Synthese
https://doi.org/10.1007/s11229-018-1721-z

S.I. : ABSTRACTION AND IDEALIZATION IN SCIENTIFIC MODELLING

Idealization and abstraction: refining the distinction

Arnon Levy[1]

Received: 30 August 2016 / Accepted: 7 February 2018

Abstract Idealization and abstraction are central concepts in the philosophy of science and in science itself. My goal in this paper is suggest an account of these concepts, building on and refining an existing view due to Jones (in: Jones MR, Cartwright N (eds) Idealization XII: correcting the model. Idealization and abstraction in the sciences, vol 86. Rodopi, Amsterdam, pp 173–217, 2005) and Godfrey-Smith (in: Barberousse A, Morange M, Pradeu T (eds) Mapping the future of biology: evolving concepts and theories. Springer, Berlin, 2009). On this line of thought, abstraction—which I call, for reasons to be explained, abstractness—involves the omission of detail, whereas idealization consists in a deliberate mismatch between a description (or a model) and the world. I will suggest that while the core idea underlying these authors' view is correct, they make several assumptions and stipulations that are best avoided. For one thing, they tie abstractness too close to truth. For another, they do not allow sufficient room to the difference between idealization and error. Taking these points into account leads to a refined account of the distinction, in which abstractness is seen in terms of relative richness of detail, and idealization is seen as closely connected with the knowledge and intentions of idealizers. I lay out these accounts in turn, and then discuss the relationship between the two concepts, and several other upshots of the present way of construing the distinction.

Keywords Idealization · Abstraction · Omission · Deliberate misrepresentation

✉ Arnon Levy
arnon.levy@mail.huji.ac.il

[1] Philosophy Department, The Hebrew University of Jerusalem, Mt. Scopus, 91905 Jerusalem, Israel

Published online: 16 February 2018

1 Introduction

As the title suggests, this paper discusses idealization and abstraction. It is largely a suggestion about how we should understand these notions, which have become increasingly central in philosophy of science. My discussion draws on two existing accounts—the only extended, fully explicit accounts I am familiar with—one by Martin Jones (2005), the other by Peter Godfrey-Smith (2009).[1] The key idea in both is that abstraction concerns a description's degree of detail whereas idealization consists in introducing simplifying misrepresentations. But there is, I think, room for improving upon these accounts and refining them. The result is greater clarity about idealization and abstraction. This, as I illustrate in the latter part of the paper, can benefit a number of important discussions.

I begin with a few preliminary remarks to clarify the aims and contours of the discussion, starting with a methodological point. The sort of task I am engaged in, explicating significant notions and distinctions, requires tracing a fine line. On the one hand, the idea is to take existing usage, both scientific and philosophical, as a starting point. On the other hand, there are variations, at times inconsistencies, in how philosophers and scientists understand the notions of abstraction and idealization, and in how they use the corresponding terms. While I am confident that my proposal is similar to, and continuous with, some extant understandings of idealization and abstraction (examples are given below), I am also sure that it diverges from others. Therefore, my goal isn't, and cannot be, to capture the common usage of 'idealization" and 'abstraction'. Instead, I will attempt to characterize what seems to me the most effective way of understanding these notions.

Philosophers, and to some extent scientists too, rely on the notions of idealization and abstraction in order to formulate ideas and argue for theses about central scientific practices such as modeling and explanation. In these kinds of discussions idealization and abstraction are central, but they are not, typically, the analytically interesting categories in and of themselves. The discussion relies on them to address other issues—the nature of explanation, reduction and emergence, realism. To this end, then, it is best have a characterization of idealization and abstraction that remains as neutral as possible between the substantive philosophical issues at stake. That is the sort of explication I will attempt to offer. Indeed, I take this to be a central argument for my proposal: by regimenting 'abstraction' and 'idealization' in the way I suggest, we can clarify various issues in the philosophy of science, and unify our discussion of them. (This is elaborated upon in Sect. 4.)

Another assumption I will be making that idealization and abstraction are attributes of representations. But I will use 'representation', throughout, in a low-key way—i.e. with no specific theory of representation in mind and without, I think, making controversial assumptions. Texts, mathematical models, graphs and other forms of visual depiction, as well as physical models—these all count as representations for present purposes (and in general, I think). They can all, in principle, contain idealization and/or

[1] Others have expressed similar views—e.g. Cartwright (1999), Elliot-Greaves and Weisberg (2014), Frigg (2006), Nowak (1992), Strevens (2008) and Weisberg (2013). But since these authors do not provide extended, argued-for accounts of the distinction itself, I won't discuss their work in detail.

abstraction. There may be differences in how the notions of abstraction and idealization apply to different items in this family—what counts as more/less detail may differ between graphs and mathematical models, for instance. But these differences have more to do with questions about the nature of representation and related notions such as depiction, denotation and reference, a set of topics I will not address here.

A final preliminary concerns a different use of 'abstract' and its cognates. This term functions in metaphysical discussions to designate a kind of entity—an abstractum (Rosen 2012). Abstracta are (by most lights) devoid of causal powers and do not have properties such as spatiotemporal location, shape, size and color. I note this to highlight that *this is not* the sense of 'abstract' I will be speaking of. My discussion is not metaphysical—this isn't a paper about abstract entities.

The plan going forward is simple: the next Sect. 2 characterizes abstraction and a related notion—abstractness. The following Sect. 3 characterizes idealization; and the penultimate Sect. 4 looks at several upshots of these characterizations. Section 5 provides a short summary.

