1 Introduction

1.1 Models

Imagine that you're a shipbuilder working with ocean liners like SS *Monterey* (Figure 1a). Matson Navigation Company has re-purchased the liner from the US government, to whom they had sold the ship following its utilisation during the Second World War. The company now deems the ship too slow for its San Francisco–Los Angeles–Honolulu run, and tasks you to redesign its engine so that it will be able to sail at a certain speed whilst carrying a certain load. What is the minimal power the engine must have to ensure it's up to the task? You could of course just make a guess, install a certain engine, and see whether the ship runs at the right speed when it's back in the water. If you're lucky, the ship works as it is supposed to. But there's a good chance it won't, and that would be a costly failure. A better way to proceed is to construct a model of the ship, a scaled-down version of the real ship you're overhauling, and perform experiments on that model. The model must be carefully constructed: it has to have a shape that reflects the shape of the full-sized ship you're ultimately interested in. And the experiments that you perform on the model have to be carefully designed. In this case, you want to measure the complete resistance, R_C, faced by the model ship as it is propelled through the basin at velocity V, because this gives you information about how powerful the engine needs to be. Experiments of this kind are standard practice in the process of designing ships. In Figure 1b, we see a model ship being moved through a towing tank.

But however carefully you construct your model, and however carefully you perform experiments on it, you're not investigating the model for its own sake. Ultimately, you want to use the results of your investigations to inform you of another system: SS *Monterey* at sea. So you have another task at hand: you have to translate your experimentally discovered facts about the model into claims about the full-sized ship. This translation procedure is subtle and complex: it is informed by our theoretical background knowledge about fluid mechanics, and clever ways of thinking about things like scale, length, and resistance. We'll come back to these later in the Element. For now, what's important is the general pattern of reasoning: the shipbuilder first constructs a scaled-down model of the ship, investigates how the model behaves, and then translates facts about their model into claims about the actual ship.

Now leave ships behind and imagine you've landed a job as a stunt planner for the next instalment of the 007 franchise. You are planning an exhilarating car chase through the streets of London, culminating in the British spy launching

Cambridge University Press & Assessment
978-1-009-00915-7 — Scientific Representation
James Nguyen , Roman Frigg
Excerpt
More Information

Figure 1a The target system: An ocean liner like SS *Monterey*

Figure 1b Model ship in towing tank

Figure 2a Tower Bridge half open[1]

their Aston Martin across Tower Bridge.[2] The bridge is a drawbridge and the car is supposed to jump over the bridge when it's half open, as displayed in Figure 2a. The producers tell you the angle, α, at which they want the leaves (i.e. the arms) of the bridge for dramatic effect. How fast does 007 have to drive at the jump off point to ensure that the car lands safely on the other side?

Unlike our shipbuilder, you don't produce a scale model of the situation in which you make a small remote-controlled model car jump across a model bridge. Instead, you revert to the power of the imagination and the laws of Newtonian mechanics. You imagine a scenario with two inclined planes facing each other with a gap in the middle. You then imagine a perfect sphere moving up one of the planes with constant velocity v. You imagine that both the planes and the sphere are on the Earth's surface, that all other material objects in the universe have vanished, and that the planes and the sphere are in a vacuum. On the basis of these assumptions, the Earth's gravity is the *only* force acting on the sphere; that is, the sphere is not subject to all the other forces that a real car jumping across a real bridge would experience, such as air resistance and the

[1] Credit: 'Tower Bridge Open' by Tony Hisgett from Birmingham, United Kingdom is licensed under CC BY 2.0

[2] Readers might be familiar with similar stunts, including a jump over Tower Bridge in the 1975 movie *Brannigan*; the jump on the 95th Street Bridge in Chicago in the 1980 movie *The Blues Brothers*; the bus jump over an incomplete overpass in the 1994 movie *Speed*; or the jump involving two cars simultaneously in the 2003 movie *2 Fast 2 Furious*.

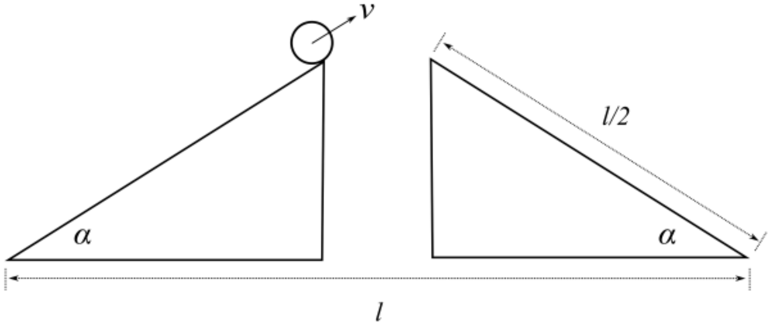

Figure 2b Sketch of the bridge jump model

gravitational pull of other pieces of matter in the universe. This is your model of the bridge jump, where, of course, the two inclined planes stand for the two leaves of the bridge and the sphere for the car. The model is sketched in Figure 2b.

