

MODELS AS FELICITOUS FALSEHOODS

CATHERINE Z. ELGIN

Harvard University, USA

catherine_elgin@harvard.edu

Abstract. I argue that models enable us to understand reality in ways that we would be unable to do if we restricted ourselves to the unvarnished truth. The point is not just that the features that a model skirts can permissibly be neglected. They *ought* to be neglected. Too much information occludes patterns that figure in an understanding of the phenomena. The regularities a model reveals are real and informative. But many of them show up only under idealizing assumptions.

Keywords: Felicitous falsehoods • epistemic access • distortion • as if

RECEIVED: 11/10/2021

ACCEPTED: 22/02/2022

1. Introduction

Modern science is one of humanity's greatest epistemic achievements. It constitutes a rich, variegated understanding of the natural world. Epistemology could readily accommodate its extraordinary success if that understanding were expressed in accurate representations of the phenomena. But it is not. Science couches its deliverances in models that are purposefully inaccurate — models that simplify, augment, exaggerate and/or distort. Nor does it construe modeling as a temporary expedient. Although scientists anticipate that current models will be supplanted, they expect today's models to be replaced by better models, not by the unvarnished truth. That being so, science's pretension to epistemic eminence may look a bit dodgy. The regular, unabashed divergence of accepted, highly touted models from the truth seems intellectually suspect. How can science claim to supply understanding *of* the phenomena, if its representations are not and do not purport to be literally true of the phenomena they concern?

Maybe my concern is misplaced. Nature is complicated; understanding it is hard. Perhaps we should just own up to our limitations and construe scientific models as approximations. Then we might contend that scientific progress consists in devising ever better approximations.

Although construing models as approximations is often plausible, it is not as helpful as we might hope. A variety of sorts of divergence from accuracy count as approximations. Some models are approximately accurate in all the cases that they



pertain to; there's just a small margin of error. Others are accurate (or close to accurate) in nearly all the cases, but admit of a few outliers. Yet others are accurate (or nearly accurate) in nearly all the cases that scientists particularly care about, even though these are a relatively small subset of all the cases that the models purport to pertain to. Snell's law is a case in point.¹ It is accurate only for isotropic media. But most media are anisotropic. Still, scientists insist, isotropic media are interesting and important. Some models aggregate; although they are approximately true of the phenomena taken collectively, they are not approximately true of any particular case. No one comes close to being an economically rational agent. But models that treat of such agents are effective because the ways individuals diverge from economic rationality cancel out in the aggregate. At best then, the operative notions of approximation show a family resemblance; handwaving about approximation does not get us very far. If we are going to account for the epistemic success of science by saying that models are good approximations, we need to explain what makes a particular approximation and a particular sort of approximation good, recognizing that there are a variety of things we might reasonably mean by 'approximation' and a variety of ways an approximation might be good.

We might think that, like the anomalies that plague current science, models are temporary expedients that will be eliminated when science reaches its goal. Maybe at the limit of inquiry, science has no need of such devices. In ultimate science, every truth-apt representation would be literally true. Teller strongly suggests otherwise, arguing that truth and precision inevitably pull apart. A simple example illustrates his point. The velocity of sound in dry air depends on the temperature, the pressure, and the proportions of the different sorts of gas molecules that compose the air. Although we can give averages, we are in no position to precisely specify the temperature, pressure, and proportions of the different sorts of molecules in a sample of air. Nor, Teller maintains, is this just an epistemic problem. He says, "No real world sample of air has such precise values, if only because the values would vary from place to place. So at best one is talking about the speed of sound in some idealized condition, not in the real world".² Because the temperature, pressure, and proportions of different sorts of gas molecules vary slightly from one region of the sample to another, we cannot say truly that sound travels with exactly velocity v throughout the sample. To say something true here we need to idealize, treating the magnitudes as uniform across the sample. This is what we do and what we should do. But we do not thereby mirror the actual (fluctuating) speed of sound across the air sample. Teller's point is not just that contemporary science contains idealized models that are strictly false; rather, science *consists* of them (see Teller, 2018).

The widespread use of inaccurate models raises important epistemological questions. First, why does science rely on models that are designed to deviate from the truth? Second, how — if at all — do such models contribute to understanding?

I will argue that models enable us to understand reality in ways that we would be unable to if we restricted ourselves to the unvarnished truth. The point is not just that the features that a model skirts can permissibly be neglected. They *ought* to be neglected. Too much information occludes patterns that figure in an understanding of the phenomena. The regularities a model reveals are real and informative. But many of them show up only under idealizing assumptions.

Potochnik (2017) maintains that science's dependence on idealized models is due to the confluence of two factors: first, nature is enormously complicated; and second, human minds are limited. Both claims are true. Nevertheless, I think that the emphasis on human limitations undervalues and distorts the epistemic significance of modeling. Rather than being cognitive crutches on which scientists regrettably lean, or heuristics which they are unfortunately fated to deploy, I believe that effective models are powerful tools that extend our epistemic range. It might seem that in saying this, I concede Potochnik's point. Perhaps it is only because our range is limited that we benefit from its extension. I'm not convinced. Before turning to models in particular, let us consider a couple of analogous cases.