2 Abstraction

Jones and Godfrey-Smith, as noted above, understand abstraction as incompleteness. An abstract description says less about its subject matter than there is to say. More specifically, they have in mind the provision of *an incomplete truth*. "An abstract description of a system leaves a lot out," says Godfrey-Smith, "[b]ut it is not intended to say things that are literally false." (2009, p. 48). Jones puts the point more bluntly (emphasis in the original): "[O]n the proposal I am putting forward… abstractions involve omission *without* misrepresentation." (2005, p. 175). I think the basic idea conveyed in these quotes is correct and important: Paucity of detail often seems to be what references to 'abstract' and 'abstraction' designate. We say that the Lotka–Volterra model 'abstracts away' from many details about predators and prey—it does not provide information about feeding strategies or animals' spatiotemporal distribution, not to mention the critters' colors, sizes or other distinguishing features. Similarly, network diagrams such as those that are used to model the interconnections in our brains or the regulatory structure of genomes, are typically regarded as abstract because they merely depict the causal connections between elements (neurons, genes), saying little about what the elements look like and how interconnections work (Levy and Bechtel 2013).

While I accept the basic idea of abstraction as lack of detail, I think it can be improved upon in two significant ways. First, we should draw a clear process/product distinction. Second, it is best not to tie the notion of abstraction to truth—a representation can be *both false and abstract*. Let me discuss these points in turn. This will lead up to a more explicit account that refines the insights of Jones and Godfrey-Smith.

2.1 Process versus product

In scientific and philosophical discourse, 'abstraction' sometimes denotes a standing feature of a given representation—poverty of detail. On the other hand, talk of 'leaving out' or 'omitting' detail brings to mind a process, where one starts with a representation that contains certain details and ends up with one that lacks them. Such a process can take diverse forms: there could be a literal omission of a specific piece of information from a description that formerly contained it, such as leaving out the precise value of a variable (and, say, providing a range of possible values). But abstraction may also consist in applying certain formal operations—like reducing the dimensionality of a problem, or calculating an average.

There may be philosophically interesting aspects to the process of abstraction. But most philosophical discussions involving abstraction appear to aim at abstraction-qua-product, and so I will spend most of my time on that. This focus is also more general, in one way, because of an asymmetry between process and product in this context. By definition, the process of abstraction results in an abstract product. So the process can be understood in terms of the product—what makes it a process of abstraction is the end product, a detail poor representation. But not the other way around: an abstract representation need not be the product of a process of abstraction. For instance, the statement "The speed of light in a vacuum is constant" is less detailed, hence more abstract, than "the speed of light in a vacuum is c = 299,792,458 m/s." This is so irrespective of whether the former was arrived at by leaving out information contained in the latter, or via some other route—e.g., by derivation from general premises. More generally, one may identify a model, description or other representation as abstract without worrying about how it was created, and whether details that are now missing were once included.

In view of the potential for process\product ambiguity, and more generally for the sake of clarity, I will henceforth use slightly different terminology. 'Abstraction' will designate the process—moving to a detail-poor representation. When speaking of products (representations) I will speak in terms of 'abstractness'. More precisely, 'abstractness' will designate a property of the representation in question—its level of detail.

2.2 Abstraction and truth

Godfrey-Smith, and more explicitly Jones, speak of abstractness as involving a lack of *true* detail—as they see it, a representation cannot be both abstract and idealized.[2] But consider, to continue with the simple example used above, the statement "the speed of light in a vacuum is several hundred meters per second." It is more abstract, i.e. less detailed, than the statement "the speed of light in a vacuum is 300 m/s." But both statements are false. More generally, there is no in-principle reason to tie abstractness to truth in this way: a given representation can be both abstract and idealized with respect to one and the same (feature of one and the same) thing. An account that connects abstractness and truth too closely has trouble allowing for this common situation.

[2] Jones: "[A] given representation can contain an idealization, or an abstraction, or neither, but it cannot contain both." (2005, p. 176).

More fundamentally, this observation reveals that abstractness is essentially a comparative notion: the issue isn't whether some aspect of the world is included in the representation per se. It is whether, given two representations, one includes more detail than another vis-à-vis the same subject matter. Put differently, abstractness is not about fidelity to reality, but about *relative informativeness*. Holding the subject matter fixed—be it a real-world system, or some idealized or imaginative relative of reality—we ask: which of the following descriptions of it is more detailed (i.e. less abstract).

Thus, I accept the basic idea that abstractness is connected to the detailedness of a representation. But I differ from Jones, Godfrey-Smith (and, I think, quite a few others) in two ways: I give priority to product over process, and I regard the notion as comparative. My suggestion is that the more basic notion isn't abstraction per se but rather *relative* abstract*ness*.[3]

2.3 A three-place relation

Abstractness-as-incompleteness—which I am treating here as a (partial) foil for my view—is a two-place relation between a description and what it is about. In contrast, relative abstractness is a three-place relation between two descriptions and some portion of the world they are (both) about.

Informally, one representation has greater abstractness—it is more abstract—relative to another representation if it contains less detail with regard to the same subject matter or topic. Thus the statement "this is a mammal" rates higher, in terms of abstractness, than "this is a Red-tailed Chipmunk"; a model that notes an object's momentum is more abstract than one that specifies its mass and velocity; and likewise for a report that lists the height of each and every tree in some patch of forest over one stating the average tree height and its standard deviation.

Let me emphasize that on the present suggestion abstractness pertains to two (or more) representations of *one and the same subject matter;* or (speaking loosely) of one and the same thing.[4] It seems intuitively hard to make judgments about relative abstractness with respect to statements that are about entirely different topics: Is the second law of thermodynamics more or less abstract than the principle of evolution by natural selection? Sometimes, perhaps, we can guesstimate the answer to such questions. But strictly speaking, I think, the idea that a concrete representation fleshes out, or fills in the details of, an abstract counterpart cannot be applied when the representations concern different subject matters. Moreover, this way of looking at abstraction

[3] Godfrey-Smith does note that "ignoring some features in a description of a system is inevitable to some extent in any description. The question is only how much is left out, and what is retained." (p. 48) but does not construe the notion of abstraction in the comparative way I have. Jones devotes more space to degrees of abstraction (2005, §4), but explicitly regards that as derivative from the idea of abstraction as omission of detail.

