

Simplicity in the Best Systems Account of Laws of Nature

James Woodward

ABSTRACT

This article discusses the role of simplicity and the notion of a best balance of simplicity and strength within the best systems account (BSA) of laws of nature. The article explores whether there is anything in scientific practice that corresponds to the notion of simplicity or to the trade-off between simplicity and strength to which the BSA appeals. Various theoretical rationales for simplicity preferences and their bearing on the identification of laws are also explored. It is concluded that there are a number of issues about the role of simplicity within the BSA and its relation to strength that need to be addressed before the BSA can be regarded as an adequate account of laws.

- 1 *Introduction*
 - 2 *The Best Systems Account*
 - 3 *The Trade-Off between Simplicity and Strength: Preliminary Considerations*
 - 4 *Alternative Conceptions of the Relationship between Simplicity and Strength*
 - 5 *Two Roles for Simplicity*
 - 6 *Simplicity in the Best Systems Account: Curve-Fitting*
 - 7 *Simplicity as a Corrective for Overfitting*
 - 8 *Descriptive Simplicity in the Best Systems Account?*
 - 9 *Simplicity as Due to Human Intellectual Limitations*
 - 10 *Summary*
 - 11 *Concluding Remarks*
-

1 Introduction

Many philosophers of science regard the best systems account (BSA), according to which laws of nature are the theorems or axioms common to systemizations of the full Humean basis (HB) that achieve a best balance of simplicity

and strength, as the most attractive contemporary approach to laws of nature. The appeal of the BSA does not, I believe, mainly derive from its demonstrated descriptive adequacy as a treatment of detailed aspects of scientific practice involving laws. It rather has to do with its overall fit with other ideas to which many philosophers are committed: two of these are a picture of scientific reasoning as involving a trade-off between simplicity and strength, and a ‘Humean’ programme of reduction of the nomic to the non-nomic. Yet another consideration has to do with the character of the main competing accounts of laws of nature. These tend to postulate objects and relations, which seem (to naturalistically minded philosophers) to be metaphysically extravagant—for example, relations of necessitation between universals (Dretske [1977]; Armstrong [1983]; Tooley [1987]) or dispositional essences (Bird [2007]) and their ilk. By contrast, the BSA appeals to features that (it is argued) play a role in scientific methodology, apparently avoids extravagant metaphysics, and seems to yield a reduction of the modal commitments of physical theorizing commitments that many find puzzling. This makes many philosophers think that something along the lines of the overall package must be right and perhaps that they ought to pay less attention than they should to the details of exactly how the account is supposed to work.

This article attempts, in a partial way, to redress this balance. My aim is modest: I do not claim to have produced arguments that ‘refute’ the BSA. Rather, what I try to do is to explore a set of questions concerning the role of ‘simplicity’ and ‘best balance’ in the BSA, focusing on how these notions are best understood, the normative rationale (if any) for their deployment, and whether they correspond to anything in the actual practice of science. In particular, I will focus on questions like the following: Is it true that scientific practice involves a trade-off between simplicity and strength of the sort that is described in the BSA? Are there normative reasons why science should exhibit such trade-offs? Is there any reason to suppose that it is via engaging in such trade-offs that scientists discover or identify laws? Although much of my discussion will be sceptical or critical in tone, my primary goal is constructive: to encourage defenders of the BSA to be more explicit and precise than they have hitherto been about the core concepts in their doctrine.¹

To focus my discussion, I will for the most part accept, for the purposes of this article, many aspects of the BSA that might easily be challenged, such as the ideas that laws themselves should be conceived as generalizations

¹ For a brief but trenchant discussion arguing that the BSA trade off between simplicity and strength corresponds to nothing in scientific practice, see Roberts ([2008], pp. 9–10). The arguments that follow are in the spirit of Roberts’s remarks, but I have tried to show in more detail why we should not expect to find such trade-offs and to raise further questions about what is meant by simplicity.

describing exceptionless regularities, rather than as ‘effective’ generalizations that hold within certain domains or regimes and break down outside of these, as arguably is the case for all known laws. I will also accept the idea that systemization of the HB is a useful idealization for how science, as practiced by human beings, works. I will also forego the common practice of appealing to ‘intuitions’ about lawful relations in allegedly metaphysically possible worlds that are different from the actual world (worlds consisting just of ‘one lone proton’, as in Lange ([2000]), and so on) to criticize the BSA.