You now apply Newtonian mechanics to your model and find that the equation of motion for the sphere is:

$$m\begin{pmatrix}\ddot{x}\\ \ddot{y}\end{pmatrix} = \begin{pmatrix}0\\ -mg\end{pmatrix}, \tag{1}$$

where the x is the horizontal, and y the vertical, coordinate of the ball; g is the gravitational constant on the surface of the Earth; m is the mass of the sphere; and the two dots on x and y indicate the second derivative with respect to time (and recall that this second derivative of a position is acceleration, and so $\ddot{x}$ and $\ddot{y}$ correspond, respectively, to the sphere's acceleration in the horizontal and vertical directions). According to this equation, the sphere moves at a constant horizontal velocity, and accelerates towards the ground at a rate equal to mg. The general solution to this equation is:

$$\begin{pmatrix}x\\ y\end{pmatrix} = \begin{pmatrix}v_x t + x_0\\ -\frac{g}{2}t^2 + v_y t + y_0\end{pmatrix}, \tag{2}$$

where x_0 and y_0 are the initial conditions (that is, the position of the sphere when it starts the jump); v_x and v_y are the components of the sphere's velocity in the x-direction and the y-direction, respectively; and t is time. From this general solution, you can calculate the minimal velocity that the sphere must have to land on the other inclined plane without falling into the gap between them:

$$v = \sqrt{\frac{gl(1-\cos\alpha)}{\sin 2\alpha}}. \tag{3}$$

Given the angle α that the producers want for dramatic effect, and the length l of the bridge, this formula tells you the minimum velocity with which the sphere must move up the inclined plane to fly across the gap and land safely on the other side.

But real cars (even Aston Martins) aren't spherical and don't move in a vacuum. And you're not interested in spheres in vacuums per se. What you really want to know is the velocity at which the actual car has to move up the half open bridge to avoid plummeting into the Thames. So the question is: what can you learn about real cars jumping across half open drawbridges from spheres moving on inclined planes in a vacuum? To answer this question, you have to face the task of translating facts about the model into claims about the actual situation, just like you did when you faced the task of redesigning a ship. Again, we will see in Section 4 how this should be done.

Despite their differences, both of these scenarios rest on a common style of reasoning. You investigate one system, a *model*, and use the results of that investigation to inform yourself about another system, a *target*. In the first case, the model is a concrete physical system itself; it is constructed out of steel (or wood, or paraffin wax, depending on the era) and the shipbuilder performs physical experiments directly on the model. In the second case, the model is a combination of an imagined scenario and mathematical equations (derived from Newtonian mechanics), which the stunt planner can investigate, and there are facts about this model and the solutions of its equations. But neither the shipbuilder nor the stunt planner is interested in their models per se; they are interested in what the models are about. So to reach the end point of their investigations they have to translate their model results into claims about another system: the full-sized ship moving through water and the car jumping the bridge.

The aforementioned examples are not isolated instances of some outlandish style of reasoning. Models are used across the sciences; they are one of the primary ways in which we come to learn about the world. Scientists construct models of atoms, black holes, molecules, polymers, populations, DNA, rational decisions, financial markets, climate change, and pandemics. Models provide us with insight into how selected parts or aspects of the world work, and they act as guides to action. Much of scientific knowledge, and understanding, is ultimately based on the results of some modelling endeavours.

How do models work? How can the investigation of a model possibly tell us anything about something beyond the model, some system out there in the world? Our answer here is that models do this because they *represent* their targets. Just as the scale model in the shipbuilder's tank represents the full-sized ship and the stunt planner's model represents the actual car, the physicist's model of a black hole represents parts of the universe from which no light can escape; the economist's macroeconomic model represents the actual economy of a given country; and the epidemiologist's model represents how a disease will spread through a country as a result of various policy interventions by a government.

As these examples suggest, models can be used to represent *particular* target systems like a particular government's economy, a specific ship, and so on. But they can also be used to represent *types* of targets. Depending on the details, economists might employ models to reason about economies in general; a physicist can use a model to represent a type of atom like hydrogen; the model ship can be used to represent a certain type of ship; and the mathematical model of 007's stunt can also be used to represent a type of stunt involving car jumps across bridges (a few exemplars of which are mentioned in Footnote 2). So by 'target system' we can mean both specific systems and types of systems.

In each of these cases the model stands in for 'the' target system; the model is the secondary system that scientists investigate, with the hope that the results of their model-based investigations will deliver insight into their targets. So in order to understand how model-based reasoning works, we need to understand how models represent. In this Element, we provide a philosophical investigation into this question.