[4] This requirement is explicit in Michael Strevens' view of abstraction (2008, Ch. 3). A related claim is Cartwright's requirement that an "satisfying the associated concrete description that applies on a particular occasion is what satisfying the abstract description consists in on that occasion" (1999, p. 39). For it follows from Cartwright's claim that the abstract and concrete description have the same truth maker. On a natural understanding of subject matters, they should thereby have the same subject matter.

allows us to sharply distinguish abstractness from generality—a point to which I'll come back towards the end.

Is it possible to give a more rigorous treatment of relative abstractness? Perhaps in some contexts. Michael Strevens (2008), for instance, suggests that abstractness should be understood in terms of logical entailment, roughly as follows: representation A is more abstract than representation B iff (i) B's subject matter is contained within A's subject matter. And (ii) all propositions in A are entailed by propositions in B.[5] This covers an interesting range of cases, but it is doubtful that all instances of abstractness can be handled this way. For instance, it seems that non-indicative statements can exhibit relative abstractness. Moreover, non-propositional representation, like concepts and notions, perhaps even pictures, can be abstract. In these cases, entailment is beside the point. Perhaps Strevens' account can be modified. Alternatively, a different way of characterizing abstractness could be pursued. For example, we may try to give an account in terms of some measure of information content—where, very roughly, abstractness would be understood as relative paucity of information, vis-à-vis the same subject matter.[6] Whatever technical analysis is given, it seems that statements can be compared in terms of their degree of abstractness. If there is a useful formal analysis of abstractness, therefore, it should allow for some kind of ranking of representations.

While I have suggested that abstractness is the more basic notion, I do not mean to imply by this that judgments about abstractness are made against a background of two *pre-specified* representations, one of which we judge to be more abstract than the other. That can but need not be the case. Instead, I think that in making judgments of abstractness with regards to a given representation we typically only presuppose that there *could be* a representation with a different degree of detail. In other words, we often rightly speak of a representation as abstract even though we are not comparing it to some specific, more concrete alternative. In so doing we simply express the judgment that a more detailed description is possible, i.e. that more *could* be said. And since a strictly complete description is rarely if ever feasible, it will virtually always be the case that a given representation is more abstract than *some* possible alternative. That said, a judgment about abstractness will, naturally, be more interesting and informative if it specifies, or at least suggests, what a less abstract representation would look like.

To summarize the discussion so far: I suggest that abstraction—or rather, abstractness—is best understood as a three-place comparative notion, namely which of two (possible) representations of some subject matter is less detailed.

[5] Two comments: First, we could weaken (i), having it state that A's subject matter is *either identical to, or contained within*, B's subject matter. In that case we'll have defined "weak abstractness", i.e. a definition of when A is at least as abstract as B. Second, Strevens speaks in terms of causal models and requires that "all causal influences described by [A] are also described by [B]" (Ibid, 97). My account is not restricted to causal representations.

[6] I should note that here, and throughout, I am understanding the information (and/or detail) contained in a representation in objective terms. Or at least, in terms that do not pertain to any individual's state of belief, knowledge etc. If two representations differ in their degree of detail than (all else equal) that should be so irrespective of who produces or consumes the representations.

3 Idealization

Let me turn now to idealization, which I understand in terms of deliberate misrepresentation. Again, my account will track the basic ideas of Jones (2005) and Godfrey-Smith (2009), as well as suggestions from a number of other philosophers (e.g. McMullin 1985; Laymon 1995; Weisberg 2007; Strevens 2008, Ch. 8). A description is idealized inasmuch as what it says is *known not to be true* of its intended target. A model of a brick sliding down a frictionless inclined plane is idealized, since everybody—at least, everybody with basic knowledge of mechanics—knows that no actual (earth-bound) body is exempt from friction. An account of gene flow in a population that assumes an infinite population size is idealized, in that, obviously, no real-world population is infinite. Similarly for a treatment of a brain function in terms of a three-layered network or a model of a marketplace as populated by fully rational agents. I take idealization simpliciter to involve no more (and no less) than these quick examples illustrate: the representation is deliberately false, i.e. it misrepresents the target, and this is known to be so by competent practitioners in the relevant area.

Note that some authors tie idealization to more specific kinds of misrepresentations. For instance, Strevens (2008, Ch. 8) focuses on idealizations that involve extreme or 'default' values—an assignment of zero, or infinity, or some such value. I agree that this is common—the examples I just gave attest to that—but I do not see it as constitutive of idealization or even as an especially important aspect of it. There are many types of idealizations that do not involve "extremities", such as assuming that one's target is symmetric in some respect (e.g. treating particles as spheres); assumptions about the separability of various factors or components, and depicting a continuous magnitude as discrete (or vice versa.)

It might be thought that while there is no specific connection between idealization and extreme values, there is a more general feature of which this is an instance—-namely, simplification. Godfrey-Smith, for instance, says that "idealization involves a departure from reality in the direction of some kind of simplicity…" (2009, p. 49. See also Elliot-Greaves and Weisberg 2014). There is certainly something to this, but it is important to be precise about what. It is not that an idealized description is *itself* simpler. There is nothing inherently simple, I think, about an infinite population relative to a finite one, or about a discrete variable relative to a continuous one. What such assumptions simplify is one's *handling of the relevant model*. The idealizations contained in a model are often geared towards allowing one to bring certain tools to bear—primarily formal tools. How to do so depends, of course, on the analytical tools at one's disposal, given the historical and scientific context. So the connection between idealization and simplicity, while it exists and may be important, is indirect and context dependent.

A further point is that idealization often occurs in the context of mathematical modeling. But on the present suggestion, mathematics isn't essential either. Mechanistic models in molecular biology, for instance, often portray proteins and other large molecules as having simpler structures or activities than they in fact do. This is an idealization, but it need not be part of a mathematical analysis.

