Brit. J. Phil. Sci. **0** (2020), 1–28

Concrete Scale Models, Essential Idealization, and Causal Explanation

Christopher Pincock

ABSTRACT

This paper defends three claims about concrete or physical models: (i) these models remain important in science and engineering, (ii) they are often essentially idealized, in a sense to be made precise, and (iii) despite these essential idealizations, some of these models may be reliably used for the purpose of causal explanation. This discussion of concrete models is pursued using a detailed case study of some recent models of landslide generated impulse waves. Practitioners show a clear awareness of the idealized character of these models, and yet address these concerns through a number of methods. This paper focuses on experimental arguments that show how certain failures to accurately represent feature *X* are consistent with accurately representing some causes of feature *Y*, even when *X* is causally relevant to *Y*. To analyse these arguments, the claims generated by a model must be carefully examined and grouped into types. Only some of these types can be endorsed by practitioners, but I argue that these endorsed claims are sufficient for limited forms of causal explanation.

doi:10.1093/bjps/axz019

1 Introduction

Some philosophers of science have recently emphasized that model-based science is helpfully distinguished from a more direct investigation or experimentation with target systems.[1] This paper considers and resolves a worry about a certain kind of target-directed modelling. The worry arises when scientists are focused on a repeatable type of target system which, following Bogen and Woodward ([1988]), I will call a phenomenon. When a model is devised for a phenomenon, the scientists using the model often believe that the model is misrepresenting the phenomenon in some ways. In certain cases, this sort of misrepresentation is unavoidable. As I clarify in Section 2, the model is thus essentially idealized. The worry about essentially idealized target-directed models is that they do not seem to be able to extend our knowledge of phenomena. For this sort of model to extend our knowledge of the phenomenon, its misrepresentations would have to be somehow identified and corrected. But in model-based science we typically lack the prior knowledge of the phenomenon that would allow this sort of correction. The upshot of these considerations is the conclusion that model-based science is not actually able to extend our knowledge. At best, modelling is a heuristic source for possible features of phenomena. The worry seems especially clear for explanations of features of the target phenomenon. An essentially idealized model does not seem able to license explanations, but only suggest them.[2]

I reject this sceptical worry: sometimes an essentially idealized model can license causal explanations of aspects of phenomena. To argue for this conclusion, I must first clarify how I am using some contentious terms like 'model' and 'idealization'. That is the aim of Section 2. I focus on a special kind of target-directed modelling, where the model is physical or concrete. The detailed case study that I develop in Sections 3 and 4 considers models of landslide generated impulse waves. These are naturally occurring waves in bodies of water that result from landslides. Although these models involve a range of idealizations, there are ingenious ways to identify and avoid endorsing the falsehoods associated with these idealizations. This means that the indirect strategy of studying these models to learn about these waves can be vindicated. There is good experimental evidence that the model is reliably generating accurate claims about some of the causes of these waves. These accurate claims are sufficient for causal explanations of certain features of landslide generated impulse waves.

One upshot of this discussion is that we can appreciate why concrete models remain important to science and engineering. They afford the means to extend

[1] See especially (Godfrey-Smith [2006]; Weisberg [2007b]; Weisberg [2013]).

[2] See, for example, (Jebeile and Kennedy [2015]; Rice [2019]). Others draw the conclusion that explanation does not require truth, as in (Cartwright [1999]; Suárez and Cartwright [2008]).

our knowledge of phenomena, even when the character of the phenomena precludes direct investigation and experimentation. The indirect strategy of model-based science introduces some obstacles in the form of essential idealization, which in our case are tied primarily to the smaller spatial scale of the concrete models. But this smaller size also permits careful manipulation and experimentation on the models that is sufficient to generate assurances that many of the claims generated by the model are true of relevant aspects of the phenomenon of interest.[3]

2 Models, Essential Idealization, and Selective Endorsement

In a recent paper, Frigg and Nguyen ([2018], p. 206) consider what they call a 'material' model that involves 'a material object [being] used as a model that represents a certain target system'.[4] Weisberg ([2013], p. 7) has also singled out the same kind of model using the term 'concrete model': 'Roughly speaking, concrete models are physical objects whose physical properties can potentially stand in representational relationships with real-world phenomena'. Frigg, Nguyen, and Weisberg all agree that these models may stand in representational relationships to target phenomena, but they propose somewhat different accounts of what these representational relationships amount to. Frigg and Nguyen develop what could be called an exemplification/translation account of representation. First, a model comes to exemplify some of its features. Second, some of the exemplified features of the model may be attributed to a target system, typically after they have been transformed using a translation key. When these two stages are combined, we have a model that represents a target in a certain way. These ways can be associated with propositions of the form 'exemplified model feature *X* is found in the target as feature *Y*', where *Y* is what results when the translation key is applied to *X*. Weisberg develops a somewhat different proposal for model representation that he calls a feature-matching account. First, scientists identify a set of features that are potentially shared by a model and its target. Second, the features are assigned weights in terms of a weighting function that reflects their relative importance for the modelling task at hand. Once these two elements are in place, there is some measure of the similarity between the model and the target. As with Frigg and Nguyen, this measure can be decomposed into propositions of a certain form: 'model feature *X* is present (or absent) in the target'.

Despite their disagreement on the nature of the representational relationship, there is a clear consensus on two points. To start, there are concrete

[3] My discussion is very much indebted to Sterrett's pioneering work on concrete models; see especially (Sterrett [2017a], [2017b]).

[4] See also (Frigg and Nguyen [2017]) for their survey of recent accounts of models.

models whose core is a material or physical object, and some of these models stand in representational relations to specific target systems. This is the sort of situation that I am focused on here: the models that I discuss are physical objects, and they are targeted at the phenomenon of landslide generated impulse waves. The second point of agreement is that the physical object does not bear any intrinsic relation to its target system. For a model to stand in a representational relation to its intended target (assuming it has one), agents must add the elements that constitute this representational relation. I will say that these elements constitute the interpretation of the model: when the physical object is supplemented with an interpretation, the result is a model that represents its target in a certain way. In what follows, I will suppose that the way that a model represents its target to be can be captured by a collection of propositions. My discussion will tend to follow Weisberg's approach of supposing that the propositions take the form of 'model feature *X* is present in the target'. However, I also draw on the translation key that is emphasized by Frigg and Nguyen. This is because the model features that are connected to the target features are different from one another in a systematic way. For example, the lengths of objects found in the model have corresponding lengths in the target that will be many times larger. There is no reason to suppose that Weisberg could not extend his feature-matching account to handle these sorts of interpretations as all that would be required would be a 'key' that would indicate how salient features of the model are supposed to be reflected in a 'similar' target.[5]

Once a model stands in a representational relation to an intended target it makes sense to ask to what extent it misrepresents that target. I will say that an idealization of a model is a special kind of false statement about the target that is generated by the features of the model and the representational relation that it stands in to that target. For such a false statement to count as an idealization, the agents using the model must believe that the statement is false. If the agents lack this belief, then we just have a model misrepresenting its target in a certain way, but not an idealization of the model. Suppose, for example, that an interpreted model generates a feature-matching claim of the sort 'model feature *X* is present in the target', but that in fact the feature is absent in the target. If the agents believe that *X* is not present in the target, then the statement that model feature *X* is present in the target is an idealization of the model.[6]

[5] My discussion in this paper is meant to address a worry about cases where concrete models are directed at specific target phenomena. Nothing I say here is meant to endorse any broader account of how models represent or the nature of models more generally.

[6] Cf. (Thomson-Jones [2005]). Thomson-Jones does not impose this belief requirement, and so in this sense his notion of idealization is much broader. However, he also emphasizes the importance of the misrepresentation pertaining to the relevant features of the target, which potentially restricts what counts as an idealization.

Some of the idealizations of a model will be found in the explicit statements that are used to interpret the model, but others will be merely implicit. For both an exemplification/translation approach and a feature-matching approach, an interpretation of a model will not simply enumerate all the ways that the model represents the target to be. Instead, some basic elements of the interpretation will be made explicit, with the provision that the full interpretation follows from these elements, along with some perhaps unknown features of the model. This is especially clear with a small-scale concrete model whose target is a much larger concrete system. A spatial scale will indicate how lengths in the model are to be associated with lengths in the target, but the lengths of all the objects in the model may not be known in advance. The purported length of some object in the target will then be discovered once a modeller combines this spatial scale with a measurement of a part of the model. Investigating a model that stands in a representational relation to a target can thus help to flesh out how the model is representing the target to be. In the course of these investigations, a modeller may come to realize a new respect in which the model is idealized.