At this point, one might worry that our focus on models is too narrow. Scientists use plenty of other kinds of representations to reason about systems in the world: doctors use MRI scans to learn about brain structure; particle physicists pore over bubble chamber photographs to learn about the nature of subatomic particles; and astronomers study the images produced with telescopes. But it pushes the limits of language to deem any of these kinds of representations models. Moreover, there are plenty of non-scientific representations that function in a similar way. All of us are familiar with using maps to navigate new cities, and we regularly make inferences about the subjects of photographs based on features of the photograph. In general, we call a representation that affords information about its target an *epistemic representation*, and it is clear that models are just one kind of such representation. Given this, there is a question whether our analysis in this Element covers things beyond models; whether it covers epistemic representations more generally. By and large we think it does, and most of the existing accounts of how models represent end up being accounts of epistemic representation more generally.[3] In this Element, we primarily focus on models because models are crucial to the

[3] See the discussion of the representational demarcation problem in Frigg and Nguyen (2020, Ch. 1), and the discussion of how the different accounts handle demarcation in later chapters of that book.

working of modern science, and because discussing different kinds of representations side by side would end up using more space than we have. We will briefly broaden our scope in Section 4, and readers who are interested in other kinds of epistemic representations – images, and certain works of art, for example – are encouraged to consider how what we say applies to these as they proceed through the following sections.

1.2 Questions Concerning Scientific Representation

At first glance it might seem like there is only one question to be asked here: how does a model represent its target? But looking a little closer, we see that this question breaks up into several different questions. To have a clear focus in our investigation, it's important to disentangle these and clarify how answers constrain the shape of the conceptual landscape concerning how models work. This is the project for this section, in which we lay out the relevant questions; in the next, we discuss what it takes to answer them appropriately.

The first, and most fundamental, question to investigate is: in virtue of what does a model represent its target? Call this the *Semantic Question*.[4] Answering this will allow us to understand how the steel vessel that is dragged through the towing tank comes to represent a real ship, and how the shipbuilder manages to translate results of model experiments into claims about a full-sized ship; likewise, it will allow us to understand how an imagined scenario consisting of two inclined planes and a sphere in a vacuum comes to represent a real-world bridge jump, and how the stunt planner manages to plan the jump based on the model.

It's of paramount importance that we don't confuse this question with a closely related one, namely: what makes a model an *accurate* representation? A model can represent its target, without doing so accurately. To see this, alter the aforementioned examples slightly. You may not have a very good initial idea of the ship's shape because you haven't been able to measure it up and no plans are available. So you may decide to use an empty barrel as a model of the ship. When you finally see the ship, you realise that this is a bad model because it doesn't have the form of the ship at all. Nevertheless, the barrel is a representation of the ship; it's just not an accurate representation. Or assume that an error has been made in measuring the angle of the open leaves of the drawbridge, and the angle in reality is twice the angle in your model. As a result, the model will underestimate the velocity needed to get to the other side, and the

4 We're using the term 'semantic' in the broad sense of referring to the relationship between symbols and what they are about. Thus understood, a discussion of how models relate to their targets falls within the scope of semantics.

car, along with the stunt driver, will plummet into the river. If this happens, the model is not an accurate representation of the target system, but it is a representation of it nevertheless.

The lesson is that we should distinguish between the question of what turns something into a representation of something else to begin with (the Semantic Question), and the question of what turns something into an accurate representation of something else. Call the latter the *Accuracy Question.* The Semantic Question is conceptually prior in that asking what makes a model an accurate representation presupposes that it is a representation in the first place: a model cannot be a *mis*representation unless it is a representation. But once this has been established, there is a genuine question about what it takes for a representation to be accurate.

The distinction between representation and accurate representation is not an artefact of the simple examples we have used to illustrate it. The history of science provides us with a wealth of examples of inaccurate, but nevertheless representational, models. Ptolomy's model of the solar system represents the solar system even though it is inaccurate with respect to the orbit of the Earth around the Sun. Thompson's 'plum pudding' model represents atomic structure, but it is inaccurate with respect to the distribution of charge within an atom. Fibonacci's model of population growth is inaccurate with respect to the long-term growth of a population because it assumes that organisms are immortal and food supplies are unlimited. But despite their inaccuracies, all of these models represent their targets.