3.1 Anti factivity

I now turn to contrasts between idealization and abstraction. There are several of these. Most basically, idealization is a matter of distortion rather than level of detail. But, in addition, idealization concerns the relation between a representation and *what the world is actually like*. Furthermore, idealization involves *deliberate* misrepresentation.

Let me take each point in turn. Recall, first, my suggestion about abstraction—it should be seen in terms of comparative level of detail. In contrast to this, idealization isn't primarily a matter of the relation *between* representations, but of how a representation relates to the world. Another way to put the point is to say that in idealization, an author (a modeler, say) introduces a known falsehood. Thus, unlike with abstraction, there is an intimate connection to truth. Or rather, to falsity. It might be said that idealization is "anti-factive"—that a description is an idealization implies that it is false.

For this reason, a given model is typically an idealization only relative to a *specific* target in the world—or, often, a specific kind of target, since models often cover a range of similar systems. So a given model may be an idealization with regards to one (kind of) target, and yet be an accurate description of another target. For instance, a population genetics model that deals with an asexually reproducing population may be accurate when applied to bacteria, but idealized if applied to a population of multicellular, sexually reproducing rabbits.

This is the point to note two potential complications. First, some authors view idealizations, idealized models in particular, in a way that precludes attributions of truth and falsity altogether. I have in mind advocates of indirect approaches to modeling (Giere 1988; Godfrey-Smith 2006; Weisberg 2007). On this approach the texts and mathematical equations that appear in scientific books and articles specify model systems, which are seen as objects or things of sorts. Model systems are then used as a means for representing the world—where that is done via some sort of mapping or model-world similarity. So on this view models are not, properly speaking, truth-apt and cannot be false (or true). Does this show my suggestion to be incorrect, or commits me to rejecting the indirect view?

Although I do have reservations about the indirect view, they are independent of the present issues (Levy 2015). For I take it that even on an indirect view there should be some way to express the idea that idealization is a matter of deliberately simplifying one's representations of the world. I have cashed this out in terms of falsehood, and I believe that is the best way to understand what is going on. Indeed, those who hold the indirect view, including the authors references above, often describe idealization in terms of falsehood or the introduction of distorting assumptions themselves. But a proponent of an indirect view could also say that a model is idealized to the extent that it exhibits deliberate (and, typically, significant) discrepancies with the world. Perhaps this formulation can be refined, but I take it that something of this sort can (indeed, must) be said within the indirect view. So, as far as I can see, all the claims I make here about idealization can be adapted to that way of understanding model-based representation.

Second, some authors have argued that, appropriately interpreted, many or all idealizations should not be regarded as false, or even, strictly speaking, as representing anything in world (truly or falsely). A recent case in point is Jones (2013; not the same Jones alluded to earlier), who distinguishes a model's *actual* and *apparent* content. The apparent content consists of the claims the model appears to be making, taken at face value. The actual content is more subtle, and requires taking into account the intended scope of the model and various other restrictions on its interpretation. Jones argues that while an idealization's apparent content *is* typically false, its actual content is *never* false. This is because the idealizations function as scope restrictors and do not form part of the actual content.

If the distinction between actual and apparent content is accepted, then my account of idealization is best understood as an account of an idealization's apparent content. That, I think, is also the best way to understand the accounts of Martin Jones, Godfrey-Smith and most other authors on the subject. My claim is that, taken at face value, idealizations are deliberate misrepresentations of some aspect of the world (in contrast to abstractions, of course, which are detail-poor representations). I think this claim, especially given the contrast with abstraction, is valuable even if, in the final account, we can extract actually-true content from apparently-false (i.e. idealized) models.

3.2 Intentions and the process–product distinction

The second feature of idealization I want to highlight concerns its relativity to the intentions and knowledge of a particular author or authors. In contrast to (Martin) Jones and Godfrey-Smith,[7] I understand idealization to involve a *deliberate* introduction of falsehood into a representation. If someone mistakenly thought that rabbits reproduce asexually then they wouldn't, in proposing a model that contained such an assumption, be idealizing. They would simply be mistaken. Admittedly, with some idealizations it is hard to see how anyone could propose them in earnest, as it were. It is hard to imagine a person who knows enough about population genetics (and rabbits) to do work in that area, but who would mistakenly think that there exist infinite populations, not to mention that there were infinitely many rabbits in our world. However, it is perfectly possible, indeed it has happened time and again in the history of science, that a model or theory is proposed at one time (as true), discovered to be false, and later gets treated as an idealization. Current use of Newtonian mechanics is arguably such a case, and similarly, for instance, for some applications of Mendelian genetics. These theories are nowadays known to substantially mis-describe gravitational interactions and inheritance, respectively. Still, in many contexts they are used in order to simplify various calculations or to isolate features of interest. At any rate, the key contrast here is between idealization and error, in that the latter involves non-deliberate misrepresentation.

These points are closely connected with another contrast between idealization and abstraction, concerning the distinction between process and product. In discussing abstraction, I proposed a strict separation of process and product: the abstractness

[7] Jones explicitly denies that idealizations need be intentional. Godfrey-Smith states that idealization consists of "treating things as having features they clearly do not have" (2009, p. 47), which may suggest an intentional element. But he does not emphasize the contrast with error, which I take to be important.

of a description has nothing to do with whether it was arrived at via abstracting, i.e. via stripping away detail. Not so for idealization. Indeed, in one important sense, in idealization the process is constitutive of the category. Suppose two persons offer models that make the exact same claims, claims which are at least partly incorrect, about one and the same system in the world. But suppose these identical models are offered in rather different spirits. One is proposed as a true theory, the other as a deliberate misrepresentation (intended, say, to simplify some calculation.) I think it is natural and helpful to treat the former as a case of error, and the latter as a case of idealization. More importantly, the difference matters insofar as it has consequences for what one does with the model, how one tests it, which aspects of it one hopes to confirm and so on. In this kind of situation, note, the supposed models are identical and they are directed at the same system in the world; the difference consists in the intentions of their authors. Intentions are one aspect of the process through which the two models came into being—or at least the manner in which they are handled. Therefore, it seems that whether a description is idealized depends, in part, on its circumstances of creation and use—about whether it is put forward as a simplifying distortion or as a sincere, if wrongheaded, claim about what the world is like. So process and product are more closely connected in the case of idealization than in the case of abstraction\abstractness.