In this article, I stipulate a special sense in which a model is essentially idealized. In this sense, a model is essentially idealized when any model of that type with that purpose has an idealization of some kind. That is, the modellers aim to model some target system in some way for some purpose, and there is no known way to do that without using a model with that kind of idealization. By contrast, a model is not essentially idealized when each idealization of the model can be avoided in a way that preserves the intended modelling purpose. When a model is essentially idealized, it is not clear how to use the claims that the model makes about the target to explain features of the target. For we believe that some of these claims are false, and we suppose that false claims are not apt to explain. My proposed solution to this problem is to say that scientists selectively endorse the claims that the model makes about the target. That is, they assemble evidence that some of the claims that the model makes about the target are true even though other claims that the model makes about the target are false. If this evidence is adequate, then modellers can explain some features of their target phenomenon using essentially idealized models. The model, even though it is essentially idealized, will represent the causes operating in the target to be a certain way, and in this respect the model will be an accurate representation.

3 Giant Waves in Lituya Bay, Alaska

While water waves or ‘tsunamis’ generated by earthquakes are well known, there is a rarer category of water waves caused by solid materials rapidly entering some body of water via a landslide. The most discussed and spectacular instance of such landslide generated impulse waves occurred on 9 July, 1958 in Lituya Bay, Alaska (see Figure 1).[7] Eyewitnesses reported a massive wave that originated at the head of the bay, rapidly running through the length of the bay and out into the ocean. The wave was so powerful that it scoured the sides of the bay, removing trees and soil. Subsequent investigation showed that this wave run-up reached the extraordinary height of 524 m in one localized region of the bay in Gilbert Inlet. This remains the largest wave run-up ever recorded (Fritz *et al.* [2009]).

As its name suggests, the phenomenon of landslide generated impulse waves is identified partly in terms of the primary cause of the wave: an event only counts as an instance of this phenomenon when the wave is generated by a landslide. The primary source of the difficulty in understanding these waves is the complex interactions between the landslide material, the water and the ambient air. As the landslide material enters the water, its impact is affected by the buoyancy of the water. The extent of the impact is also controlled by the depth of the body of water. As energy is transferred from the landslide to the water, an air cavity is generated that results in a thorough mixing of water and air. This 'air bubble entrainment' changes the character of the resulting water waves, and greatly complicates any direct treatment of the wave evolution.[8] On the one hand, the power of the wave is clearly due to the massive amount of energy injected into the body of water by a massive landslide. On the other hand, the mixing of solid, liquid and gas in this process blocks any simple theoretical treatment.

One research strategy to deal with these complexities involves the construction of and experimentation on small-scale concrete models of such landslide generated impulse waves. My discussion here considers the investigations that were initially conducted at ETH Zurich by a group that includes W. Hager, H. Minor, A. Zweifel, H. Fritz, and V. Heller, with special emphasis on (Fritz *et al.* [2001]).[9] The abstract of this paper summarizes the conclusion of their investigations using a concrete model: 'The laboratory experiments confirm that the 1958 trimline of forest destruction on Lituya Bay shores was carved by a giant rockslide generated impulse wave. The measured wave run-up perfectly matches the trimline of forest destruction on the spur ridge at Gilbert Inlet' (Fritz *et al.* [2001], p. 3). I interpret the second sentence to provide part of the evidence for the first sentence: a conclusion is drawn about the instance of the phenomenon based, in part, on the capacity for the concrete model to

[7] For a more recent instance, see (Schiermeier [2017]).

[8] Some of the challenges in handling such cases are surveyed in (Chanson [2009]).

[9] See also (Fritz *et al.* [2004]; Heller *et al.* [2008]; Fritz *et al.* [2009]; Heller and Hager [2010]; Heller [2011]). Space considerations preclude any discussion of the more recent investigation of three-dimensional concrete models in (McFall and Fritz [2016]; McFall *et al.* [2018]). It is important to emphasize that additional tools, such as computer simulations, are an important part of this broader research program.

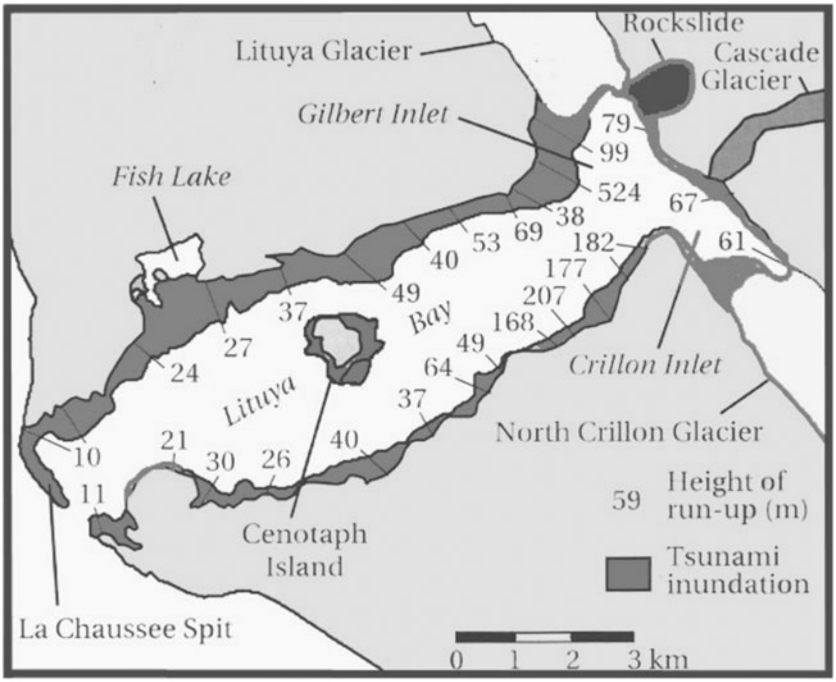

Figure 1. Lituya Bay with wave heights in metres.

reproduce a feature of that phenomenon. But obviously the 'match' here between the model and the target holds only when the model is appropriately interpreted. Some background is needed to appreciate how this interpretation works.

Two theories are deployed to motivate the construction and interpretation of the model: fluid mechanics and dimensional analysis. The theory of dimensional analysis plays a significant role in justifying the conclusions that these practitioners extract from their modelling activity. To appreciate this role, one should be aware of the central result of the theory, Buckingham's theorem (or the pi-theorem). Consider any equation $y = f(x_1, \ldots, x_n)$ that relates variables $x_1, \ldots, x_n$ to y via some function f. For this equation to represent a relationship between some quantities that obtains in some phenomenon, each variable must be related to an intended quantity via a system of units. For example, to represent the period of a pendulum, the variable T may be assigned the units of time of seconds. Once the equation is interpreted, it may be correct, and if it is correct, it may or may not be 'dimensionally homogeneous'. A dimensionally homogeneous equation is an equation that remains correct when variables $x_1, \ldots, x_n, y$ in a given system of units are transformed into a different system of units, resulting in variables $x_1', \ldots, x_n', y'$. To continue our example, consider the equation for the period of a simple pendulum, $T = 2\pi\sqrt{L/g}$. Suppose the variable for period T is assigned the units of seconds, the variable for length L is assigned the units feet, and the variable for gravitational acceleration g is assigned the units feet per second per second. If the equation is dimensionally homogeneous, $T = 2\pi\sqrt{L/g}$ will continue to hold even if we shift to a different system of units, such as seconds and metres. Some equations clearly are not dimensionally homogeneous. For example, $T = 1.11\sqrt{L}$ follows from our first equation if $g = 32.2$ feet per second per second. But this equation will fail if we transform our units of length from feet to metres.[10]