Several other preliminary remarks: First, in what follows, my focus will be on the descriptive and normative adequacy of the BSA as an account of the role of simplicity and laws in science as practiced in roughly the period from the scientific revolution to the present. Some philosophers hold that there are metaphysical considerations that support the BSA, independent of its adequacy as an account of scientific practice.² I am sceptical of such claims but will not try to argue against them here. One justification for this focus is that adherents of the BSA often attempt to motivate it by claiming that it is rooted in or describes a methodology that is in fact used in many areas of science. Another justification is this: defenders of the BSA rarely spend much time trying to precisely characterize the notions of simplicity, strength, and best balance on which they rely.³ My guess is that failure to be precise about

² Such a view is urged by one of the referees for an earlier version of this article. The referee also suggests that we may think of the BSA as proposing a ‘revisionary’ view of laws. According to this interpretation, ‘there are philosophical problems with the notion of law that has come down to us through the history of physics’, and these problematic features are corrected or revised by the BSA. Thus, it does not necessarily tell against the BSA if it fails to capture any notion of law used in science because such notions embody philosophically problematic elements. I will not address this possibility in this article because my focus is science as currently practiced, which reflects those features of the notion of law that the referee considers problematic. I nonetheless agree with the referee that if the BSA is not adequate as an account of law in current science, then the possibility remains open that we should respond by revising our concept of law rather than the BSA. I will add, however, that there is an obvious danger of trivialization if we follow this strategy unless the claim that our present concept of law is problematic can be motivated independently of the BSA. In other words, if we hold that our present concept of law is problematic simply because it incorporates non-BSA features and therefore needs to be replaced by a revised concept, that of a BSA-law, which is exactly what is captured by the BSA, thereby vindicating the BSA, we will be arguing in a small circle indeed.

³ This is arguably less true for the notion of strength than for the notions of simplicity and best balance, although here too there are obscurities, only some of which I will consider below. One important unclarity that I will not discuss in the main body of this article is whether strength is a function of the variety of different types of phenomena derivable from the best system (BS) or whether it instead also depends on the number of instances (tokens) of those types that occur in the Humean basis (HB). In other words, if P is some type of phenomenon derivable from a candidate best systemization, is that systemization to be counted as stronger to the extent that there are many, as opposed to only a few, instances of P occurring in the HB? Or does derivability of some P with one instance count just as much as the derivability of the same P with many instances, with gains in strength tracking only the extent to which different types of phenomena are derivable? Counting the number of instances is arguably more in the spirit of the BSA, but counting just types seems to be more in accord with scientific practice. However, in the latter case, how does one tell when one is dealing with ‘different’ types? Another issue is this: when the deductive consequences, C_1 , of systemization, S_I , that are within the HB are a proper subset of

these notions strikes them as relatively undisturbing at least in part because it is assumed to be obvious that these notions are exemplified in scientific practice. If this is so, we can, so to speak, look to science itself to specify what counts as a 'simple' systemization or one that achieves a best balance of simplicity and strength, without being explicit about what these notions involve. In other words, vagueness in the official definitions of simplicity and balance are thought of as tolerable because the concrete details of what simplicity, strength, and best balance involve are supplied contextually in particular cases by science itself. Obviously, this strategy depends on it being plausible at an overall level that scientists themselves are guided by the balancing goal that characterizes the BSA. If it is dubious that science aims at any such goal or identifies laws via such a goal, then we lose an important motivation for the BSA and, given the absence of any detailed explicit specifications of its central concepts, it becomes unclear just what the account amounts to.