3.3 Degrees of idealization

A further contrast between idealization and abstraction concerns the notion of *degree or level* of idealization. Recall that with respect to abstraction I argued that the more important concept was that of relative abstractness. The parallel claim does not hold for idealization—comparative judgements of relative idealized-ness are not, it seems to me, basic or important.

Indeed, it is not even clear what the notion of a degree of idealization comes to. On one way of understanding this notion, the extent to which a claim is idealized amounts to something like how far it strays from the truth. But what does "distance from the truth" mean? Maybe it is a form of approximate truth or verisimilitude, though given the state of the literature on that topic this isn't necessarily good news. On another way of understanding the notion, a model's degree of idealization is a function of the number of idealizing assumptions it makes. This way of construing degree of idealization assumes that the idealizing assumptions can be cleanly individuated and counted. But will this assumption generally hold? I agree with Jones that it does not: consider a model that "represents the gravitational force of the Earth on [a] cannonball as constant in both direction and magnitude throughout the region in which the cannonball moves, whereas in fact there will be variation in both respects. Is that one idealization, or two? Or an uncountably infinite number, one (or two) for each spatial point at which the model misrepresents the Earth's gravitational field? There would seem to be little prospect of settling upon a non-arbitrary answer to such questions…" (2005, p. 184).

Perhaps sometimes individuating idealizations and counting them can be carried out. For instance, the assumption that a population is infinitely large appears to be one idealization on any reasonable way of counting. Still I doubt that even in such cases counting idealizations will give us a significant and useful notion of degree of idealization. For one thing, one can readily point to cases where it gives apparently incorrect results. For instance, two important models in polymer physics are, first, the freely-jointed chain model (FJC), in which a polymer is envisioned as a chain of rigid monomers, whose orientations are independent of each other, like a chain of strung-together paper clips. And, second, the worm-like chain model (WLC), in which the polymer is conceived as a long homogenous rod, continuously flexible at every point along its length. Assuming one can count the constituent idealizations, it seems both models contain the same number of idealizations.[8] But the WLC model is usually treated as the more realistic (i.e. less highly idealized) of the two.

But the most important point is this: even if one could give a sound general characterization of degrees of idealization, I do not think much is to be gained by doing so. For the significance of an idealizing assumption consists in what it allows one to do—the computational or other techniques that can be brought to bear on the model, the kinds of features it allows one to highlight and so on. But there is no reason to suppose that the power or fruitfulness of the techniques and insights that an idealizing assumption affords are related to its degree of idealization in the intuitive sense (or in any other sense I am familiar with). A model that portrays a population as infinite would appear farther from the truth than one that portrays it as large but finite. But the former makes the model amenable to various mathematical techniques, and is therefore more valuable for many purposes. So I do not see a reason to place much weight on the notion of a degree of idealization, and I do not think there is much to lament if a defensible version of this notion is, as I suspect, hard to formulate.

Summing up the discussion of idealization: I propose to see it as involving deliberate misrepresentation. It is a non-comparative notion, but one that essentially pertains to the intentions of authors and users (of the idealized representation) and therefore exhibits a process–product dependency, unlike abstraction.

4 Implications and relationships

Having laid out the basic distinction, I now want to consider some implications of this way of looking at idealization and abstraction. My aim is not to explore the issues below in detail or reach definite conclusions, but to illustrate some of the benefits of working with a refined distinction of the sort I have suggested. As noted in the Introduction, this is essentially the argument I want to offer for proposed distinction: It is fruitful, illuminates nearby issues and allows us to pose problems in a clear way.

4.1 A representation can be both idealized and abstract

The first point I want to look at concerns the compatibility of idealization and abstraction (focusing, as before, on abstractness). Recall that both Jones and Godfrey-Smith presented abstractness as the omission of *true* information. They also took idealization

[8] Either one or two, depending on whether one counts flexibility as separate from the assumption that the polymer is a jointed chain/long rod.

to consist in the introduction of falsehood. One consequence of this is that a model cannot be both abstract and idealized, at least not with respect to one and the same fact or feature of its subject matter (Jones explicitly accepts this—see footnote 3, above). By my lights this is not so: While the presence of abstractness does not *imply* falsehood or misrepresentation, it is compatible with it. To return to our simple example, saying that the speed of light in a vacuum is somewhere in the range of 200,000–250,000 m/s is more abstract than putting it at 220,510 m/s. But both statements are false.