If equation $y = f(x_1, \ldots, x_n)$ is dimensionally homogeneous, then there is a so-called reduced equation that pertains to the very same phenomenon, but that relates dimensionless products of variables instead of dimensioned variables. To comprehend this contrast, it is essential to pin down what it means for a variable to have a dimension. When a variable in an equation is related to a physical quantity, it must be related via a system of units such as seconds and feet in the pendulum equation. More specifically, period T is assigned a unit of time, seconds, while length L is assigned a unit of distance, feet. When we transform our equation from one system of units to another system of units, the kind of unit that is appropriate for each variable remains the same. For example, length L is first assigned to feet, and then assigned to metres. The 'dimension' of a variable is a means of indicating how the numerical value of a variable in one system of units should be changed when shifting to another system of units.[11] If x is a variable, then $[x]$ is the dimension of that variable. A system of units comes with a set of fundamental units which permit the determination of the units of all variables. In our case, the fundamental units are feet and seconds, which together determine the units of acceleration as feet per second per second. This induces an algebraic structure on dimensions themselves which will be preserved across various systems of units: dimension $[L]$ of L and $[T]$ of T yield dimension $[L/T^2]$ for g, an acceleration. In addition, we can use the dimensioned variables to arrive at the dimensions of products of variables, such as $T/\sqrt{L/g}$. The dimension of product $T/\sqrt{L/g}$ is $[T]/\sqrt{[L]/[L/T^2]} = [1]$. Whenever such a product is assigned the dimension [1] through this sort of algebraic calculation, we say the product is dimensionless.

If equation $y = f(x_1, \ldots, x_n)$ is dimensionally homogeneous, then there is a reduced equation, $\Pi = F(\Pi_1, \ldots, \Pi_p)$, where $\Pi, \Pi_1, \ldots, \Pi_p$ are

10 This illustration and much of my exposition follows (Langhaar [1951]). This textbook draws extensively on Buckingham's ([1914]) classic paper, but also embeds the discussion in a more purely algebraic setting. I follow Langhaar in treating equations as relations between variables which are in turn tied to physical quantities via a system of units.

11 Some practitioners, such as Langhaar ([1951], Section 3), suppose that a dimension is nothing but this. For a more realistic attitude towards dimensions, see (Skow [2017]). I do not think this issue needs to be addressed to handle the case I develop in this paper.

dimensionless products of the original variables, $y, x_1, \ldots, x_n$. In addition, $p = n - r$, where r is the number of fundamental units in the system of units.[12] When $n = r$, the reduced equation will take the form of $\Pi = k$, where k is some constant. This is what happens with the simple pendulum case: a reduced equation is obtained from $T = 2\pi\sqrt{L/g}$ by dividing both sides by $\sqrt{L/g}$, yielding $T/\sqrt{L/g} = 2\pi$. As product $T/\sqrt{L/g}$ is dimensionless, the reduced equation holds for any system of units. This sort of simplicity is not necessary for dimensional analysis to be useful. In our landslide case, Fritz *et al.* begin by supposing that there is some dimensionally homogeneous equation that applies to landslide generated impulse waves. The character of this equation is not known, but the theory of fluid mechanics and past experience provide good evidence for the variables that it will relate. Assuming that the variables can be identified, Fritz *et al.* can use Buckingham's theorem to arrive at a list of p dimensionless products. To see how this works, we start with a system of units with three fundamental units whose dimensions are length [L], mass [M], and time [T]. Starting with 11 variables, Buckingham's theorem ensures that $11 - 3 = 8$ dimensionless products of these variables are sufficient for their target phenomenon (Heller [2011], pp. 298–9). As the reduced equation is adequate for this target phenomenon, it follows that agreement between the model and target on these 8 dimensionless products is sufficient for the model and target to agree in all other relevant respects. In particular, the causes found to be operating in the model will also be found to be operating in the target.

Some of these dimensionless products are quite intuitive. For example, for any distance x, there is a dimensionless 'relative distance' that results from dividing x by the depth of the bay h. Fritz *et al.* elected to build their concrete model at a spatial scale of 1: 675 ($\lambda = 675$). The measured depth of the bay is $h = 122$ m while the measured width of the bay is 1342 m. The model depth was thus set at $122/675 = 0.18$ m, and the model width was set to $1342/675 = 1.99$ m. The choice of h is of course not mandated *a priori*: any length variable with dimension [L] could be used to generate a family of relative distances. In this case, modellers used h to arrive at relative distances, but they could have used the width instead. The basis for choosing h is that modellers believe that the depth, h, is the most significant spatial variable. It figures into a variety of causal processes that determine the characteristics of the impulse wave. This importance is made clear by some of the other dimensionless products that are identified by a dimensional analysis of the phenomenon.

[12] For some clarification of r, see (Langhaar [1951], Chapter 3).

The most important dimensionless product for this phenomenon is called the Froude number, F:

$$F = \frac{V_s}{(gh)^{1/2}}. \tag{1}$$

This is the ratio between landslide velocity V_s and the square root of the product of the gravitational acceleration, g, and the depth, h. The numerator has dimension $[L]/[T]$, while the denominator has dimension $([L]/[T]^2 \cdot [L])^{1/2}$, which simplifies to $[L]/[T]$. The Froude number is thus a dimensionless product that can be used to characterize some important features of a landslide generated impulse wave. It can be conceived as a ratio between two of the forces that are at work in the phenomenon (Heller [2011], p. 295). The numerator, V_s, reflects the strength of the momentum carried by the landslide as it enters the water, while the denominator, $(gh)^{1/2}$, indicates the effects of the gravitational forces that further accelerate the landslide as it falls through the water to the bottom of the bay. Estimated slide velocity V_s for the Lituya Bay landslide is 110 m/s, while gravitational acceleration is $9.8\,\mathrm{m/s^2}$. This sets the Froude number for the landslide at $\frac{110\,\mathrm{m/s}}{(9.8\,\mathrm{m/s^2} \cdot 122\,\mathrm{m})^{1/2}} = 3.18$. The trials conducted with the concrete model aimed to match this Froude number by carefully setting the model slide velocity. To obtain a Froude number of 3.18 with a 1: 675 scale model, the model slide velocity must satisfy:

$$\frac{z\,\mathrm{m/s}}{(9.8\,\mathrm{m/s^2} \cdot \frac{122}{675}\,\mathrm{m})^{1/2}} = \frac{110\,\mathrm{m/s}}{(9.8\,\mathrm{m/s^2} \cdot 122\,\mathrm{m})^{1/2}}$$

$$z\,\mathrm{m/s} = 110\,\mathrm{m/s} \cdot \frac{1}{\sqrt{675}}$$

$$z\,\mathrm{m/s} = 4.22\,\mathrm{m/s}.$$

The model slide velocity was thus set to 4.22 m/s. When a model and target agree on their Froude numbers, the two are said to be Froude similar.

When modellers say that a model and target are Froude similar, there is a corresponding interpretation of the model that allows one to take claims about the model and translate them into corresponding claims about the target. This makes it possible to evaluate the claim that the model run-up 'perfectly matches' what happened in the target. For a spatial variable like the run-up, this means that the model variable is $1/\lambda = 1/675$ times the target variable. That is, the concrete model showed a run-up of 526/675 m or 0.77 m. But times are interpreted differently. We saw how the requirement of Froude similarity mandated that model slide velocity V_s be set at $1/\sqrt{675}$ times target slide velocity V_s. More generally, this is how all model velocities are interpreted as generating claims about target velocities. For this interpretation of model lengths and model velocities to be coherent, model times must be interpreted as $1/\sqrt{675}$ target times. This ensures that the claims about velocities in the target that are based on the model will agree with the claims about distances and times, as velocities just are ratios of distances and times. For example, if an element of the model moves 1 m in 1 s, for a model velocity of 1 m/s, this generates the claim that the corresponding element of the target moves 675 m in $\sqrt{675} = 26$ s, for a target velocity of 26 m/s. These aspects of the interpretation of the model are needed to determine, for example, how the velocity of a wave in the model is to be related to the velocity of a wave in the target.