So far in this section we've distinguished between the question of what makes a model represent, and what makes it accurate. But some might worry that there's an even more prior question lurking in the background: what is a 'model'? Call this the *Model Question*. In the case of the model ship the answer is relatively straightforward: the model is the concrete material object towed through the tank. But many models aren't like this; they are, to use Ian Hacking's memorable phrase, things that we 'hold in our heads rather than our hands' (1983, 216). It has become customary to refer to such models as 'non-concrete' models. The model of the car jump is of this kind. Earlier we said it was an imagined scenario combined with mathematical equations from Newtonian mechanics, and we also said that there were facts about this model that the planner would investigate. One option then, would be to identify this model, and others like it, with mathematical entities (and then leave mathematicians and philosophers of mathematics to work out what they are). But there's a lingering worry that this isn't the whole story. The car jump model might involve mathematics, but it's not obviously purely mathematical. Before writing down equations, the stunt planner had to imagine a scenario with a sphere moving in the vacuum on an inclined plane. And it's not clear that this can be accounted for by identifying the model with something purely mathematical.

A philosophical account of modelling should have something to say about how we should think about models of this kind. This goes beyond metaphysical bookkeeping; as we will see, how one answers the Model Question has implications for how we understand the Semantic Question and the Accuracy Question.

1.3 What Does Success Look Like?

The next thing to establish is the success conditions on answers to these questions. What counts as a successful answer to the aforementioned questions?

We begin with conditions on answers to the Semantic Question. First and most straightforwardly, there is a *direction* to the representation relationship that holds between models and their targets. Typically at least, models represent their targets, but not *vice versa*: the scale model in the towing tank represents SS *Monterey*, but SS *Monterey* doesn't represent the scale model. We say 'typically' here because we are not assuming that it's a conceptual impossibility; in some special cases a representational relationship can hold both ways. Rather, we require that any answer to the Semantic Question should not entail that the model–target representation relation is always symmetric. We call this the *Directionality Condition*.

Another important condition of success on answering the Semantic Question is accounting for the fact that models are *informative* about their targets. Some representational relationships do not work this way: the term 'atom' can be seen as representing (at least in some sense) atoms, but it's uninformative: no investigation into the term itself will allow us to extract any information about what it refers to. In contrast, a model represents its target in a way that does allow such information extraction (although, of course, that information doesn't have to be true, because models don't have to be accurate). To use Swoyer's (1991) phrase, models allow for *surrogative reasoning*: by investigating the behaviour and features of the model, scientists can generate claims about the behaviour and features of its target. We call this the *Surrogative Reasoning Condition*.

The distinction between representation and accurate representation motivates a further condition for success: any answer to the Semantic Question should be compatible with the fact that models can misrepresent their targets; no answer should entail that all representations are accurate, nor that inaccurate models are non-representations. In brief, a viable answer to the Semantic Question must distinguish between misrepresentation and non-representation. This is not to say that all models are inaccurate (although there is reason to think that no model represents with perfect accuracy). Indeed, much of the motivation for investigating how models work stems from the fact that at least some of them are accurate; some of them are paradigmatic instances of the cognitive success of the scientific endeavour. But it remains that some models misrepresent, and thus whatever it is that establishes a representational relationship should not equate representation with accurate representation. We call this the *Misrepresentation Condition*.

The previous condition relates to models that represent actual targets in the world (whether specific systems, or types of systems), but do so inaccurately. We should also recognise that some models don't represent any actual target whatsoever. Straightforward examples of models of this sort include engineering models of structures never built. Vary our initial example slightly and assume that you are tasked with designing a new ship rather than redesigning an existing one. But for some reason the ship is never built. In that case your model doesn't represent anything in the world. The same goes for architectural models of buildings that have never been constructed and models of spacecraft that have never been realised. Targetless models aren't unique to fields that are in the business of constructing something; such models also appear in theoretical science. Population biologists construct and investigate models involving a population consisting of four different sexes to see how such a population would develop; elementary particle physicists study models of particles that don't exist to learn about techniques like renormalisation; and philosophers of physics construct models in accord with the principles of Newtonian mechanics in order to demonstrate that the theory is consistent, under certain conditions, with indeterminism, without the model representing, or indeed being intended to represent, any system in the world.[5] A philosophical account of modelling should accommodate models that don't have real-world targets. We call these 'targetless models', and the condition that they be allowed for the *Targetless Models Condition*.

We now turn to the Accuracy Question. Unlike truth, which (many people think) is an all-or-nothing matter, accuracy comes in degrees. Models can be more or less accurate, depending both on the scope of the features they represent, and on how well they represent those features. And the fact that a model misrepresents some aspects of its target doesn't entail that it misrepresents all aspects. For example, whilst the Ptolemaic model of the solar system might misrepresent the structure of the celestial orbits, it accurately represents, to some degree, the relative position of the celestial bodies as seen in the night sky

[5] For discussions of n-sex populations, see Weisberg (2013). For discussion of elementary particles, see Hartmann (1995). For discussions of Newtonian mechanics and indeterminism, see Norton (2008) for 'Norton's Dome' and Xia (1992) for a model involving charmingly named 'space-invaders'.