One consequence of this is that it is possible for two idealized models (of the same target) to differ in their relative abstractness. Indeed, it is often the case that an idealized model of some phenomenon is presented and theorists in the field then go on to develop less detailed versions of that model, retaining some or all of its idealizations. Here is an example, briefly recounted. It concerns models of the action potential, i.e. neural "spikes". In seminal work, Hodgkin and Huxley (1952) introduced a model that explains the action potential in terms of the interactions of underlying ions (especially sodium and potassium) with the neuron's membrane. More specifically, the HH model describes the neuron's axon as an electrical circuit, which dynamically changes its ionic conductance(s) during the process of the action potential, thereby generating a self-reinforcing "spike" of current. The HH model is idealized in some significant respects; among other things, it assumes that the membrane is perfectly insulating and treats the axon as perfectly cylindrical. Even so, it is fairly concrete, i.e. detailed, describing the incoming and outgoing ionic currents individually and viewing the overall current flowing through the membrane as a composite of them. Because of this complexity, solving the HH equations for all but simple spiking patterns is very difficult. So a variety of less detailed, i.e. more abstract, models of the action potential have been proposed. For instance, Izhikevich (2003) omits all mention of the membrane's structural aspects as well as all information about the constituent ionic currents. Instead, it settles for a single variable to track membrane potential and another one that represents the "resetting" of the membrane's properties after a spike has occurred. However, like other models in this category, Izhikevich's model retains many of the idealizations that Hodgkin and Huxley made—such as simplified axonal geometry and no leakage. Thus, such newer models are abstract, relative to earlier work; they provide less information than the more concrete HH-style model. But they employ many of the same idealizations.[9]

Movement in the other direction is also possible, and quite common too: investigation starts out with an abstract model and moves in the direction of concretization. In ecology, for instance, much early theoretical work consisted of abstract aggregative models of population growth, predation and so on. Perhaps the best-known model in this category is the Lotka–Volterra model, which describes predator–prey interactions. It consists of two coupled differential equations:

[9] Note that I am not claiming that the two models have similar explanatory or predictive power. They do not. I'm only highlighting the fact that while both are idealized, one is more abstract than the other.

$$\frac{dV}{dt} = rV - (aV)\,P \tag{1}$$

$$\frac{dP}{dt} = b\,(aV)\,P - mP \tag{2}$$

Equation (1) tracks the abundance of prey (V): the first term represents the prey's growth rate, and the second the rate at which prey are captured by predators. Equation (2) tracks the abundance of predators (P): the first term represents the rate at which prey is "converted" into new predators, and the second the rate of predator mortality. This model has served as a case study in a number of recent philosophy of science papers (e.g. Weisberg and Reisman 2008; Levy and Currie 2015; Knuuttila and Leottgers, forthcoming). The point of bringing it up here is to note that such models handle populations (of predators and prey) at an abstract level: while populations grow or shrink over time as a result of the myriad actions and interactions among individuals, the model depicts this at an aggregate level, via population-level variables—population size, i.e. V, P—and parameters—birth and death rates, conversion rate etc. i.e. a, b and r. This way of representing a population abstracts away from what specific individuals are like and what they are doing.

In contrast, more recent models in ecology often have a more concrete character, depicting individual-level happenings (though population-level models are still present and important). Such so-called individual-based models (IBMs; see Railsback and Grimm 2011) depict each and every individual in the population, representing their specific properties—whether they are predator or prey, their number of offspring, potentially their location and so on. Moreover, IBMs typically depict dynamics at the individual-level: they assign a set of behavioral rules to each individual and the process then unfolds due to the individual behaviors of members of the populations—where each moves, whether it eats or gets eaten, whether it reproduced and in what numbers etc. So these models can be seen as filling in (some of) the abstract aspects of statistical models. Nonetheless, while models of the latter sort are concrete relative to earlier, population-level models, they often retain many of the idealizations present in the earlier models. For instance, both early and late models typically assume that birth\death rates are constant, independent of population size and other extrinsic factors. Now, to be sure, the individuals depicted in IBMs are depicted abstractly too: they are assigned a sparse set of properties (whether they are predators or prey, their location on a grid, a set of behavioral rules and not much else). Designing and running the model would be difficult otherwise. And this may be significant for understanding their content and evaluating their success. The present point is not that IBMs depict organisms at a maximal (or even a very high) degree of detail, but merely that they provide more detail relative to earlier (idealized) population-level models.

Noticing that idealization and abstractness are compatible can matter for our understanding of various historical developments in science, for instance when there is movement between more and less abstract models of the same phenomenon. By employing the distinction in the right way—by, for instance, doing so without assuming that idealization and abstractness are mutually exclusive—we can describe how a certain area of science has developed. Relatedly, making the distinction between in a way that permits both notions to apply at once can allow us to better assess certain kinds of scientific progress. For instance, in (Levy 2011) I distinguish internal progress from target-oriented progress. The former concerns progress in our understanding of a model, or set of models, while the latter consists in the betterment of our understanding of the model's target in the world. Importantly, one can often achieve a kind of internal progress by filling in details of previously existing models. But if this kind of work retains the idealizations embodied in the older models, then it will not necessarily provide a better depiction *of the target* (this would depend on the details, of course, but the possibility is what I wish to highlight here). The result may be a decoupling of internal progress and target-oriented progress, a situation that may matter a great deal for our assessment of the success or failure of a research program.

4.2 Abstractness and generality

Abstractness and generality are sometimes seen as one and the same, or at least as going hand in hand. On the present way of understanding abstractness, this isn't so. Generality is concerns scope—the number of things (objects, processes, phenomena) to which a representation applies. Abstractness, as we have seen, concerns detail—not which things a representation covers, but how much it says about them. More formally, representation A is more *abstract* than representation B just in case B provides more detail than A about *the same set of objects*. But if A is more *general* than B then the former describes *more things* than the latter. In other words, that A and B bear a relation of relative abstractness to each other implies that both are about the same set of things. But A and B stand in a relation of greater generality relative to each other only if the things to which one of them (the more general one) applies are a proper subset of the things to which the other applies. Thus: "all birds are black" is more general than "all ravens are black". But "all ravens have black parts" is more abstract then "all ravens have black feathers".