The remainder of this article is organized as follows. Section 2 briefly reprises the basic ideas of the BSA. Section 3 then raises a puzzle concerning the role of information about initial conditions in the best systemization, and explores some implications for the relationship between simplicity and strength. Section 4 considers some possible alternative views about the relationship between simplicity and strength besides the trade-off view incorporated in the BSA. Sections 5–10 then explore various issues about the role played by simplicity in the BSA. Section 11 concludes.

2 The Best Systems Account

I assume that the BSA is sufficiently well-known that it does not require detailed description.⁴ I will largely follow David Lewis's development of the basic framework (Lewis [1999]) because other advocates largely accept his formulation.

Begin with the HB, a full description of what 'actually' happens everywhere and at all times throughout the entire history of the universe, but with all modal elements removed from this characterization. This means (at least) that the HB does not contain information about causal or lawful relationships or relationships of counterfactual dependence, or claims about physical possibility and necessity. Positively, in Lewis's version, the HB consists of something like a specification of the 'perfectly natural monadic properties' instantiated at each space-time point, together with the facts about the spatio-temporal

the deductive consequences C_2 of S_2 that are within the HB, it is uncontroversial that S_2 is stronger than S_1 , but how do we evaluate comparative strength in cases in which neither S_1 nor S_2 is a proper subset of the other?

⁴ I will focus in what follows just on deterministic laws and put aside best systems treatments of 'objective chance'.

relations among those instantiations.⁵ Given the HB, we are then to consider alternative deductive ‘systemizations’ of it. The best systemization or systemizations are those that achieve a ‘best balance’ of two virtues: strength and simplicity. Strength is, roughly, a matter of informativeness or logical content, although as we shall see below, the way in which this notion is to be interpreted in the BSA requires some clarification. Lewis says relatively little about the relevant notion of simplicity but he seems to have in mind a notion that is at least in part ‘objective’, and not just relative to human psychology, and that is rather domain-general—even (to a substantial extent) syntactic. He writes:

I suppose our standards of simplicity and strength and balance are only partly a matter of psychology. It’s not because of how we happen to think that a linear function is simpler than a quartic or a step function; it’s not because of how we happen to think that a shorter alternation of prenex quantifiers is simpler than a longer one; and so on. Maybe some of the exchange rates between aspects of simplicity, etc., are a psychological matter, but not just anything goes. ([1999], p. 232)

According to Lewis, assessments of simplicity must be made relative to a language or representational scheme—a hypothesis that looks simple in one language may look complex in another. Lewis’s response to this is to suggest that there is a privileged or canonical language in science, which is the language that should be used for making comparisons of simplicity—this is the (or a) language containing predicates describing ‘perfectly natural properties’.

As Lewis sees it, simplicity and strength are competing virtues that trade off against one another: increasing strength generally decreases simplicity, and increases in simplicity are generally bought at a cost in strength. The best systemization (or systemizations) is (are) the one (or ones) that consist(s) only of true propositions and achieves an optimal balance of simplicity and strength. Although, according to Lewis, the best systemization may contain claims that do not describe regularities,⁶ only regularities count as laws.⁷

⁵ A number of commentators (for example, Maudlin [2007]; Hall [forthcoming]) have noted respects in which this characterization is problematic, but little will turn on these in what follows. However, it is worth noting that the inclusion of information about spatio-temporal relations in the HB, in conjunction with the requirement that all such information be non-nomic and non-modal in character, apparently commits the defender of the BSA to the claim that spatio-temporal relationships can be adequately specified without reference to the laws of any background spacetime theory and without reference to notions like ‘possible positions’ of bodies.

⁶ Lewis ([1999], p. 41) writes:

The ideal system need not consist entirely of regularities; particular facts may gain entry if they contribute enough to collective simplicity and strength. (For instance, certain particular facts about the Big Bang might be strong candidates.) But only the regularities of the system are to count as laws.