Now that the model is fully interpreted, it makes sense to wonder to what extent the model is idealized and if the model is essentially idealized. The model's interpretation involves two significant idealizations. As this sort of idealization is present in any alternative way of modelling the bay for these purposes, I conclude that the model is essentially idealized. Both idealizations are associated with so-called scale effects. Scale effects arise when a small-scale concrete model must be dissimilar to its intended target in ways that are known to be causally relevant to the phenomenon in question. In the Lituya Bay case, we have focused on the Froude number and the aim of having a match in the Froude numbers of the model and target. However, there are two other dimensionless products that are known to be important for water waves. These are the Reynolds number and the Weber number. The Reynolds number involves the kinematic viscosity of water, ν, while the Weber number involves the density of water, ρ, and the surface tension of water, σ:

$$R = \frac{g^{1/2} h^{3/2}}{\nu}, \tag{2}$$

$$W = \frac{\rho g h^2}{\sigma}. \tag{3}$$

Just as the Froude number can be conceived as reflecting a ratio between two forces, the Reynolds and Weber numbers relate other forces at work in water waves (Heller [2011], p. 295). The Reynolds number reflects the relative importance of the inertial forces and the viscosity of the water: at low Reynolds numbers, fluid flows can be divided into regular layers, while at high Reynolds numbers turbulent flows arise where fluid layers mix and eddies are generated. The Weber number reflects the comparative strengths of inertial forces and the surface tension of water: the larger the Weber number, the less a wave's behaviour is affected by this surface tension.

The scale effect tied up with the Froude, Reynolds, and Weber numbers is that a small scale concrete model that is Froude similar to its target must be quite dissimilar with respect to both the Reynolds and Weber numbers. To see why, note that the ratio between the actual depth, h, and the model depth is given by the spatial scale of the model, in our case 675. The remaining variables g, ν, ρ, and σ are fixed by the character of water (for ν, ρ, and σ) or the laboratory environment (for g). It follows that if the target depth is 675 times the model depth, then the target Reynolds number is $(675)^{3/2} \approx 17500$ times greater than the model Reynolds number. Similarly, the target Weber number is $(675)^2 \approx 455{,}000$ times greater than the model Weber number. This is a significant mismatch in a causally relevant feature of water wave dynamics, and there is every reason to suspect that the dissimilarity in Reynolds and Weber numbers is distorting the claims about the target that are made on the basis of what is found in the model. The claims that the model makes about the target phenomenon are generated using the features of the model and the interpretation of those features based on Froude similarity. But Froude similarity precludes a match in Reynolds and Weber numbers. This is an unavoidable consequence of the materials used and the choice to interpret the model in this way. I conclude that the model is essentially idealized: there is no known way to develop a concrete model that aims to model the causes of this wave without having that model generate some known falsehoods about its intended target.

4 Mitigating Scale Effects

One way to use an essentially idealized concrete model to come to know a causal explanation is to note that scale effects are themselves a phenomenon that can be studied and understood. Through this sort of reflective investigation, modellers can assemble evidence that they have minimized the problematic aspects of scale effects, even if they cannot be entirely eliminated. Here I will summarize Heller's ([2011]) discussion of this 'avoidance' strategy for this case in his insightful 'Scale Effects in Physical Hydraulic Engineering Models'.[13] The avoidance strategy involves a seemingly paradoxical result: while the variables that define the Reynolds number are causally relevant to the behaviour of water waves, in this case the differences between the model and target 'makes no difference' to the aspect of the phenomenon in question. In Section 5 I will analyse this situation using Woodward's notion of conditional causal irrelevance. The variables that define the Reynolds number, R, are conditionally causally irrelevant to certain features of the wave in specifiable circumstances, even though those variables are causally relevant to those very features.

Heller's avoidance strategy involves building a concrete model so that the scale effects will be unimportant for the target phenomenon. We can consider

[13] Heller also discusses compensation and correction strategies, but I will not pursue these here for reasons of space; see also (Sterrett [2017a]) for a richer investigation of a wider range of cases.

a range of size scales, λ. The larger λ is, the smaller the concrete model will be when compared to the target. This presumably makes the model easier to build and investigate. But if the model is too small, then scale effects will spoil the claims about the target that are extracted from the model. In Fritz *et al.*'s case, the modellers chose $\lambda = 675$. The worry is that they may have made their model too small. Heller notes that modellers have developed 'rules of thumb' for how large λ can be, consistent with the reasonably accurate modelling of some type of target phenomenon. However, Heller ([2011], p. 299) also helpfully points out that these rules of thumb must be tied to the specific aims of the modeller: 'These guidelines may be misleading if, for example, just a limiting scale factor λ or water depth h on its own is applied without considering to which prototype features [target features] they were defined'. Heller illustrates this point using the landslide generated impulse waves. Consider two small scale concrete models of such waves where the depths are 0.4 m and 0.2 m. We can conceive of the smaller concrete model as a model of the larger concrete model. Then $\lambda = 2$, so we would expect only negligible scale effects. But when we compare the character of the two waves immediately after the landslide enters the water, significant differences emerge. There are significant scale effects that spoil the model with respect to the extent of air entrainment at the time the landslide enters the water. Remarkably, though, this failure is consistent with scale effects not being significant for the purposes of modelling other features of the wave. Here I will focus on the maximum amplitude achieved by the primary wave A_M, the primary wave velocity, and the run-up height of this primary wave, Y.

The key assumption needed to link acceptable spatial scales to features of interest is that these features arise in regular and predictable ways for extended ranges of dimensionless products. To show that this is the case for A_M and Y, Heller and his collaborators experimented with a 'scale series' of concrete models of landslide generated impulse waves in order to determine how various features of interest arise once certain thresholds are reached. Going from a depth of 0.4 m to 0.2 m and then 0.1 m can induce significant scale effects. If we again think of the smaller models as models of the larger models, then λ ranges from 2 to 4. But these scale effects fail to induce a mismatch for A_M and Y. Heller ([2011], p. 299) summarizes these experimental researches by claiming that 'scale effects are negligibly small (< 2%) relative to the maximum wave amplitude a_M, if $R_I = g^{1/2}h^{3/2}/\nu \geq 300{,}000$ and $W_I = \rho g h^2/\sigma \geq 5000$ resulting in the rule of thumb $h \geq 0.200$ m for typical laboratory conditions'. That is, when the Reynolds number and the Weber number in the model exceed some thresholds, the predictions generated by the model (under the interpretation tied to Froude similarity) for primary wave amplitude A_M will closely match what will be found in the target phenomenon. Crucially, these thresholds are not a simple function of λ. As we have seen, the Reynolds number and

Table 1. Reynolds and Weber numbers. Adapted from (Heller *et al.* [2008], p. 698)

S4	*h*	*R*	*W*
S4/1	0.4 m	836,495	21,382
S4/2	0.2 m	313,103	5345
S4/3	0.1 m	109,414	1336

the Weber number take account of other features, such as the viscosity and surface tension of water. Table 1 summarizes the Reynolds and Weber numbers for Heller *et al.*'s scale series S4/1, S4/2, and S4/3: When h is decreased from 0.4 m to 0.2 m, the Reynolds number also decreases, from around 800,000, to around 300,000. The additional decrease from 0.2 m to 0.1 m takes the Reynolds number from around 300,000 to around 100,000. If the primary wave behaviour is essentially unchanged once we have passed a Reynolds number of 300,000, then there is every reason to trust a concrete model once its depth has exceeded 0.2 m. A similar pattern emerges for the Weber numbers, where 0.1 m, 0.2 m, and 0.4 m correspond to Weber numbers of around 1300, 5300, and 21,300, respectively. Again, once the threshold of 5000 is reached, we may have a reliable way to generate reasonably accurate claims about the primary wave using observations of the concrete model.

The data arrived at through experiments with these three concrete models is summarized in Figure 2 (Heller *et al.* [2008], p. 698). Heller *et al.* are able to establish a very close match between the primary waves in S4/1 ($h = 0.4$ m) and S4/2 ($h = 0.2$ m), indicating that scale effects are relatively insignificant for A_M. By contrast, the gaps between S4/1 and S4/3 ($h = 0.1$ m) remain relatively large, indicating the presence of important scale effects. The upshot of this analysis is that Froude similar models of landslide generated impulse waves can be used when the depth is $h = 0.2$ m or greater. But these models should only be targeted at A_M, the velocity of the primary wave, and consequent primary wave run up Y. Nothing in this analysis licenses the use of these models to understand the immediate effects of the landslide impact. As we have seen, even when $\lambda = 2$ there are significant scale effects that undermine any extrapolation from the model's behaviour right after the landslide enters the water.[14]

Experimental investigation of the concrete models in a scale series can show that, even though the models are idealized in important ways, they are still reliable guides to certain features of the targets, even when those features are

[14] It is important to note that the data summarized in Fig. 2 are not the raw data, but the data after it has been analysed; see (Heller *et al.* [2008], p. 698). One could question the details of this analysis, but I do not have the space to address those concerns here.