Eye-glasses are called 'corrective lenses' because they are devices for bringing people with defective vision to the level of visual acuity that other humans naturally achieve. They compensate for a deficiency. Lenses in ocular microscopes and telescopes are not *corrective*. They *extend* perceptual access. With their aid, humans see what otherwise none of us could. Other instruments do more. Radio telescopes and electron microscopes, MRI scanners and the like do not extend perception; they augment it, enabling us to detect what human perception cannot reveal. Still, Potochnik might say, such extensions are needed only because we do not naturally have the ability to see very small things or very distant things or directly observe the internal workings of living organisms or whatever. Their value derives from our perceptual limitations. Maybe so.

Consider then a non-perceptual case. Prior to the invention of calculus, physics was hampered by the inability to compute the instantaneous rate of change of an accelerating object, the area under a curve, the exact tangent to a curve, and so on. Calculus enables us to perform such calculations, among a host of others. It massively extends our epistemic range. No one construes differential equations as mere heuristic devices, unfortunately necessary crutches, or modes of representation that will be eliminated at the end of inquiry. We fully expect ultimate science, if there is such a thing, to express its findings in mathematical formulas, some of which will probably be differential equations. If they are not, it will be because mathematicians have invented even more powerful methods. Newton invented calculus expressly to extend science's epistemic range — to enable it to represent phenomena in terms of continuous functions, to compute instantaneous rates of change, to make other calculations that could otherwise not be done.³ Calculus affords a cognitively rich and

rewarding way to represent and reason about things. It is not a cognitive prosthesis that enables us to accommodate our unfortunate natural failings.

I want to say the same about modeling. Like both instrumentation and mathematics, modeling extends our epistemic range. Like mathematics, it provides representations that characterize phenomena in ways that facilitate informative reasoning about them.

Still, even if we recognize that instruments expand epistemic access, we might doubt that models do too. Scientific instruments are, and are designed to be, data gatherers. The more powerful our instruments, the more or better data they afford access to. But models do not gather data. Like theories, they process or interpret data. If epistemic access is fixed by our data-gathering abilities, models do not afford epistemic access. Nor, presumably, do theories. In this respect the distinction between theories and models is idle.⁴

This objection goes too fast. First, it is not the case that the limits of our epistemic access are fixed by the data we can gather. What matters is the data we can process. That depends on both our data-gathering and our data-processing capabilities. If we devise new ways to process old data, we expand our access to, and often our understanding of, the phenomena they concern. This is why innovations in statistical modeling are often fruitful. They enable us to glean more information from data we already have. Models tell us what to measure, how to measure, with what precision to measure, and where to focus attention. Theories do the same. According to Hacking, “Once the theory of pulsars was in place, older astrophysical records were shown to be rich in evidence of pulsars” (Hacking 1988, p.511). The theory transformed what was once dismissed as noise into data.

Second, data are not things out there in the world that we sometimes manage to trip over. They are products of curation. Information qualifies as data only because it answers to specific requirements set by the models or theories it bears on. These requirements specify how items are to be individuated, how they are to be conceptualized, and how they are to be detected and/or measured. Seemingly relevant information is justifiably disregarded if it does not fit the demands set by the relevant models or theories. Theoretical commitments thus constitute information as data.

Third, scientific instruments do more than gather data. Some process what is found in nature to generate pure samples. Others interfere with such samples in the lab. That is, they do something to the samples to produce a detectable and, it is hoped, informative change (see Hacking, 1988). The access to nature that scientific instruments provide is informed and shaped by theories that dictate what is to be eliminated in purifying material, and by theories of instrumentation which both enable scientists to design suitable devices and that explain their affordances and limitations. To be sure, some scientific instruments are designed to register and gather data. They detect and record the information they do because they were expressly

constructed to answer to theoretical requirements. This is so even for relatively simple instruments like thermometers (see Chang, 2004). But it is particularly vivid in the case of complex instruments. An MRI scan is a statistical synthesis of diverse inputs that the scanner registers. The device was designed to register exactly the information that would contribute to the statistical synthesis and to omit information that was irrelevant. Clearly theoretical commitments lie behind the design specifications. The Large Hadron Collider uses models to synthesize outputs — to transform them into data. Completely unprocessed, its outputs would be unintelligible. Through a series of iterations, it produces tractable data.⁵

The point about instrumentation is an elaboration of the uncontroversial point that even unmediated observation is theory laden. Observing a tree by looking at it with the naked eye requires ignoring many inputs into our visual system. It involves synthesizing the ones that remain. It involves settling on a level of generality — seeing it as a tree rather than, for example, as an elm, or a plant, or an organism apparently thriving in New England, or whatever. It involves taking a particular perspective on the tree — not too close, not too far, not where it is occluded by the barn. Here the counterpart of a model is a perceptual schema which equips the observer to interpret visual inputs as of a tree.