In the context of discussions of scientific explanation, for instance, it is important to clearly distinguish abstractness and generality. For the sake of illustration, let us focus on two major contributors to the recent literature on explanation: James Woodward and Michael Strevens. Both advocate a causal approach to explanation, but they develop it in rather different ways. Woodward provides an account of causal relations, in terms of manipulability: an explanation is a description of those factors that can be manipulated in order to make a change to the explanandum. Woodward then emphasizes the importance of generality as an explanatory virtue: an explanation's quality is proportional to the amount of manipulability information it provides.[10] In contrast, while Strevens also provides a causal account of explanation, he emphasizes abstractness and does *not* accord much of a role for generality. Strevens does not provide an account of causal relations per se, and aims to remain neutral on this score. Instead, he holds that an account of explanation ought to describe criteria for selecting among an explanadum's causes those are explanatorily relevant. This, he thinks, is a matter of

[10] Specifically, Woodward thinks of "same-object" generality as important: the more information we get about what would happen to the system under consideration in alternate circumstances, the better. Here, counterfactual scope is at issue—scope as regards the number of "ways the world could have been".

abstractness: a good explanation is one that abstracts away from as much causal detail as possible, thereby distilling those causes that "make a difference" to the explanandum. Moreover, of two explanations that capture all (and only) difference makers for an explanandum, the one that does so in a more abstract way is deemed by Strevens a better explanation (2008, p. 134–37). Here, one sees a kind of "less is more" attitude, in which an explanation is seen as better if it directs our attention to the minimal set of conditions necessary for the explanandum to occur.

I will not discuss the relative merits of these two accounts of explanation. My aim is to highlight how abstractness and generality differ and why this matters. For one thing, we see in Woodward and Strevens appeals to abstractness and generality that stem from different underlying motivations. The former is tied to a kind of minimalism about explanation, which views economy of representation as key. The latter stems from placing value on scope, which is orthogonal to economy. A second point is this: while Strevens and Woodward describe distinct explanatory virtues, they are not incompatible. One can aim for an abstract *and* general explanation, one that provides relatively little detail and has wide scope. It is easy to take the contrast between Strevens' and Woodward's views as implying an incompatibility, but with a clear notion of abstractness in hand, we can see that this isn't the case.

The discussion so far has concerned the concepts of abstractness and generality, independent of other features of the context. I have highlighted the way in which these are distinct concepts, which may play distinct roles, for instance in discussions of explanation. But before leaving this topic I want to point to a subtler way in which abstractness and generality may become quite tightly linked—if further, contingent features are present. I have in mind especially cases of multiple realization. Roughly speaking, multiple realization occurs when a class of systems have a shared property, despite differing in their fine structure. In such situations, moving to a representation that abstracts from structural details often permits one to say something general, applying to a range of systems. But this connection, generalizing by abstracting from irrelevant specifics, depends on whether the details are indeed irrelevant, and does not obtain merely by virtue of some inherent connection between generality and abstraction.

To illustrate this, let us briefly look at an example from recent theoretical neuroscience. Canonical Neural Computations (CNCs) are "standard computational modules that apply the same fundamental operations in a variety of contexts." (Carandini and Heeger 2012, p. 51). The key idea is that some operations, which can be characterized in abstract (computational) terms, are used over and over across different brain regions, modalities and scales. Work on CNCs is ongoing, and this is by no means a settled area. I rely on it here only as an illustration.

One important CNC is divisive normalization (DN). Essentially, in DN the response of a neural element (say, a neuron) is scaled against the response of a population of related elements (neurons), either ones with a similar selectivity, (i.e. sensitivity to similar stimuli), or ones with a different selectivity. The simplest DN model takes the following form:

$$R_j = \gamma \frac{D_j^n}{\sigma^n \sum_k D_k^n}$$

R_j represents the response of element j, given a driving input D over the pool of elements k. Here γ, σ and n are parameters determining overall responsiveness, saturation, and amplification of individual elements, respectively. The equation entails that total response rises with driving input D, but is attenuated by the population's behavior, represented as a sum in the denominator. Thus we have a kind of averaging that smooths out some of the peculiarities of the response and scales relative to its surrounding field. Caradini and Heeger argue that, depending on its context of operation, normalization can allow a neural system to perform important neurcomputational functions—such as filtering out some stimuli, or fine-tuning the gain of an input to the system. Note that in this connection generality—the application of the same kind of computational model to a variety of neural systems—is achieved by abstracting from the specifics of their implementation. And we can do this because the specifics do not matter—the same computational operation can be realized in different ways by the underlying neural machinery. We leave out such information, concerning the specifics of the realizing neural machinery, and the result is a model with greater generality, one that applies to a range of phenomena. But it is important to note that abstracting will result in greater generality only if, indeed, the specifics of the underlying realization don't matter.

In sum, I have emphasized the conceptual distinctness of abstractness and generality and I have also suggested that there may be subtle connections between them. But it is important to see that these connections obtain, when they obtain, in virtue of specific features of a given domain, particularly multiple realization, and not merely in virtue of a conceptual connection between generality and abstractness.

4.3 Idealization and abstraction in real life

Finally, before summarizing, let me say a few words about distinguishing idealization and abstraction in practice. As so often in life, all the more so in philosophy, the distinction is easier to state (in the abstract?) than it is to apply to actual cases. Sometimes, as the examples given above attest, we can say fairly easily whether a certain model contains one or the other. But this isn't always the case. In particular, it is often difficult to tell whether a certain feature is (falsely) presumed absent, hence idealized. Or whether it is merely abstracted away from, such that little or nothing is said about it.

As an example, consider models of large biomolecular complexes—molecular machines, as they are sometimes called—such as ribosomes. Biology texts often portray the ribosome in terms of an ordered, "monotonic" set of steps—it binds to an mRNA molecule, recruits co-factors, then moves along the mRNA joining together amino acids to generate a protein, until it terminates. But ribosomes, like any molecule, are immersed in a highly noisy thermal environment. In actuality, they do not go through a fixed sequence of predictable steps. The ribosome in fact moves back and forth along the mRNA, attaches and detaches, constantly changing conformation. However, it is hard to tell whether standard depictions of this process are idealized or abstract (or both). Do they portray the ribosome as having a sequential, deterministic character, *contra* the realities of ribosomal action? Or is it that the molecular models exhibit a high degree of abstractness, meant to highlight certain functional states and activities? When such models are presented, there is often no indication one way or another.