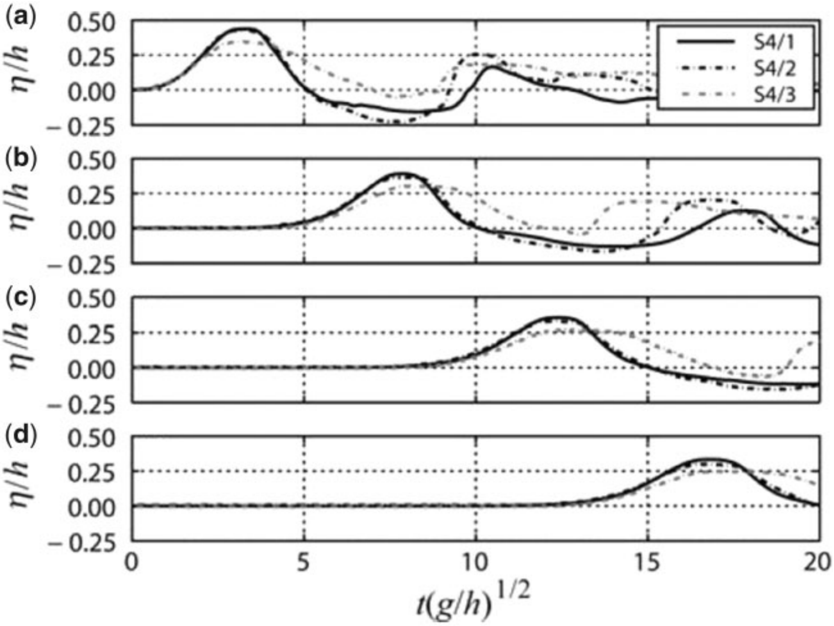

Figure 2. Scale series data (Heller *et al.* [2008], p. 698).

caused by elements that are idealized according to the interpreted model. In our case, the Reynolds and Weber numbers of the model fail to match the target, and this failure is tied to a causally relevant difference between the physical processes responsible for the waves in the model and target. Nevertheless, the mismatch can be estimated and argued to be unimportant for predicting the amplitude and velocity of the primary wave. This is because we can take care to exceed certain thresholds that are controlled by the depth of the water. In this analysis, the reliable threshold of 0.2 m is actually slightly larger than Fritz *et al.*'s choice of 0.18 m. This suggests that future work on this phenomenon should always use depths that exceed 0.2 m. Other phenomena might mandate even more demanding restrictions, corresponding to λ less than 100 or even less than 5. It all depends on the character of the relevant dimensionless products and how they relate to the features of the phenomenon that are of interest.

5 Accuracy for the Purposes of Causal Explanation

In an influential book and a number of papers Woodward has advanced an interventionist account of causation.[15] This framework applies very well to the case of concrete models and can be used to clarify how essentially idealized

[15] See especially (Woodward [2003]; [2013]; [2016]).

concrete models can extend our knowledge of phenomena. For Woodward, causation is a relation between variables, and causal facts obtain only relative to a set of variables. The metaphysical status of these variables is left open, but the basic idea is readily adaptable to our landslide case. We have treated a variable with a unit as a means of referring to a physical quantity. A variable can thus be treated as a family of incompatible values, where each value stands for a specific physical quantity. For example, depth h (in metres) is a variable whose values range from 0.4 m to 122 m. On an interventionist analysis, to say that the depth of Lituya Bay is a cause of the wave run-up of 524 m is to make a certain type of counterfactual claim. First, we select a set of variables, and set their values using the features of Lituya Bay on the day the wave occurred. The causal claim is true just in case there is some intervention on the depth of Lituya Bay that would result in a change in wave run-up. An intervention here is a special sort of manipulation of the depth of the bay which succeeds in changing the depth while holding all other potential causes of the wave run-up at their actual values. There is no requirement that this intervention be practically feasible. Woodward repeatedly emphasizes our ability to assemble indirect evidence concerning the outcomes of these interventions, such as through randomized controlled studies.

Landslide generated impulse waves are of course a case where interventions are not practically feasible and it may not be immediately clear how we can assemble evidence for interventions in such cases. Woodward ([2003], p. 125) has not said much about concrete models, but one promising strategy is mentioned in passing in an early criticism of agency views of causation: 'when we ask what it is for a model or simulation that contains manipulable causes to "resemble" phenomena involving unmanipulable causes, the relevant notion of resemblance seems to require that the same *causal* processes are operative in both'. This is the view that I am arguing for in this paper: even when a concrete model is essentially idealized, scientists can assemble evidence that the causal generalizations that apply to the model are also applicable to the target.

I claim that the evidence for these conclusions is assembled in two stages, in line with the indirect strategy of model-based science. First, practitioners establish causal claims about the concrete models themselves. Then they extend some of these causal claims to their target phenomenon, and draw conclusions about the results of impractical interventions on these targets. In both stages, the variables that mark the beginning of dimensional analysis are the potential causes of the features of interest. But as we have seen, there is an intermediate stage where practitioners deploy another family of dimensionless products of these variables. As I have reconstructed this process, the dimensionless products are primarily tools for translating the results found in the model to claims about the target. In addition, the dimensionless products can be used to assess the reliability of these extrapolations.

In our wave case, the effects of interest are the maximum amplitude achieved by primary wave A_M, as well as wave run-up Y. The potential causes that are included in the concrete model are the density, height, length, and impact velocity of the slide, as well as the depth of the bay and the density of the water. Other potential causes that are considered are the kinematic viscosity, ν, and surface tension of the water, σ. The focus on the Froude number involves the assumption that the variables that determine it are very significant. So the hope is that the slide velocity, depth and gravitational acceleration are each causes of the maximum amplitude and wave run-up. Experimentation with the concrete model can show how changes in each of the slide velocity and depth result in changes in the maximum amplitude and wave run-up. While gravitational acceleration can not be directly manipulated in this context, scientists have background theoretical motivations for concluding that this variable is also a genuine cause of the features of interest.

The most interesting step concerns how and when the causal patterns found to hold in the concrete model should be extended to the target phenomenon. Ignoring scale effects, the extrapolation is unproblematic: when the model is set so that its Froude number and other dimensionless products agree with the actual conditions in Lituya Bay, there is good reason to conclude that the actual causes in the model will correspond to the actual causes in Lituya Bay. For example, a cause of the observed model run-up of 0.77 m is the model depth of 0.18 m. That is, were the depth to have been different as a result of an intervention, then the model run-up would have been different. This causal claim about the model is directly translated to the target: a cause of the observed target run-up of 524 m is the Lituya Bay depth of 122 m. If the bay's depth had been changed through an intervention, then the wave run-up would have been different.

How can this sort of causal extrapolation be justified once we acknowledge the prevalence of scale effects? I will defend a strategy that I call selective endorsement. Using an interpretation of the concrete model, a wide range of claims are generated about the target system, including claims about the causes of various features of the target. Given her awareness of the mismatch between the model and target associated with various idealizations, a practitioner must indicate which generated claims about the target she wishes to endorse. In our wave case, the endorsement is clarified through an experimental investigation of scale effects and the conclusion that certain mismatches make no difference to the aspect of the target of interest. The practitioner is thus rational to endorse a causal explanation of the aspects of interest, while refraining from endorsing the claims about the target that are compromised by the idealization.

Recall that the experimental investigations of the scale series of concrete models showed that the concrete model involved scale effects that undermined its reliability for the period immediately after the landslide entered the body of water. However, these same investigations indicated that the scale effects are not important for the maximum wave amplitude or the consequent wave run-up. This is because the Reynolds number, R, and the Weber number, W, in the concrete model exceeded certain thresholds, even though those numbers were much smaller in the model than they were in the target. How do these experimental investigations license a selective endorsement of the causal claims generated by the model?