Observation, instrumentation and model building operate in tandem. What observations scientists can make depends on what instruments they have designed, calibrated, and validated. These in turn depend on what the relevant, available models or theories are, since they are integral to the design, calibration, and validation of instruments. Models thus inform the design of instruments that will detect the phenomena and supply data. Models process the data yielding information that can be imputed to their targets. Because models are integral to the design of scientific instruments, they are integral to the epistemic access that empirical investigations provide. The critical point, made by Suárez (2009), Knuuttila (2011), and Hughes (1996), is that models and theories are not merely reflections of the phenomena. They are things we think with.

Effective models are powerful tools, not regrettable expedients that we fall back on only because our reach exceeds our grasp. And they are powerful tools precisely because representing things as, strictly speaking, they are not can equip us with epistemic resources that foster systematic, fruitful, empirically grounded accounts of things.

2. Felicitous Falsehoods

Effective models are what I have called *felicitous falsehoods* (Elgin, 2017). They are (typically) representations of phenomena; that is, they typically have targets. But

they purposefully misrepresent those targets. Although some, being non-propositional, are not strictly false, I label them falsehoods because they misrepresent. I consider them felicitous because their inaccuracies are epistemically fruitful; they are not defects. The falsehoods are inaccurate in ways that enable them to non-accidentally provide epistemic access to obscure or occluded aspects of their targets. Not despite, but because of their inaccuracy, they afford the access that they do. When inaccurate models facilitate reasoning about the phenomena, they are epistemically felicitous. Henceforth, unless I say otherwise, I shall be talking about effective models. Ineffective models are ones that fail to perform or inadequately perform their epistemic functions.

Strevens (2008, 2017) takes a similar line. He maintains that effective models are epistemically legitimate because their inaccuracies are not difference-makers.⁶ Although real gas molecules are not perfectly spherical, their divergences from sphericity make no difference to their Boylean behavior. So when we seek to understand Boylean behavior in gases, it does no harm to treat the molecules as spheres. Similarly, although the relevant attitudes of each individual car buyer are idiosyncratic and often mutually inconsistent, collectively the divergences from economic rationality cancel out; thus seemingly intractable idiosyncrasies are not difference-makers. That being so, it does no harm to ignore them and represent each potential customer as an economically rational agent. The economic model will not, of course, provide a profile of the deliberations and actions of any individual customer. Nor will Boyle's model provide a profile of the career of any particular gas molecule. Each will, however, provide a representation of what happens in the aggregate: irrelevancies and idiosyncrasies wash out.

The adequacy of a model depends on the context of inquiry where it is deployed. For a model to be epistemically effective, Strevens maintains, inaccuracies in non-difference-makers are unproblematic (2017). If this is right, the use of idealized models is epistemically permissible. Since a felicitously false model does not impute all of its seemingly relevant properties to its target, in purporting to be a model *of* that target, it need not say or implicate anything false *about* the target.

This does not yet explain why models are epistemically advantageous. Perhaps the answer is simply that they are the best we can do. Even if we cannot devise or grasp a fully accurate, fine-grained representation of the automobile market — why Fritz bought a blue Toyota, while Sasha opted for a red Audi, and so forth — maybe we can devise a representation that will explain why overall car sales trend as they do. That would not be nothing. But, I suggest, the simplifications, distortions, and emendations of effective models do more. They afford epistemic access to patterns in the phenomena that the unvarnished truth would obscure.

Models are schematic representations that rigorously and systematically omit irrelevancies. If they did no more than omit, characterizing them as falsehoods, even in

my extended sense, would be unfounded. Every representation omits something; the truth is never the whole truth. Distortions are different. They plainly falsify. Many models effect their simplifications via distortion. Here is an example: Droplet formation is a complex hydrodynamical phenomenon, where the shape, velocity, and curvature of the fluid surface change continuously. Drawing on Batterman (2009), Potochnik describes a model of how a stream of water forms drops as it drips from a faucet. She says,

[C]ertain idealizations can be made in virtue of the system approaching a hydrodynamical discontinuity — that is, the breaking point — and those idealizations simplify matters considerably [. . .] [T]he water can be treated as if it were a vertical line [. . .] because its axial extension (length) is much greater than its radial extension (width) as the discontinuity is approached. This idealization enables a one-dimensional solution to an otherwise quite complicated set of equations (the Navier-Stokes equations). [. . .] Approaching the discontinuity, the relationship among surface tension, viscous forces, and inertial forces is such that acceleration due to gravity can be wholly neglected, even though it is present. [. . .] [T]he axial and radial lengths both become, at the discontinuity, arbitrarily small. For this reason the shape of the fluid at the point of discontinuity is expected not to depend on the size of the system, that is, of the fluid mass and surface area. [. . .] These idealizations [. . .] demonstrate why different liquids, forming droplets in a variety of circumstances, have the same shape at the breaking point. (Potochnik 2017, pp.80–1)

This constellation of simplifications is elegant, effective, and efficient. Numerous liquids share these features and thus exhibit the same behavior.