⁷ Some writers (for example, Albert [2000]; Loewer [2009]; Hall [forthcoming]) favour dropping this feature of Lewis’s formulation and allowing (what they think of as) particular matters of

The laws are those regularities that are described by the axioms and theorems that are common to all such best systemizations or, alternatively, the generalizations that describe such regularities.⁸

Why suppose that balancing simplicity and strength in this way will pick out generalizations corresponding to laws? An important part of Lewis's answer (and the only motivation with which I will be concerned in this article) is that, as a matter of descriptive fact, this assumption or standard guides scientific practice in theory choice and the identification of laws. He writes:

I take a suitable system [that is, a best systemization] to be one that has the virtues we aspire to in our own theory-building, and that has them to the greatest extent possible given the way the world is. ([1999], p. 41)

The standards of simplicity, of strength, and of balance between them are to be those that guide us in assessing the credibility of rival hypotheses as to what the laws are. ([1986], p. 123)

3 The Trade-Off between Simplicity and Strength: Preliminary Considerations

How exactly is all this supposed to work? What information makes it into a best systemization? As a sort of warm-up exercise, let me begin with an initial puzzle having to do with the fact that a great deal of information of scientific interest—in particular, a great deal of information about initial conditions—apparently cannot make it into the best systemization.

To motivate the puzzle, let us consider the notion of strength in a bit more detail. On one possible understanding of this notion, the strength of a systemization is its logical strength as measured by all of its deductive consequences; if, for example, the full set of deductive consequences of systemization B_1 is a proper subset of the deductive consequences of B_2 , then B_2 is stronger than B_1 . Although informal expositions of the BSA often invoke this notion of strength, it seems fairly clear that this is not the notion intended by Lewis or the notion that most developments of the BSA rely on. Rather, the

fact—for example, alleged facts about the early universe such as those described by the 'past hypothesis'—to count as lawful or at least nomologically necessary. I will ignore this possibility in what follows because it does not seem relevant to the criticisms I will be exploring. I emphasize, however, that these other writers agree with Lewis, that claims about particular matters of fact (initial and boundary conditions) can be part of the best systemization whether or not these are regarded as laws. This is enough to motivate the question that I pursue below: whether some of these particular facts, perhaps in conjunction with other claims in the BSA, will entail regularities that are 'accidental'.

⁸ Different authors use the word 'law' to refer either to regularities occurring in nature, as in the passages from Lewis above, or to generalizations or representations describing such regularities. I find the latter usage clearer, but to avoid unnecessary pedantry will use the word in both ways, trusting the context to sort out what is meant.

operative notion of strength should be understood as something like ‘strength with respect to the HB’. That is, a systemization is stronger to the extent that it is more informative about the HB or allows us to derive more consequences that are represented in the HB. In other words, deductive consequences of a systemization that are not represented in the HB do not count as adding to the strength of the systemization, only consequences in the HB do. This fits with Lewis’s claim that generalizations corresponding to ‘uninstantiated’ laws that follow from a systemization do not contribute to the strength of that systemization. It also fits with generally empiricist orientation of the BSA because from the point of view of such an orientation, it is hard to see why the fact that systemization has many implications regarding unrealized possibilities should, in itself, conduce to the credit of that systemization.⁹

Assuming that this is the right way to understand strength, consider a world that consists just of a structure behaving like our solar system as this is represented in Newtonian mechanics: planets orbiting a ‘sun’ and so on, with the laws associated with this structure being the gravitational inverse square law and Newton’s laws of motion. The conventional scientific representation of this world will consist, in addition to these laws, in detailed information about ‘initial conditions’ having to do with the masses, positions over time, and so on of the objects in this system.¹⁰ Call the deductive closure of all of the propositions describing these laws and initial conditions *S*. *S* will thus contain some specific claims about particular matters of fact that do not hold as a matter of law but it will also contain generalizations describing regularities that, intuitively speaking, are non-lawful. A standard example is the generalization *R* that all of the planets orbit their sun in the same direction. If the laws of nature are the axioms and theorems of the systemization *S*, we would be required to count *R* as a law because *R* is in *S*. Assuming that *R* is not a law, we need to find some way of excluding it from counting as a law, while counting generalizations like the inverse square law as genuine laws. Roughly, the BSA proposes that this can be accomplished by appealing to the notion of a best balance of simplicity and strength. The argument is that the addition of *R* to *S* would make the resulting systemization less simple in a

⁹ Put slightly differently, it seems contrary to the spirit of the HB that one should be able to improve the goodness of a candidate systemization just by adding additional claims to it that are simple and deductively strong, even though these claims have no connection with what is reported in the HB.