It is here that we can deploy Woodward's notion of conditional causal irrelevance. Woodward develops this notion as part of a debate about the relationship between a set of macro-variables, X_i, and a set of micro-variables, Y_k. In certain situations, a causal explanation of effect variable E in terms of X_i is just as good as a causal explanation in terms of Y_k. This is because all interventionist information concerning the salient values of E is condensed in the macro-variables. In this sense, once we deploy the macro-variables, there is no additional causal information to be obtained from also considering the micro-variables: 'A set of variables Y_k is irrelevant to variable E *conditional* on additional variables X_i if the X_i are unconditionally relevant to E, the Y_k are unconditionally relevant to E, *and* conditional on the values of X_i, changes in the value of Y_k produced by interventions and consistent with these values for X_i are (unconditionally) irrelevant to E' (Woodward [forthcoming], pp. 17–18). I will extend this notion to cases where there is a tighter relationship between the variables than what we have in a micro versus macro case.[16] In this extended sense, a set of variables, Y_k, is irrelevant to variable E conditional on additional variables X_i each exceeding some specified threshold. For our example, the Y_k variables are the dimensioned variables ν and σ. We have reason to believe that their causal relationships are not correctly tracked when we translate the features of the model to the target. The additional variables X_i are the dimensionless products R and W that are defined partly in terms of ν and σ. Finally, the effect variables here are A_M and Y. When R and W exceed certain thresholds, the mismatch between model and target with respect to the causal consequences of ν and σ can be discounted as their actual values are conditionally causally irrelevant to A_M and Y. That is, if we keep R and W above the requisite thresholds, manipulations on ν and σ will fail to change A_M and Y.

To see how this works, consider a causal explanation of $A_M = 152$ m and $Y = 524$ m that occurred in Lituya Bay. The dimensioned variables V_s, h, and g that enter into the Froude number are deemed causes of both of these values. Should we also say that the other dimensioned variables that enter into R and W, namely, ν and σ, are also causes of these effect values? On the one hand,

[16] Cf. (Woodward [2015], Section 6).

they are unconditionally causally relevant as there are interventions on ν and σ that change the maximum wave amplitude and the wave run-up. On the other hand, they are conditionally causally irrelevant once we have specified a regime where $R \geq 300{,}000$, and $W \geq 5000$.

With these distinctions in place, it should be clear how a selective endorsement strategy can be implemented. Based on an experimental investigation of the concrete model, a scientist may conclude that some dimensioned variables are conditionally causally irrelevant to the aspects of the phenomenon of interest. Both in the model and in the target, the values of ν and σ make no difference to the maximum wave amplitude and the wave run-up once R and W exceed their thresholds. The dimensioned variables that are conditionally causally relevant to these outcomes are all the remaining variables, including the slide velocity and slide dimensions, as well as h and g. This motivates a scientist to refrain from endorsing a claim about the target generated by the model whenever it is implicated in an idealization. But the scientist can be assured that these divergences between model and target will not undermine some other model-based extrapolations. This is because the claims that they have refused to endorse have been shown to involve variables that are conditionally causally irrelevant. These claims generated by the model are not reliable indicators of the features of the target, but this unreliability does not undermine the model's indications about the causes of the wave amplitude and run-up. We can use a fully interpreted concrete model to explain and control features of this phenomenon. Scientists selectively endorse the claims that the model makes about the target, under this interpretation, using the experimentally generated evidence that the claims that they endorse are reliable, even if many other claims generated by the model are not reliable.

The selective endorsement strategy is different from what could be called the selective interpretation strategy. The latter idea is that a modeller should limit their interpretation of their model so that the claims that the interpreted model makes about the target avoid any known falsehoods. I would concede that some idealizations warrant selective interpretation, and have examined some cases of this sort elsewhere.[17] For example, when a variable in a mathematical model is set to infinity, it is natural to suppose that the modeller intends to decouple that variable from any physical interpretation. But I do not think the selective interpretation strategy can work when we consider essentially idealized concrete models and the modelling purpose of arriving at causal explanations. For our model to generate claims that are able to explain the primary wave run-up in Lituya Bay, we need the model to generate claims about the primary wave's amplitude and velocity over an extended period of time, leading up to the time of the run-up. In interpreting the model in terms of Froude

[17] See (Pincock [2014]). An important recent discussion of this strategy is (Nguyen [forthcoming]).

similarity, modellers advance a full interpretation of the model's features, including all lengths, times and velocities. But this requires that the model generates claims about the target that are believed to be false due to the distorted model values for the Reynolds number and Weber number. If we try to calibrate or limit our interpretation to avoid these claims, then we will arrive at a heavily censored interpretation that avoids any claims about the dynamic evolution of the primary wave. That is, we will have a purely predictive model that 'black-boxes' the dynamics of the wave's development. This sort of interpretation does not generate claims about the target that are apt to explain why the primary wave run-up was 524 m. I conclude, then, that in some cases a selective endorsement approach is needed to make sense of how model-based science can afford knowledge of causal explanations.

6 Potential Alternatives

6.1 Robustness analysis

One widespread strategy for dealing with essentially idealized models is known as robustness analysis. A familiar example from recent philosophy of science is a robustness analysis of the Lotka–Volterra model of predator–prey interactions (Weisberg and Reisman [2008]; Weisberg [2013]). We can consider a case where the model is directed at a specific type of predator–prey system, such as the fish in the Adriatic Sea over some period of time. The numbers of various predator and prey are causally affected by any number of factors, and most of these factors are not present in the model. Nevertheless, the model is used to license an explanation of some changes in the numbers of predator and prey. Weisberg and Reisman ([2008], p. 114), in particular, have emphasized the explanatory power of the so-called Volterra principle: 'Ceteris paribus, if a two-species, predator–prey system is negatively coupled, then a general biocide will increase the abundance of the prey and decrease the abundance of the predators'. This principle is one of the claims about the target that is generated by the features of the model and its representational relationship to a target. How, though, can this license the use of this explanatory principle, given that the model is highly idealized?

One proposed way to answer this question is to deploy robustness analysis. Schematically, a robustness analysis starts with a model that gives rise to a principle about a target. The analysis consists in varying certain features of the model in a systematic way, and then checking to see if the same principle is consistently generated. Perhaps the simplest sort of robustness analysis varies the parameters of the model. In the Lotka–Volterra model, the parameters include r, the growth rate of the prey population, and a, which determines what proportion of prey are consumed by the predators. A parameter robustness analysis would vary the parameters of the model and see if the elements required to derive the Volterra principle remain in place. If so, then the model robustly generates that principle. A more demanding sort of robustness analysis is 'structural' in the sense that it changes a mathematical element of the original model with a new mathematical relationship. For example, Weisberg and Reisman consider the result of replacing the term rV with a more complex term, $r(1 - V/K)V$, that under the intended interpretation, varies the causal content of the model. Again, the original model is robust under this sort of structural change when the elements needed to generate the Volterra principle remain in place (Weisberg and Reisman [2008], Section 4). An even more demanding form of 'representational' robustness analysis would alter the character of the mathematical model even more dramatically. Weisberg and Reisman ([2008], Section 5) develop an agent-based model of predator–prey interactions that uses a completely different mathematical framework than the original model, and check to see if an analogue of the Volterra principle results from this new type of model.

Weisberg and Reisman ([2008], p. 129) conclude their survey of robustness analysis with this summary:

> [. . .] robustness analysis has shown that the principle is highly general and will hold under a wide variety of conditions. It is not dependent on idealizing assumptions made in various models of predation. While any given model contains idealizing assumptions, analysis across models has allowed us to control for them and factor them out.