To understand the droplet formation without such a model would require solving the Navier-Stokes equations (which no one knows how to do), calculating the continuously changing shape, surface area, and velocity of the water as the drop formed, incorporating the effect of gravity, and the size, shape, and material properties of the faucet's nozzle, taking into account the actual radial and axial dimensions, the surface tension at every instant, and so forth. Potochnik rightly emphasizes that we couldn't do all this. That underscores her claim about human limitations.

My point is different: even if we did it, that would not put us in a position to appreciate some of what the idealization discloses — that gravity, although real, is irrelevant; that a one-dimensional solution suffices; that surface tension, viscous forces, and inertial forces are equally significant. Nor would we be in a position to appreciate how the solution generalizes to other liquids, or why the generalization stops holding where it does. There is no reason to think that we would be able to recognize that beneath the vast array of details that distinguish different cases of droplet formation the same pattern holds. Patterns emerge when details are neglected. If we

accommodated all the differences in detail, the patterns of droplet formation in diverse liquids would be different. To recognize the common pattern requires ignoring specific differences. Too much information is an impediment to understanding.

Sometimes reality itself distorts. Features of the phenomena can warp underlying patterns. This is why many statistical models represent populations as infinite. Doing so swamps chance effects that afflict finite populations. To understand the roles of non-random factors, we devise a model where chance makes no difference. Thus models in population biology represent species as infinite in size. To prescind from the (currently irrelevant, but nonetheless real) consequences of genetic drift, for example, they construct a fictional scenario in which genetic drift would have no discernible effect. Both the fact that there are no infinite populations⁷ and the fact that genetic drift is always present are sometimes irrelevant. If one wants to understand how alleles would redistribute in the absence of genetic drift, representing the population as infinite is appropriate.

Models can disclose patterns that are indifferent to mechanisms or specific causal trajectories. The Lotka–Volterra model represents the interdependent fluctuations in the levels of predators and prey in a region via a pair of differential equations that describe the dynamics of interdependence. It is not peculiar to foxes and rabbits, fish in the Adriatic, or even to biological species. It cuts across disciplinary divides, holding of the relations between cheetahs and gazelles, starfish and mollusks, predatory lenders and needy borrowers, asset strippers and their targets. Assuming certain formal constraints are satisfied, whenever one population preys on another, the model applies. To be sure, there are limits. If the predators drive their prey to extinction, the pattern obviously is broken; if the predators are themselves also prey, the dynamic is more complex; and so forth. The model discloses nothing about how the pattern is realized — that is, how specific population pairs modulate their size. It discloses a regularity that holds at a level of abstraction that is indifferent to precise reproductive mechanisms. In effect, it construes species-specific reproductive mechanisms as non-difference-makers. Again, prescinding from details is revelatory.

Strevens maintains that the Lotka–Volterra model is merely a schema. Because it says nothing about mechanisms — about how the model is realized in a specific predator-prey pair — it is incomplete (2008, pp.158–9). I disagree. Biologists might like to understand the mechanisms by which different pairs modulate their populations. But the fact that the same pattern holds regardless of differences in mechanism reveals something important about the dynamics of predator-prey relations — something that would be obscured if we scrapped the model in favor of a multiplicity of independent, finer-grained accounts. An accurate, detailed representation of the modulation in populations of particular predator/prey pairs would show that reproduction rates vary in tandem. And an accurate account of the mechanisms by which different pairs modulate their population size would explain how each pair does so.

But because the details would vary, such an account would occlude the fact that the modulations in population size of the various predator/prey pairs display the same pattern. It would obscure the fact that the difference in the ways the pairs modulate their population size makes no difference (see Dennett, 1991).

A model that is effective for one purpose may be ineffective for another. If we seek to display a pattern that holds for a variety of predator-prey population pairs, the Lotka–Volterra model may be appropriate. If we seek to display the means by which a particular pair modulates population size, we need a model that supplies more detailed information about mechanisms. Depth of understanding is in tension with breadth of understanding. The proliferation of details that are peculiar to one realization of the model obscures the pattern that the different realizations share. To grasp the general pattern in the population dynamics of predators and prey, we need to prescind from the details.

Some models augment. Maxwell’s model represents the electromagnetic field as composed of rotating vortices separated by so-called idle wheels. Via the introduction of fictional elements, the model shows how the relational structure of the production and transmission of electric and magnetic forces in an electromagnetic medium parallels the relational structure of the production and transmission of mechanical forces in a mechanical medium (Nersessian 2008, p.29 ff.). This enables scientists to export aspects of their understanding of mechanical media to the electromagnetic realm. To read all features of the model back into the target would be a mistake. Maxwell was by no means suggesting that the electromagnetic field was littered with tiny, idle, electromagnetic wheels. They are fictional devices introduced to highlight a more abstract structure, neither mechanical nor electromagnetic, that the two systems share.

