are soap bubbles (in still air) spherical? Because the molecular dynamics of soap-film is area-minimizing: bubbles tend towards a shape that minimizes the area containing the air inside. But why is it the *spherical* soap bubble, of all possible soap bubbles, that minimizes the volume-containing surface area? Taking the relevant geometrical facts to be appropriately represented in Euclidean geometry by the isoperimetric inequality in three dimensions explains this by revealing how soap bubbles' property of minimal surface-area-to-volume ratio *depends on* their spherical shape.

Plateau's laws. There are much fancier facts about soap bubbles, of course, that we can take as the explanandum. We might, for instance, focus on Joseph Plateau's experimentally established laws for any configuration of soap bubbles, with or without a rigid frame. One of the laws says, for example, that there are only two ways for soap films to intersect: three surfaces can intersect along a curve, meeting at equal angles of 120° degrees, or four surfaces can intersect at a point, meeting at around 109° degrees. What explains this regularity of all 'Plateau configurations'?

There is no deep difference between the explanation of this regularity, and the explanation of the regularity concerning the approximately spherical shape of isolated soap bubbles. Both regularities are a result of the area-minimizing tension of the soap-film. We can then further ask why the bubble configurations that do minimize the surface area exhibit the geometrical regularities in question, such as Plateau's law above. The answer to the further question is an abstract, non-causal explanation, revealing how the property of minimizing the surface area *depends on* the configuration geometry ('shape'), partly specified by the intersection types and angles. The explanatory regularity is again simple: a binary explanandum variable (*area-minimizing* vs. *non-minimizing*) depends on the value of a determinable explanans variable (*Plateau-configuration* vs. *non-Plateau-configuration*). This is a true generalisation about soap bubbles (and other area-minimizing systems) that supports an explanation of Plateau's laws by answering a range of change relating what-if-things-had-been-different questions. What if things were different so that the surface intersection angles did not comply to Plateau's law? Then the configuration would not minimize the surface area. Given that the physics of soap molecules causes the bubbles to minimize their area, Plateau's laws are thus explained as a (non-causal) consequence of area-minimization.[30]

References

Frederick Almgren and Jean Taylor. The Geometry of Soap Films and Soap Bubbles. *Scientific American*, 235(1):82–93, 1976.

Robert Batterman. On the Explanatory Role of Mathematics in Empirical Science. *British Journal for the Philosophy of Science*, 61:1–25, 2010.

Robert Batterman and Collin Rice. Minimal Model Explanations. *Philosophy of Science*, 81:349—376, 2014.

David Braine. Varieties of Necessity. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 46:139–70, 1972.

Jan Faye. *The Nature of Scientific Thinking : On Interpretation, Explanation, and Understanding*. Palgrave Macmillian, 2014.

Lina Jansson. Explanatory Asymmetries: Rehabilitating Laws of Nature. *Journal of Philosophy*, forthcoming.

Marc Lange. *Laws and Lawmakers: Science, Metaphysics, and the Laws of Nature*. Oxford University Press, 2009.

Marc Lange. What Makes a Scientific Explanation Distinctively Mathematical? *British Journal for the Philosophy of Science*, 64(3):485–511, 2013.

Marc Lange. How to Explain the Lorentz Transformations. In S. Mumford and M. Tugby, editors, *Metaphysics and Science*. Oxford University Press, Oxford, 2015.

Aidan Lyon. Mathematical Explanations of Empirical Facts, and Mathematical Realism. *Austrasian Journal of Philosophy*, 90(3):559–578, 2012.

[30]As far as we can see, this is in complete harmony with the relevant experts' own assessment, who summarise the importance of their work as having *demonstrated* that "the area-minimizing principle alone is sufficient to account for the overall geometry of soap films and soap bubbles." [Almgren and Taylor, 1976, 93] (Cf. Pincock [2015])

Aidan Lyon and Mark Colyvan. The Explanatory Power of Phase Spaces. *Philosophia Mathematica*, 16(3):227–43, 2008.

Christopher Pincock. A Role for Mathematics in the Physical Sciences. *Noûs*, 42:253–275, 2007.

Christopher Pincock. On Batterman's 'On the the Explanatory Role of Mathematics in Empirical Science'. *The British Journal for the Philosophy of Science*, 62:211–217, 2011.

Christopher Pincock. Abstract Explanations in Science. *British Journal for the Philosophy of Science*, 66(4):857–882, 2015.

Angela Potochnik. Levels of Explanation Reconceived. *Philosophy of Science*, 77(1):59–72, 2010.

Collin Rice. Moving Beyond Causes: Optimality Models and Scientific Explanation. *Noûs*, 49(3):589–615, 2015.

David-Hillel Ruben. *Explaining Explanation*. Routledge, 1990.

Juha Saatsi. The Enhanced Indispensability Argument: Representational versus Explanatory Role of Mathematics in Science. *The British Journal for the Philosophy of Science*, 62(1):143–154, 2011.

Juha Saatsi. On Mathematics' 'Indispensable Explanatory Role'. *Mind*, forthcoming.

Juha Saatsi and Mark Pexton. Reassessing Woodward's Account of Explanation: Regularities, Counterfactuals, and Noncausal Explanations. *Philosophy of Science*, 80(5): 613–624, 2013.

Wesley Salmon. Scientific Explanation: Three Basic Conceptions. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume Two: Symposia and Invited Papers (1984):293–305, 1985.

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

Bas van Fraassen. The Pragmatics of Explanation. *American Philosophical Quarterly*, 14 (2):143–150, 1977.

Bas van Fraassen. *The Scientific Image*. Oxford: Clarendon Press, 1980.

Brad Weslake. Explanatory Depth. *Philosophy of Science*, 77:273—294, 2010.

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

James Woodward and Christoper Hitchcock. Explanatory generalizations, part I: A counterfactual account: A counterfactual account. *Noûs*, 37(1):1–24, 2003a.

James Woodward and Christoper Hitchcock. Explanatory generalizations, part II: Plumbing explanatory depth. *Noûs*, 37(2):181–199, 2003b.

Petri Ylikoski and Jaakko Kuorikoski. Dissecting Explanatory Power. *Philosophical Studies*, 148:201–219, 2010.