A pressing question for our discussion of essentially idealized concrete models is thus the relationship between the sort of dimensional analysis via scale series carried out by Heller *et al.* and the robustness analysis advanced by Weisberg and Reisman. Is the dimensional analysis just an instance of robustness analysis, or is there some significant difference between them?[18]

One difference is the scope of the two techniques: robustness analysis is applicable to any essentially idealized model, while the scale series analysis works only for concrete models, and only when the series is structured by the background theory of dimensional analysis. The scale series considered smaller and smaller concrete models, with depths ranging from 0.4 m to 0.1 m. Significant scale effects arose when the depth went below 0.2 m indicating that models with depths below 0.2 m could not reliably indicate the causes of the primary wave run-up in the target phenomenon. A robustness analysis of these results could thus conclude that the causal relationships found in the larger model are not robust across all these changes in the depth. However, the dimensional analysis allows additional conclusions to be drawn in terms of the well-motivated dimensionless products F, R, and W. As we saw, there are

[18] I am grateful to an anonymous referee for raising this question.

empirically established thresholds that can be given in terms of *R* and *W* that allow scientists to understand why a given causal regularity will not hold below a certain scale. With the background theory of dimensional analysis in place, then, additional, positive conclusions can be drawn when compared to a generic form of robustness analysis. These positive conclusions enable the practitioner to be more confident in extrapolating their causal claims from the small-scale concrete models to the full-scale targets than they would otherwise have been. On this reconstruction, then, the scale series analysis can be treated as a special case of robustness analysis where additional background information is provided by the theory of dimensional analysis.

One benefit of this reconstruction is that we can diagnose and address some of the sceptical worries that have been raised about robustness analysis. For example, Odenbaugh and Alexandrova ([2011]) maintain that robustness analysis is merely a means to discover certain principles, rather than to confirm them in the way that Weisberg and Reisman claim. They reach this pessimistic conclusion by pointing out that robustness analysis does not scrutinize all of the idealizing assumptions of a model. This shows the possibility of residual doubts about a principle that has passed various robustness tests. Without endorsing this general worry, it should be clear how the background theory of dimensional analysis can be used to respond to Odenbaugh and Alexandrova's scepticism for our case. This background theory tells us which dimensionless products are sufficient to characterize the phenomenon in question. There is thus genuine confirmation of a causal claim about the target through the analysis of these concrete models.

6.2 Holism about models

Rice has recently argued that a wide variety of 'decompositional' approaches to models fail for what we are calling essentially idealized models. This is because it is not feasible to wall off the falsehoods associated with the idealizations from the other aspects of the model. Rice ([2019], p. 182) targets philosophers who deploy a 'model decomposition assumption' that supposes that 'The scientific model is decomposable such that the contributions of its accurate parts can be isolated from the contributions of its inaccurate (that is, idealized or abstracted) parts'. Rice ([2019], p. 204) rejects this assumption, and concludes that models must be approached in a more holistic way: 'Many (if not most) idealized models in science ought to be characterized as holistically distorted representations of their target system(s)'. This is supposed to have immediate implications for how essentially idealized models are used to explain features of target systems. For if the model cannot be decomposed, but is indeed a holistic unit, then the explanations that arise from the use of the model bear a somewhat tenuous relationship to the model itself. As Rice

([2019], p. 205) puts it, on his view, idealized models afford the use of various mathematical techniques, and thus 'provide epistemic access to explanations and understanding that would otherwise be inaccessible'.

Rice is not focused on concrete models, but it would seem that the account of model-based causal explanation developed here is an instance of precisely the sort of model decomposition that Rice supposes is rarely possible. In response to Rice I would distinguish two stages in the generation of a genuine model-based causal explanation. In the first stage, the concrete model is constructed and interpreted. At this point I would agree with Rice that there is no way to separate out which of the claims generated by the model about the target are accurate and which claims are inaccurate. So I would agree with Rice that at this first stage of investigation it is difficult to vindicate the model decomposition assumption. But, against Rice, I would emphasize the significance of a second stage of model-based science. During this stage, a scientist can study the model and draw on a wide range of theoretical and experimental knowledge for the express purpose of 'decomposing' the claims generated by the interpreted model. This is an important part of modelling, where modellers learn more about their models, and at times reconstruct the models themselves in light of what has been found. In the Lituya Bay case, it is only through the careful experimental study of the concrete model that modellers were able to study the scale effects, and sort out their consequences. Once this stage is complete, it is certainly possible to decompose the model and meet Rice's model decomposition assumption. In our case, various causes of A_M and Y in the Lituya Bay wave were identified through the model, and a causal explanation was obtained. The idealizations associated with the scale effects were found to be limited to aspects of the wave besides A_M and Y, and a partial correspondence between causes in the model and target was obtained.

To be fair to Rice, it must be emphasized that many of his interlocutors make it appear as if it is a routine matter to sort out the accurate and inaccurate claims generated by the model. For example, Weisberg ([2007a], p. 642) talks of 'the practice of constructing and studying theoretical models that include only the core causal factors which gave rise to the phenomenon'. This combines the two stages that I have just distinguished, and makes it seem as if the modellers can tell in advance what those core causal factors might be. A more charitable interpretation of Weisberg would be that he intends just the sort of two-stage modelling process that I have described. This is how he proceeds in his discussion of the 'calibration' of a concrete model of the San Francisco Bay: 'This required a feedback process: adjustment to the model, then analysis of the model, then further modifications of the model until the model reached a specified standard of fidelity' (Weisberg [2013], p. 95). An additional resource would be the sort of robustness analysis sketched in the last subsection. Another of Rice's ([2019], p. 185) targets, Strevens, talks of dividing the 'content' of a model into two parts, where 'The first part contains the difference-makers for the explanatory target [...while] The second part is all idealization; its overt claims are false but its role is to point to parts of the actual world that do not make a difference to the explanatory target'. Again, this makes it seem as if the initial interpretation of the model effects a clear demarcation between what is supposed to be accurate and what is inaccurate. But, as with Weisberg, we could charitably ascribe to Strevens the two-stage account of modelling that I have developed: it is only after the model has been investigated and perhaps adjusted that we can have any confidence that we have sorted out the accurate from the inaccurate. At the end of this process, we can engage in the selective endorsement of the model's claims in a principled manner, even if this is not possible at the beginning of the modelling process.

This version of anti-holism about idealized models insists that it is sometimes possible to arrive at well justified claims about a target based on an examination of a model. These claims include causal generalizations, and so we have the means to causally explain features of the target using claims generated by the model. However, Rice ([2019], p. 194) raises a further challenge to this approach, over and above the worries that we have already addressed: 'even if we assume the real-world system and the idealized model are decomposable in the ways required, the model's idealizations will often distort difference-making (that is, relevant) features of the model's target system(s)'. It is at this point that I believe that Weisberg, Strevens and many others are vulnerable. For these authors have yet to confront the sort of essentially idealized models that Rice emphasizes, and that I have developed here. In our case, it is clear that ν and σ are causally relevant to the landslide generated impulse wave in Lituya Bay, and yet the values of R and W are thousands of times larger in the target than in the model. So it is just not clear how to identify Weisberg's 'core causal factors' so that ν and σ are excluded, and it is also not clear how Strevens could suppose that ν and σ are not difference makers for this 'explanatory target'. On this point, we have the resources to offer a new response to Rice's worry: we can insist that ν and σ are conditionally causally irrelevant once R and W exceed the appropriate thresholds. So we can grant to Rice that the concrete model generates claims that distort these aspects of the target and that these aspects are (unconditionally) causally relevant to the target system, that is, the wave. But the experimental evidence assembled by Heller *et al.* allows us to conclude that these distorted features are conditionally causally irrelevant to the features of the target system that we are concerned with, namely, A_M and Y.[19]

[19] Just to be clear, nothing that I say here is incompatible with Weisberg's treatment of concrete models, such as the small-scale model of the San Francisco Bay (Weisberg [2013], especially

It is important to emphasize how piecemeal this sort of response to Rice really is. On the two-stage approach to modelling that I have described, modellers begin with an essentially idealized model and with no general assurance that any useful explanation will be forthcoming. In the case of a concrete model, additional experimental investigations are needed to sort out the accurate claims from the inaccurate claims, and also to determine to what extent the inaccurate claims lead to distortions of features of interest. It may seem somewhat fortuitous that Heller *et al.* are able to identify thresholds in terms of certain dimensionless products that assure them that the model-based explanation is a genuine one. Many cases of modelling will surely end in failure, with no viable explanation of features of the target. It is a matter of skill and even luck to be able to get such modelling enterprises off the ground. My point is mainly that modellers have unexpected resources in the case of concrete models, and that they can exploit these resources to address the worries raised by Rice and others.