Some models exaggerate features in order to make them salient. According to Kepler’s first law, the Earth travels around the sun in an elliptical orbit with the sun at one focus. Diagrammatic models of the law typically represent the major axis as considerably longer than the minor axis. If the models purported to show the actual shape of the elliptical orbit, they would be badly mistaken. In fact, the two axes are almost equal in length. But interpreted correctly, the models exemplify only the property of *being elliptical*, not the precise shape of the ellipse. Under that interpretation the models are correct. Models are symbols; they require interpretation and are vulnerable to misinterpretation.

These examples should be sufficient to show that effective models are felicitously false. Still, their openness to distortion, augmentation, exaggeration, and (often) extreme simplification raises the question: How do such divergences from accuracy advance, rather than inhibit, understanding?

3. DDI

Models are representations designed to foster understanding by facilitating fruitful reasoning that illuminates the phenomena they concern. The liberties they take, the divergences from overall accuracy, are justified by their epistemic payoffs. A number of philosophers of science have emphasized that models are things we think with (see Suárez 2009); they are not mirrors. R.I.G. Hughes (2009) characterized a model as a complex symbol that performs three interanimating functions: denotation, demonstration, and interpretation. His discussion is suggestive, but schematic. I will elaborate it to bring out features he sketched.

Denotation is the relation of the model to whatever it is a model *of*. The harmonic oscillator, being a model *of* a spring, denotes a spring; the tinker-toy model, being a model *of* a protein, denotes a protein. Demonstration consists in reasoning with the model according, as Hughes says, to “its own internal dynamic”. Interpretation consists in imputing the fruits of that reasoning to the target. Denotation is familiar and straightforward. It is the relation of a name to its bearer, and of a predicate to the items in its extension. Demonstration and interpretation require explication.

Demonstration: Models are subject to conditions of epistemic adequacy, which constrain and channel reasoning in a Hughesian demonstration. A model has an epistemic end — to answer a particular question or enhance understanding of a particular range of phenomena.⁸ Its adequacy depends on how well it serves that end — that is, how well it facilitates informative, fruitful, non-trivial inferences about its target, while impeding misleading and idle inferences. This is plainly a pragmatic matter. As Knuuttila argues, a model is a tool (2011). Like any other tool, it is to be assessed on the basis of how well it enables users to do the job they want it to do.

A model’s internal dynamic delimits the cognitive resources it is permissible to deploy and the uses to which they can permissibly be put. Those resources include background assumptions, auxiliary hypotheses, forms of inference, taxonomies, standards of relevance and of precision, and so forth. The recognition that the model is supposed to afford epistemic access to the target and answer specific questions about the target guides the choice of constraints. Descriptions, inferences, and actions that take us too far afield are sidelined.

The notion of inference operative here should be understood broadly. In addition to more rigorous forms of logical inference, a model’s internal dynamic may (but need not) license analogical reasoning, associative reasoning, and/or abductive reasoning. It issues more focused licenses as well. The internal dynamic of the water-droplet formation model discussed above permits treating the droplet near the breaking point as a vertical line — that is, it licenses reasoning as though, near the breaking point, the droplet were two-dimensional. Moreover, reasoning according to an internal dynamic involves action as well as deliberation. Using a Newlyn–Phillips machine to

represent or to figure out the effects of tweaking economic policy requires physically manipulating a flow of water; for it is by seeing how the water flows through the apparatus that one draws conclusions about the flow of money in an economy. Nor are practical inferences solely the province of material models. The internal dynamic of a purely abstract model or of a computer simulation licenses certain actions when particular results are reached. One such action is terminating demonstration — ceasing to draw further inferences. The internal dynamic tells us when to stop. We can think of a model's internal dynamic of as specifying the range of permissions and prohibitions for reasoning with it.

Chains of inference are in principle endless. There is always a further conclusion that could be drawn. They proliferate in all directions. To properly use a model, we need to know what direction to take in drawing inferences and when to stop. Unrestricted inferences threaten to generate a plethora of disparate conclusions, with no obvious way to tell which ones can be legitimately imputed to the target. It follows from $pV = nRT$ that the item in question is not a giraffe; it follows that the gas model is applicable in New York. Such inferences, although sound, are idle. The proper use of the model sidelines them; it takes them off-line. A model has irrelevant features — features that ought to be ignored. If the demonstration phase promoted drawing valid inferences indiscriminately, these would swamp (and likely deflect) our thinking. The model must block irrelevant and unproductive inferences. How does it do so?

The answer lies in the way a model symbolizes. Exemplification is the mode of reference by which a symbol refers to some of its own features (Goodman 1968, Elgin 1996, Vermeulen et al. 2009). It highlights them, underscores them, makes them manifest. Exemplification is the relation of a sample to whatever it is a sample of. Models are exemplars. Like paint samples, they are designed to make some of their features salient. Those features may be monadic or polyadic, static or dynamic, abstract or concrete. By representing a population as infinite, for example, the Hardy–Weinberg model showcases the aspects of allele redistribution that are insensitive to random fluctuations. Exemplification is selective. To highlight some features, an exemplar marginalizes or occludes others.