¹⁰ Two comments about this example: First, although for convenience it involves a ‘small world’ (a solar system), I see no reason to suppose that anything essential would change if we were to ‘scale up’ to the description of some much larger structure, such as the entire universe. (At least, it seems arbitrary to simply claim that the problem disappears when we scale up without providing any further argument for this claim.) Second, note that the description of the solar-like system in the example is in terms of properties that count as ‘natural’—presumably ‘mass’, and ‘position’ count as natural properties if anything does. The problem raised by the example is thus in not any way an artifact of its ‘small world’ character or description in terms of less than ‘perfectly natural’ properties.

way that is not compensated for by any increase in strength provided by the inclusion of R . That is, R contributes only a small increase in strength at the cost of a significant loss of simplicity. Of course, it also needs to be shown that achieving a best balance of simplicity and strength requires that we not allow into the best systemization any propositions that, in conjunction with what is already in that systemization, imply R .

Let us call this best balancing systemization (that excludes R and other similarly non-lawful regularities) for the solar system world S^* . As we have just seen, it appears that a great deal of information that is in S cannot not be in S^* , at least if we want to capture generally accepted ideas about which generalizations are laws. For example, any initial condition information about the velocities of the planets or their successive positions over time that would allow us to deduce that they are all moving in the same direction must be excluded from S^* . Similarly for many other generalizations describing non-lawful regularities in S , such as generalizations about the spacing of the planetary orbits (as in Bode's law) or generalizations describing their repeated orbital trajectories (successive positions) over time, if (as I assume) the latter are 'non-lawful'.¹¹ Because these non-lawful generalizations can be deduced from uncontroversial candidates for laws like the inverse square law and Newton's laws of motion in conjunction with appropriate information about initial conditions in S , and because we can hardly drop these uncontroversial laws from the best balancing systemization S^* , our only alternative seems to be to exclude any information about initial conditions that might permit such derivations. In doing this, we of course end up with a limited capacity to deduce much of anything about our solar system world from what remains in S^* . Indeed, it seems entirely possible that in this particular case, virtually all information about initial conditions will be excluded from S^* and that S^* will consist of little more than the laws of motion, the inverse square law, and what follows deductively from these alone. The resulting systemization, S^* , will not be strong (with respect to the HB) in absolute terms (indeed, it will be weak in comparison with S), but it will satisfy the desideratum of permitting the deduction of only those generalizations that are laws.¹²

¹¹ Some may be inclined to regard some of these generalizations as laws. I assume, however, that if the BSA project is to make sense, there must be some contrast between lawful and non-lawful regularities that the BSA captures. Pick your favourite examples of non-laws; the issue I mean to raise is how the BSA avoids incorporating information about initial and boundary conditions that allows for the derivation of the latter, while also incorporating the idea that the BS must be strong.

¹² A referee for an earlier version of this article drew attention to the fact that conditionals that state that if certain initial conditions are realized, other conditions would follow just from the generalizations in the BS that are intuitively laws, without any need for additional information about initial conditions. The referee's suggestion is that we can think of these conditionals as contributing to the strength of the BS and, because they include no claims about particular matters of fact, they will not be such that they will allow for the deduction of accidental regularities. For example, according to this suggestion, we may think of Newton's laws, N , as

How troubling is this? The partisan of the BSA might respond that if, as we have been supposing, the BS consists solely of laws and contains virtually no information about initial and boundary conditions, then (as suggested above) it is true that it is less strong than one might initially have supposed. Nonetheless (it might be argued), such a BS imposes some constraints on what can occur in the HB; for example, any law of form ' $(x) (Px \rightarrow Qx)$ ' rules out the occurrence in the HB of pairs of propositions of the form Pa and not Qa .¹³ In this, and perhaps other respects, such a BS will be informative about (or have some strength with respect to) the HB. So it might well be true that such a BS achieves a best balance of simplicity and strength, making up for a relatively low-strength score with a high-simplicity score.