7 Conclusion

In this paper I have used a case study of concrete models of landslide generated impulse waves to illustrate how essentially idealized concrete models may be used to come to know causal explanations of real world phenomena. The concrete models of these waves are essentially idealized due to ineliminable scale effects that arise when a small-scale concrete model is built using the same materials that are found in the target. There is a mismatch in some dimensionless products that are known to reflect causally relevant features of the process being studied. This idealization cannot be eliminated because there is no feasible alternative means of modelling these waves using such concrete models. However, practitioners were able to experiment with these models and provide evidence that the mismatch between model and target does not undermine the accurate representation of some of the causes of some of the features of these waves.

This success highlights (i) the ongoing scientific value of concrete models, (ii) their essentially idealized character, and (iii) how idealizations can be made to cohere with model-based causal explanations. It remains to be seen how widely this strategy can be extended, both in terms of other kinds of concrete models and also for abstract mathematical models where the sources of essential idealization will be different. In the landslide case, a generalization of Woodward's notion of conditional causal irrelevance proved crucial in vindicating a principled strategy of selectively endorsing the claims generated by the

pp. 41–2, 84–8, 153, 167–8). However, I do not think Weisberg has yet clarified the benefits of dimensional analysis in this sort of case.

idealized model. But there is no *a priori* assurance that selective endorsement will work in other cases. Additional investigations are needed to appreciate why practitioners work with idealized models and what their strengths and limitations might turn out to be.

Acknowledgements

An earlier version of this paper was presented in March 2018 at Models and Simulations 8, held at the University of South Carolina. I am grateful to the audience for their questions and suggestions, especially James Nguyen, Michael Poznic, Julia Sanchez-Dorato, Susan Sterrett, Collin Rice, and Michael Weisberg. I would also like to thank James Woodward for permission to quote from his forthcoming paper. Two anonymous referees for this journal helped a great deal during the final revisions of this paper.

Department of Philosophy
Ohio State University
Columbus, OH, USA
chrispincock@gmail.com

References

Bogen, J. and Woodward, J. [1988]: 'Saving the Phenomena', *The Philosophical Review*, **97**, pp. 303–52.

Buckingham, E. [1914]: 'On Physically Similar Systems: Illustrations of the Use of Dimensional Equations', *Physical Review*, **4**, pp. 345–76.

Cartwright, N. [1999]*: The Dappled World: A Study of the Boundaries of Science*, New York: Cambridge University Press.

Chanson, H. [2009]: 'Turbulent Air-Water Flows in Hydraulic Structures: Dynamic Similarity and Scale Effects', *Environmental Fluid Mechanics*, **9**, pp. 125–42.

Frigg, R. and Nguyen, J. [2017]: 'Models and Representation', in L. Magnani and T. Bertolotti (*eds*), *Springer Handbook of Model-Based Science*, New York: Springer, pp. 49–102.

Frigg, R. and Nguyen, J. [2018]: 'The Turn of the Valve: Representing with Material Models', *European Journal of the Philosophy of Science*, **8**, pp. 205–24.

Fritz, H. M., Hager, W. H. and Minor, H. E. [2001]: 'Lituya Bay Case: Rockslide Impact and Wave Run-Up', *Science of Tsunami Hazards*, **29**, pp. 3–23.

Fritz, H. M., Hager, W. H. and Minor, H. E. [2004]: 'Near Field Characteristics of Landslide Generated Impulse Waves', *Journal of Waterway, Port, Coastal, and Ocean Engineering*, **130**, pp. 287–302.

Fritz, H. M., Mohammed, F. and Yoo, J. [2009]: 'Lituya Bay Landslide Impact Generated Mega-tsunami 50th Anniversary', *Pure and Applied Geophysics*, **166**, pp. 153–75.

Godfrey-Smith, P. [2006]: 'The Strategy of Model-Based Science', *Biology and Philosophy*, **21**, pp. 725–40.

Heller, V. [2011]: 'Scale Effects in Physical Hydraulic Engineering Models', *Journal of Hydraulic Research*, **49**, pp. 293–306.

Heller, V. and Hager, W. H. [2010]: 'Impulse Product Parameter in Landslide Generated Impulse Waves', *Journal of Waterway, Port, Coastal, and Ocean Engineering*, **136**, pp. 145–55.

Heller, V., Hager, W. H. and Minor, H. E. [2008]: 'Scale Effects in Subaerial Landslide Generated Impulse Waves', *Experiments in Fluids*, **44**, pp. 691–703.

Jebeile, J. and Kennedy, A. G. [2015]: 'Explaining with Models: The Role of Idealizations', *International Studies in the Philosophy of Science*, **29**, pp. 383–92.

Langhaar, H. L. [1951]: *Dimensional Analysis and Theory of Models*, New York: Wiley and Sons.

McFall, B. C. and Fritz, H. M. [2016]: 'Physical Modelling of Tsunamis Generated by Three-Dimensional Deformable Granular Landslides on Planar and Conical Island Slopes', *Proceedings of the Royal Society A*, **472**, p. 20160052.

McFall, B. C., Mohammed, F., Fritz, H. M. and Liu, Y. [2018]: 'Laboratory Experiments on Three-Dimensional Deformable Granular Landslides on Planar and Conical Slopes', *Landslides*, **15**, pp. 1713–30.

Nguyen, J. [forthcoming]: 'It's Not a Game: Accurate Representation with Toy Models', *British Journal for the Philosophy of Science*, available at <doi.org/10.1093/bjps/axz010>.

Odenbaugh, J. and Alexandrova, A. [2011]: 'Buyer Beware: Robustness Analyses in Economics and Biology', *Biology and Philosophy*, **26**, pp. 757–71.

Pincock, C. [2014]: 'How to Avoid Inconsistent Idealizations', *Synthese*, **191**, pp. 2957–72.

Rice, C. [2019]: 'Models Don't Decompose That Way: A Holistic View of Idealized Models', *British Journal for the Philosophy of Science*, **70**, pp. 179–208.

Schiermeier, Q. [2017]: 'Huge Landslide Triggered Rare Greenland Mega-tsunami', *Nature*: *News*, available at <doi.org/10.1038/nature.2017.22374>.

Skow, B. [2017]: 'The Metaphysics of Quantities and Their Dimensions', *Oxford Studies in Metaphysics*, **10**, pp. 171–98.

Sterrett, S. [2017a]: 'Experimentation on Analogue Models', in L. Magnani and T. Bertolotti (*eds*), *Springer Handbook of Model-Based Science*, New York: Springer, pp. 857–78.

Sterrett, S. [2017b]: 'Physically Similar Systems: A History of the Concept', in L. Magnani and T. Bertolotti (*eds*), *Springer Handbook of Model-Based Science*, New York: Springer, pp. 377–411.

Suárez, M. and Cartwright, N. [2008]: 'Theories: Tools versus Models', *Studies in the History and Philosophy of Modern Physics*, **39**, pp. 62–81.

Thomson-Jones, M. [2005]: 'Idealization and Abstraction: A Framework', in M. Jones and N. Cartwright (*eds*), *Idealization XII: Correcting the Model*, New York: Rodopi, pp. 173–217.

Weisberg, M. [2007a]: 'Three Kinds of Idealization', *Journal of Philosophy*, **104**, pp. 639–59.

Weisberg, M. [2007b]: 'Who Is a Modeler?', *British Journal for the Philosophy of Science*, **58**, pp. 207–33.

Weisberg, M. [2013]: *Simulation and Similarity: Using Models to Understand the World*, New York: Oxford University Press.

Weisberg, M. and Reisman, K. [2008]: 'The Robust Volterra Principle', *Philosophy of Science*, **75**, pp. 106–31.

Woodward, J. [2003]: *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.

Woodward, J. [2013]: 'Mechanistic Explanation: Its Scope and Limits', *Aristotelian Society Supplementary Volume*, **87**, pp. 39–65.

Woodward, J. [2015]: 'Interventionism and Causal Exclusion', *Philosophy and Phenomenological Research*, **91**, pp. 303–47.

Woodward, J. [2016]: 'The Problem of Variable Choice', *Synthese*, **193**, pp. 1047–72.

Woodward, J. [forthcoming]: 'Explanatory Autonomy: The Role of Proportionality, Stability, and Conditional Irrelevance', *Synthese*.