The inferences licensed by a model's internal dynamic are vehicles of exemplification. They show how, for example, changes in one parameter affect changes in others, how a system evolves over time, how robust or fragile linkages are. They disclose patterns and discrepancies that might otherwise to be hard to discern. The model does not exemplify the results of irrelevant inferences; its internal dynamic does not license them. So even when they are logically impeccable, they are idle. By being an exemplar then, the model constrains and directs its internal dynamic toward features that can be responsibly imputed to the target.

Interpretation involves identifying the features exemplified in the model's demonstration phase, and imputing them to the target. Hughesian interpretation is not lit-

eral denotation. We know perfectly well, for example, that gas molecules are not spherical. So in imputing sphericity to the molecules in the target gas, — in interpreting the actual gas molecules as spherical — we do not maintain that they really are spherical. Rather we construe actual gas molecules as, in effect, spheres with distortions. In general, in imputing features of a model to a target, we construe the target as having the features exemplified by the model, distended, distorted, or overlaid by confounding features. Then we ignore the confounds as, in the circumstances, irrelevant.

A model is designed to make particular features of its target salient. Its effectiveness depends on whether the features it exemplifies illuminate the target, enabling model users to understand the phenomena it bears on.⁹ By exemplifying a feature, a model affords epistemic access to it, enabling us to recognize it and appreciate its significance. The model can thereby equip us to see the target in a new light. $pV = nRT$ exemplifies the relation between temperature, pressure and volume. It omits any mention of attractive force. If the results of our calculations hold up when we impute them to the target, we have reason to think that intermolecular forces play no significant role in the thermodynamics of the system we are investigating. We know, of course, that every material object attracts every other. So we do not conclude from the effectiveness of the model that there is no attraction. Rather, we conclude that for the sort of understanding we seek, at the level of precision that concerns us, intermolecular attraction is negligible. It is not a difference-maker. To represent gas molecules as spherical does no harm. Indeed, it helps. For by treating the molecules as spheres we prescind from complications that would impede our understanding of the relations between pressure, temperature, and volume of a gas. This suggests that it is fruitful to think of the target in terms of the features the model exemplifies. It invites us to think of actual gases as ideal gases with distortions, of springs as harmonic oscillators with friction as a confounding feature, of automobile buyers as rational economic agents with (perhaps irrational but anyway irrelevant) quirks, and so forth.

Because models omit, distend, distort, and emend, they are context- and purpose-relative. An inaccuracy that is illuminating in one context or for one purpose may be misleading in or for another. A psychologist interested in why teenage boys like flashy cars would not represent her subjects as rational economic agents. Such a model would elide the very features that she sought to study. To devise an appropriate model requires recognizing what factors are and what factors are not difference-makers for the question one is investigating. To use a model correctly requires understanding how it functions — what phenomena it denotes, what range of features it has the capacity to exemplify, what sorts of inferences it supports and what sorts it blocks, what assumptions it makes, what scaffolding it relies on. The very same phenomenon can be modeled in mutually inconsistent ways, each of which is appropriate for a

different range of problems. The nucleus of an atom can, for example, be modeled as a liquid drop or as a rigid shell. Each model exemplifies some features of the nucleus. Each facilitates some inferences and blocks others. The question for the user is which, if either, suits her current epistemic purposes.

It might seem that the best strategy would be to always use the most accurate model we can devise. Even if we are forced, for the sorts of reasons Teller adduces, to deviate from wholly accurate representations, we might think that we should deviate as little as possible. This is not as obvious as it might look. Models are epistemically effective precisely because they sideline factors that make no difference. As Knuuttila says, “There is, from a cognitive point of view, something wrong in the idea that scientific representation should aim for as accurate a representation as possible” (2011, p.267). To incorporate as many factors as we can, and then explain that they do not matter, is to invite confusion. To see this, compare gas models. Boyle’s gas model is

$$pV = nRT$$

It is highly idealized — far too idealized for some purposes. But where pressure is low, so that intermolecular forces are slight, it is often appropriate. Over time, refinements were introduced that de-idealize some of its assumptions. The van der Waals model, for example,

$$k = n^2a/V^2$$

takes intermolecular forces into account. Further refinements in the equations of state yield increasingly accurate representations. Eventually we arrive at the virial equation

$$pV/NkT = 1 + B/V + C/V^2 + D/V^3 + E/V^4 + \dots$$

“which can be rendered arbitrarily precise by extending the equation indefinitely, with each added term being derivable from increasingly detailed assumptions about the intermolecular forces” (Doyle et al. 2019, p.5).