Nonetheless, I think it is a source of concern that the notion of strength associated with this version of a BS fails to make much connection with what one would naturally think of as the primary strength-like virtue possessed by paradigmatic scientific theories. For example, much of what most impressed successive generations of scientists about Newtonian theory was its ability, when combined with detailed information about initial conditions, to provide detailed predictions of the behaviour of the planets. Intuitively, this is a central part of what the 'strength' of Newtonian theory consists in. One would think that this systematic predictive success ought to play some important role in supporting the judgement that generalizations like the inverse square law are genuine laws. But if we exclude virtually all information about initial conditions from what we are prepared to regard as a best systemization, we cannot deduce such detailed predictions about the content of the HB from the BS itself.

One way out of this difficulty is to give up on the idea that 'strength' has to do just with what can be deduced about the HB from the BS by itself. Here is an alternative proposal: the strength of a systemization is (among other things

adding strength to the best system by virtue of having implications of the form 'if A , then P ' for various predictions, P , regarding planetary positions, where the A s are claims about initial conditions. One problem with this suggestion is that conditional claims of this sort allow one to deduce little about the HB and in this respect do not seem to capture the idea that the best systemization should be strong with respect to the HB. Another problem is that many of the conditionals that are so deducible seem irrelevant to strength, at least if the strength of a systemization has something to do with the information it provides about the HB, rather than its logical implications more generally. This is true, for example, of uninstantiated conditionals of form 'if A , then P ' for which A is not realized in the HB, some of which will be deductible from the BS.

¹³ This was suggested by a referee for an earlier version of this article. This assumes, of course, that a systemization is stronger to the extent that it allows for the deduction of claims to the effect that something does not occur in the HB (in addition to the deduction of claims about what does occur in the HB). In other words, it adds to the strength of a systemization if it rules out possibilities that are not represented in the HB, rather than being silent about them. This assumption seems to be a natural way of interpreting the notion of strength, but I do not know if most defenders of the BSA wish to endorse it. This is another respect in which the notion of strength stands in need of clarification.

but perhaps primarily) a function of the extent to which claims in the systemization can be combined with some facts reported by the HB to deduce other facts in the HB; the more such facts in the HB that can be so deduced from a ‘small number’ of other facts in the HB, along with premises in the BS, the stronger this systemization. Thus, we give up on the idea that the strength of a systemization just has to do with what we can derive about the HB from the systemization alone; instead, it has to do with the extent to which we can use the systemization to derive some facts in the HB from others. For example, if a candidate BS is such that the claims in it can be combined with facts from the HB reporting what happens at a particular slice of time to deduce all remaining facts concerning what happens at all other times that are contained in the HB, this systemization will count as relatively strong. (The facts reported in the HB about what happens at a particular time slice are thus not contained in the BS.) If the BS is such that the claims in it, in conjunction with many facts from the HB can be used to deduce only a few remaining facts in the HB, such a BS will count as weaker. The overall strength of a systemization will then have to do with this consideration, presumably combined in some way with considerations about which facts, if any, about the HB can be deduced from the systemization alone, without any other information. This way of understanding strength adds additional layers of complication to the BSA and will make judgements of the relative strength of competing systemizations more difficult (because, among other things, we now need some way of measuring when we have deduced ‘many’ facts from a ‘smaller number’ of others), but it does capture a notion that seems more directly connected to the role of laws in empirical prediction. Whether, when ‘traded off’ against simplicity, this (or any of the alternative understandings of strength considered above) yields an account that successfully picks out those generalizations that are regarded as laws in scientific practice is an issue I will explore below.