Why would this matter—why care whether standard models idealize the thermal environment of the ribosome or abstract from it? First, and perhaps fairly obviously, the answer should affect how we, and more importantly the relevant scientists, understand the state of knowledge concerning the ribosome and consequently shape future research in this area. If standard models misdescribe the ribosome's actions, a natural question is whether this misdescription can be corrected for, and if so what would be the costs and benefits of such a correction. Can we obtain a more accurate model, free of idealization? If so, can we use it to make the same predictions and explain the same phenomena? On the other hand, if standard depictions of the ribosome are best seen as abstracting from thermal noise and its effects, then the question arises whether the missing details are known and, if so, why standard depictions omit them—is it that they do not matter? If so, for what purposes and why?[11]

The question is also relevant for more philosophical reasons. For instance, one central debate in recent philosophy of biology concerns the scope and character of the mechanistic approach to explanation. In part, the issue depends on how one interprets current work in biology, and specifically whether explanation in areas such as cellular and molecular biology conforms to the image of macromolecules as "miniature machines". To settle this interpretive issue it is crucial to understand how current scientific representations of ribosomes (and similar macromolecular structures) are to be read—whether they knowingly misdescribe it, or rather simply lack detail about certain aspects of its structure and operation. For this will matter for our assessment of the norms underlying such explanations: from what is missing in an explanation, and from how the missing aspects are viewed by practitioners, we can learn what they view as necessary to advance their knowledge and understanding (Skillings 2015).

5 Summary

Let me summarize the main points briefly. I proposed to distinguish the process of abstraction from the abstractness of its products, and to construe the latter as relative degree of detail. Idealization, I suggested, is best seen as deliberate misrepresentation. In idealization, due to the role of the idealizer's knowledge and intentions, the process is more intimately tied to the product. A proper understanding of these important categories can make a difference to our understanding of certain issues and situations in science, such as the state of knowledge in some domain, and affects how we conceptualize more abstract philosophical claims, such as the nature of scientific explanation and the role of theoretical virtues like generality.

[11] Details of the ribosome's molecular structure, for instance, matter greatly for some purposes (Ramakrishnan 2014).

References

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13,* 51–62.

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Oxford: Oxford University Press.

Elliot-Greaves, A., & Weisberg, M. (2014). Idealization. *Philosophy Compass, 9,* 176–185.

Frigg, R. (2006). Scientific Representation & the Semantic View of Theories. *Theoria, 21*(1), 49–65.

Giere, R. N. (1988). *Explaining science: A cognitive approach*. Chicago: The University of Chicago Press.

Godfrey-Smith, P. (2006). The strategy of model-based science. *Biology and Philosophy, 21,* 725–740.

Godfrey-Smith, P. (2009). Abstractions, Idealizations, and Evolutionary Biology. In A. Barberousse, M. Morange, & T. Pradeu (Eds.), *Mapping the future of biology: Evolving concepts and theories*. Berlin: Springer.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology, 117*(4), 500–544.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks, 14*(6), 1569–1572.

Jones, M. (2005). Idealization and abstraction: A framework. In M. R. Jones & N. Cartwright (Eds.), *Idealization XII: Correcting the model. Idealization and abstraction in the sciences* (Vol. 86, pp. 173–217). Amsterdam: Rodopi.

Jones, N. (2013). Don't blame the idealizations. *Journal for General Philosophy of Science, 44*(1), 85–100.

Knuuttila, T., & Leottgers, A. (forthcoming). Modelling as indirect representation? The Lotka–Volterra model revisited. *British Journal for Philosophy of Science*.

Laymon, R. (1995). Idealizations and the testing of theories by experimentation. In P. Achinstein & O. Hannaway (Eds.), *Observation, experiment, and hypothesis in modern physical science*. Cambridge: MIT Press.

Levy, A. (2011). Game theory, indirect modeling and the origins of morality. *Journal of Philosophy, 108*(4), 171–187.

Levy, A. (2015). Modeling without models. *Philosophical Studies, 172*(3), 781–798.

Levy, A., & Bechtel, W. (2013). Abstraction and the organization of mechanisms. *Philosophy of Science, 80*(2), 241–261.

Levy, A., & Currie, A. (2015). Model organisms are not (theoretical) models. *British Journal for Philosophy of Science, 66*(2), 327–348.

McMullin, E. (1985). Galilean idealization. *Studies in History and Philosophy of Science, XVI,* 247–273.

Nowak, L. (1992). The idealizational approach to science: A survey. In J. Brzeziński & L. Nowak (Eds.), *Idealization III: Approximation and truth*. Amsterdam: Rodopi.

Railsback, S. F., & Grimm, V. (2011). *Agent-based and individual-based modeling: A practical introduction*. Princeton: Princeton University Press.

Ramakrishnan, V. (2014). The ribosome emerges from a black box. *Cell, 159,* 979–984.

Rosen, G. (2012). Abstract objects. In E. N. Zalta (ed.) *The stanford encyclopedia of philosophy* (Fall 2017). http://plato.stanford.edu/archives/fall2014/entries/abstract-objects/.

Skillings, D. J. (2015). Mechanistic explanation of biological processes. *Philosophy of Science, 82,* 1139–1151.

Strevens, M. (2008). *Depth: An account of scientific explanation*. Cambridge: Harvard University Press.

Weisberg, M. (2007). Who is a modeler? *British Journal for Philosophy of Science, 58,* 207–233.

Weisberg, M. (2013). *Simulation and similarity*. New York: Oxford University Press.

Weisberg, M., & Reisman, K. (2008). The robust volterra principle. *Philosophy of Science, 75*(1), 106–131.