Why then do we keep the less accurate models in our cognitive toolbox? Why didn’t the virial equation supplant Boyle’s, Charles’s, and van der Waal’s formulations, consigning them to the dust bin of the history of science? Sometimes the virial equation supplies too much information. There are cases where all we need in order to understand the gas dynamical situation is $pV = nRT$. The added information the virial equation provides is, in such contexts, inert. Still, it might seem, we could in principle accommodate this by adjusting our focus. In some contexts, only a few terms of the equation exemplify. In others, more do. So we could restrict ourselves to a single model, but supply different interpretations of it in different contexts.

Such a strategy threatens to mislead. Grice’s second maxim of quantity is “Do not make your contribution more informative than is required” for the purposes at hand

(Grice 1989, p.26). When a speaker or writer imparts information, she implicates that it is relevant. In imparting the information, she makes it salient, giving her audience reason to think that in the current context it is worthy of attention. If in fact it is not worthy of attention, the implicature is unwarranted. Adducing the virial equation with its full power threatens to mislead, even if the adducer insists that we need only attend to the first few terms. Her audience is likely to wonder why the further terms were even mentioned if they aren't doing any work. The more streamlined models are preferable when appropriate, because they omit the irrelevancies.

Grice's maxim directly pertains to communication. I suggest, however, that the reason it is valuable for communication is that it reflects something important about understanding. By eliminating or marginalizing irrelevancies, we are better able to grasp relations among relevant factors. We understand better when we respect the maxim; we do not just communicate better. The reason Grice's maxim is a conversational maxim is that it is an epistemic maxim.¹⁰

Streamlining in modeling is not just due to an aesthetic preference for desert landscapes. The omission of irrelevancies figures in a model's capacity to advance understanding of its target. The message is not: it is permissible to omit these (irrelevant) factors. It is: omitting these factors promotes the grasp of something about the phenomena that otherwise we would not, or not easily appreciate.

Models figure in the understanding of a range of phenomena when it is epistemically fruitful to represent the phenomena as if they had the features the model imputes to them, where something is epistemically fruitful only if it either fosters or challenges the integration of the behavior of the phenomena into our evolving understanding of the world.

Every object has indefinitely many properties and stands in indefinitely many relations to other things. The vast majority of these are of no interest whatsoever. Some of the interesting and important ones are neatly labeled by our literal vocabulary. They can be directly and literally represented. Others are semantically unmarked. Prior to Newton, for example, the property shared by a falling apple, the varying heights of the tides, and the moon's staying in orbit was unmarked. If we want to recognize such properties and relations, we need to indicate them indirectly. One way to do so is by characterizing the objects that display them as-if-ishly (see Vaihinger 2009). It is as if gas molecules were spheres, or as if predators were insatiable, or as if the moon were falling toward the earth. Such as-if-ish representations can be epistemically important. The reason is not just that you won't go wrong in a particular context if you think of molecules as spherical or of the electromagnetic field as composed of vortices separated by idle wheels or the moon as falling; it is that *the fact that you won't go wrong* discloses something important about the phenomenon. The effectiveness of the model discloses that a particular aspect of the molecule's geometry — its being somewhat spherical — is significant. The model then provides emphasis and

focus. It affords insight not only into what property the molecule has, but which of its properties are worth registering if we want to integrate the behavior of the molecule into our understanding of gas dynamics.

I began by saying that modeling is a powerful epistemic tool. The power lies in its ability to simultaneously generate representations that afford focus, and show why that focus (even if provided only as-if-ishly) is valuable. In effect models not only say, “This is what you should be looking at”, they also say, “This is why you should be looking at it and ignoring factors that interfere with looking it”.

Because modeling involves Hughesian demonstration — drawing inferences and perhaps performing other actions that the representation facilitates — a science that relies on models cannot be wholly spectatorial. We cannot passively look *through* the model to see the world, as we might look through a transparent pane of glass. Nor can we look *at* the model, expecting it, like a flat mirror, to reflect the phenomena without distortion. Models are epistemic tools; they enable us to grasp phenomena. Grasping is not looking. To grasp something is to be able to manipulate it to serve one’s ends (Hills, 2016). In scientific models, as Knuuttila (2011) and Suárez (2009) emphasize, manipulating involves inferring — reasoning with the model, taking advantage of affordances it provides. This does not preclude taking a realist stance if that means only holding that the items mature science is committed to — be they entities, structures, mechanisms or whatever — really exist. What it does preclude is taking the view that scientific success affords direct, aperspectival access to the exact way the mind-independent world is. Rather than enabling us to see how things are in themselves, models enable us to apprehend how things appear from a certain orientation — an orientation whose justification lies in its affording us the sort of access we need to answer the questions we want to ask, and solve the problems we seek to solve.

I have argued that scientific understanding cannot, even in the long run, deliver a precise, accurate view from nowhere. Rather than being spectatorial, science is agential. Nature is complex. It displays a plethora of patterns. Some of these are, in their natural habitats, masked by irrelevancies. They emerge only under idealizations where inconsequential complications are set aside. Scientific understanding involves idealized models. They are effective to the extent that the features they exemplify and impute to their targets make a difference to the characteristics and/or behavior of the phenomena we seek to understand. There is no benefit in imputing non-difference makers to targets. Indeed, such imputations can be epistemically deleterious. Too much information can impede understanding by obscuring patterns and features that streamlined models disclose. So we have reason to retain and deploy more schematic models even when more detailed models are available.