4 Alternative Conceptions of the Relationship between Simplicity and Strength

To further explore the commitments of the BSA, consider for purposes of comparison the following alternative picture of how, at least sometimes, although perhaps not always, strength and simplicity trade off.¹⁴ Often in so-called mature sciences, there is some pre-specified domain of results (or phenomena) that competing theories are expected to account for—this would

¹⁴ This proposal is advanced just for purposes of comparison, I do not claim that this alternative is unproblematic or accurately describes all cases of theory-choice; merely that for some interesting range of case it is at least as plausible as the simplicity/strength trade-off envisioned in the BSA. And it does seem to fit the methodological pronouncements of some famous scientists better than the BSA view—see below.

include facts about planetary trajectories in the example under discussion. To count as satisfactorily strong, a theory must account for or successfully predict all (or virtually all) of these results (or, alternatively, perhaps at least as many as rival theories predict). A theory that meets this condition may still be judged unsatisfactory if it is sufficiently non-simple or *ad hoc*, but this does not involve trading off simplicity and strength in the manner assumed in the BSA. Instead, to the extent that strength and simplicity come into conflict (and they may not),¹⁵ strength has something close to lexical priority: for a theory to be regarded as satisfactory at all it must have a high level of strength (it must meet a ‘strength threshold’) with respect to the phenomena in its domain; competing theories meeting this threshold are then evaluated with respect to simplicity. If a theory falls below this strength threshold (that is, if it fails to account for many of the established phenomena in its domain), no compensating gain in simplicity will be enough to make up for this shortfall. I do not claim that this understanding of the relationship between simplicity and strength is found in all areas of science—in fact, I think it is often not adopted in the so-called special sciences when these deal with complex systems, on which point more below—but I think it is frequently adopted in areas of science that are regarded by philosophers as ‘fundamental’, such as portions of physics. (Of course, it is such fundamental science that the BSA purports to reconstruct, insofar as it purports to reconstruct any part of science.) Something like this lexical priority or threshold idea is perhaps the most natural interpretation of the first of Newton’s ‘Rules of Reasoning in Philosophy’, which is often understood as a claim about the role of ‘simplicity’ or ‘Ockham’s razor’ in science: ‘We are to admit no more causes of natural things than such as are both true, and sufficient to explain the appearances’ ([1968]).¹⁶ This rule seems to say that if a cause is required to explain appearances, we should admit it, but not otherwise. The rule does not say that we should not introduce causes that are

¹⁵ Interestingly, many fundamental laws, such as the laws of Newtonian mechanics or Maxwell’s equation seem to be both simple and strong in the sense that they can be used in conjunction with information about initial and boundary conditions to derive precise claims about what happens. In such cases, high simplicity and high strength (in the sense described) both seem attainable.

¹⁶ Ockham’s razor is often glossed as the advice not to multiply entities beyond necessity. In the current context the crucial question is what ‘necessity’ means. The BSA assumes, in effect, that if one increases strength through additional entities in a way that fails to achieve a ‘best balance’, one has multiplied unnecessarily, even if omission of these entities leaves some phenomena unaccounted for. A different view, epitomized in the quotation from Einstein in the main text, is that adding additional entities when these are required to account for known phenomena is never a case of multiplying beyond necessity, even if this leads to a considerably more complex theory. My guess is that part of the apparent appeal of the idea that fundamental science involves trading off simplicity and strength in the manner envisioned in the BSA is the result of a failure to distinguish the BSA conception of this trade-off from the different (and more plausible) idea that one should not introduce additional complexity unless required to do so to gain additional strength, but that if even small gains in strength are achievable, additional complexity to achieve such gains is mandatory.

sufficient to explain appearances if the resulting loss in strength is outweighed by the gain in simplicity of postulating fewer causes, which is what the trade-off postulated in the BSA seems to require. Instead, introduction of additional causes is always justified (perhaps required) if this leads to an improvement in strength. This ‘threshold’ idea that strength has priority over simplicity is even more clearly endorsed by Einstein ([1934], p. 165):

It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience.

To underscore how this differs from the sort of simplicity/strength trade-off contemplated in the BSA, note that if the latter were the norm, it would be regarded as satisfactory for a theory to fail to account for a large number of results in its domain if it makes up for this failure with a sufficiently large gain in simplicity. One would expect to find arguments like the following: my theory does not predict much about planetary positions or movements even given information about initial conditions, but it is so much simpler than theories that do so that it achieves a better overall balance of simplicity and strength and is preferable for that reason. I take it that, at least in many contexts, good scientific methodology does not endorse arguments like this. Instead, it often seems to make recommendations that are closer to the threshold picture: unless your theory predicts a lot about the planetary motions, it is unsatisfactory no matter how simple it is.