References

- Batterman, R. 2009. Idealization and Modeling. *Synthese* **169**: 427–46.
- Chang, H. 2004. *Inventing Temperature*. Oxford: Oxford University Press.
- Dennett, D. 1991. Real Patterns. *Journal of Philosophy* **88**: 27–51.
- Doyle, Y.; Egan, S.; Graham, N.; Khalifa, K. 2019. Non-Factive Understanding: A Statement and a Defense. *Journal for General Philosophy of Science* **50**: 345–65.
- Elgin, C. 1996. *Considered Judgment*. Princeton: Princeton University Press.
- Elgin, C. 2017. *True Enough*. Cambridge, MA: MIT Press.
- Goodman, N. 1968. *Languages of Art*. Indianapolis: Hackett Publishing Co.
- Hacking, I. 1988. The Stability of the Laboratory Sciences. *Journal of Philosophy* **85**: 507–14.
- Hughes, R. I. G. 1996. Models and Representations. *PSA 1996*, vol. **2**: S325–36.
- Knuuttila, T. 2011. Modelling and Representing: An Artefactual Approach to Model-Based Representation. *Studies in History and Philosophy of Science* **42**: 262–71.
- Lange, M. 2020. *Because Without Cause*. Oxford: Oxford University Press.
- Nawer, T. 2021. Veritism Refuted? Understanding, Idealization, and the Facts. *Synthese* **198**(5): 4295–4313.
- Nersessian, N. 2008. *Creating Scientific Concepts*. Cambridge, MA: MIT Press.
- Potochnik, A. 2017. *Idealization and the Aims of Science*. Chicago: University of Chicago Press.
- Strevens, M. 2008. *Depth*. Cambridge MA: Harvard University Press.
- Strevens, M. 2017. How Idealizations Provide Understanding. In: S. R. Grimm; C. Baumberger; S. Ammon (eds.) *Explaining Understanding*, pp.37–50. New York: Routledge.
- Suárez, M. 2009. Scientific Fictions as Rules of Inference. In: M. Suárez (ed.) *Fictions in Science: Philosophical Essays on Modeling and Idealization*, pp.158–78. New York: Routledge.
- Teller, P. 2018. Measurement Accuracy Realism. In: I. F. Peschard; B. C. van Fraassen (eds.) *The Experimental Side of Modeling*, pp.273–98. Minneapolis: University of Minnesota Press.
- Vaihinger, H. 2009. *The Philosophy of 'As If'*. Mansfield Center CT: Martino Publishing.
- Vermeulen, I.; Brun, G.; Baumberger, C. 2009. Five Ways of (not) Defining Exemplification. In: G. Ernst; J. Steinbrenner; O. Scholz (eds.) *From Logic to Art*, pp.293–319. Frankfurt: Ontos.

Notes

¹Snell's law holds only for isotropic media. Since most media are anisotropic, it is wrong more than 50% of the time. But many media are nearly isotropic, so it is often not off by all that much. Moreover, the media scientists care about are mostly isotropic, so for their purposes the other cases don't matter.

²Teller discusses this case in an early draft of his (2018). To my regret, he did not include it in the published version. The quotation is from his 2016 manuscript.

³Leibniz invented calculus to solve metaphysical problems. That too was an extension of our epistemic range, but one that does not concern me here.

⁴I am grateful to Otávio Bueno for articulating this worry.

⁵I am grateful to Anjan Chakravartty for help articulating this point.

⁶Strevens (2008, pp.158–9) maintains that all difference-makers in the natural sciences are causal. Although models may ‘black box’ details about actual causes, they are committed to its being the case that whatever is in the black box makes a difference to the causal explanation of the phenomena in question. I disagree. I believe that there are non-causal patterns, regularities, and relations of dependence, independence, and interdependence in nature. Science looks for them too. See Lange (2020).

⁷Here I will register an objection to construing models as approximations. A very large finite number is not approximately infinite. It is as far from infinity as a small finite number is. So construing these models as approximations strictly makes no sense.

⁸It may have other ends — perhaps practical ones — as well. Here I am concerned only with its epistemic ends.

⁹The illumination may be indirect. The model may exemplify something that exemplifies something else that ... bears on the target phenomena. Or the model may have no target. A biological model of how a species with five sexes would redistribute its alleles is one. Such a model is effective if the features it exemplifies shed indirect light on actual allele distribution — for example, if the features it exemplifies makes salient something subtle about how alleles redistribute or perhaps fail to redistribute in a species with fewer sexes. We may discover something important about the actual case by considering a suitably constructed counterfactual case.

¹⁰I am grateful to Tamer Nawer for encouraging me to clarify this point. See Nawer 2021.