Suppose that something like the threshold view of the relation between simplicity and strength is correct as an account of how fundamental theorizing in much of science proceeds, and suppose that we also accept that because this trade-off is the one found in scientific practice, it should also be the one on which the BSA relies. The problem this creates for the BSA is that because on the threshold view simplicity is given much less weight than strength, many more generalizations that look accidental look likely to make it into the best systemization. For example, if the threshold conception of the trade-off between simplicity and strength is adopted, generalizations like *R* may be good candidates for incorporation into the BS. We can no longer argue that although inclusion of *R* would add to the strength of the BS, this gain is outweighed by the loss in simplicity that would result from its inclusion. This is because, on the threshold view, strength gains count for much more than simplicity gains. It is for just this reason, I take it, that adherents of the BSA do not understand the notion of a ‘best’ trade-off in terms of the threshold view—instead their account requires a notion according to which significant decreases in strength can be compensated by appropriate gains in simplicity. The concern raised in this section is that this conception of how

simplicity and strength trade off is apparently not a conception that plays a role in fundamental science.

Here is a possible response to this concern, urged by a referee in response to a previous draft of this article. We should distinguish two possible roles played by appeals to a trade-off between simplicity and strength in science. One is as a criterion for theory-choice (where a ‘theory’ is understood as consisting of claims about both laws and some accidental truths about initial and boundary conditions). Here the threshold view (or something close to it) is correct: in evaluating whether a theory is true (or plausible or worth pursuing), we trade off simplicity and strength in a way that gives priority to strength. However, we should separate this role from the role played by simplicity and strength in distinguishing between laws and accidental claims. Here the trade-off posited by the BSA provides the correct account of this distinction. In other words, given a true theory, *T*—picked out by threshold considerations and consisting of claims that may be laws as well as claims that may be accidents—the BSA trade-off in which losses of strength can be compensated by gains in simplicity can then be used to distinguish among the claims in *T*, classifying some as laws and others as accidents. So the BSA remains plausible as an account of laws, even though theory-choice in science generally conforms to the threshold view.

What should we say about this proposal? If defensible, it certainly represents an important clarification of BSA doctrine. However, I do not find the proposal compelling. To begin with, there is the concern that the proposal looks ad hoc and defensive. The defender of the BSA begins by noticing that insofar as trade-offs between simplicity and strength occur in science, they often seem to conform to the threshold view, and that threshold trade-offs cannot be used to distinguish laws from accidents. She responds by claiming that there is also a second sort of simplicity/strength trade-off that fits what is described in the BSA and that can be used to distinguish laws from accidents. But what is the evidence that this second sort of trade-off is used in science and that it can be used to distinguish laws and accidents in the way claimed? (It is not an answer to this question to simply say that simplicity and strength are regarded as important in science because this is compatible with their role being fully described by the threshold view. Instead, what we are asking for is evidence that simplicity and strength are traded off in the distinctive way postulated by the BSA.) Why, for example, are there no passages from Newton and Einstein (or whoever) comparable with those quoted above in which the authors discuss this second role for simplicity/strength trade-offs? Why are there no explicit arguments in scientific treatises appealing to this second role to distinguish laws and accidents?

A related worry is that it is hard to see how to make sense of a sharp distinction between the two roles in practice. For the distinction to work, we must think in terms of a simplicity/strength trade-off procedure in

which a theory, T —understood as a set of claims that have not been distinguished into laws and accidents—is chosen (or accepted or whatever) and then a distinct procedure is used to distinguish the laws and accidents in T . It seems to me, on the contrary, that typically and perhaps always in choosing or evaluating T , one thinks in terms of a T in which the laws and accidental claims are already distinguished. That is, in evaluating T , one evaluates a candidate theory that is articulated in such a way that some of its claims are described as laws and others as accidents. We then draw upon or presuppose this law/accident distinction made by T in assessing both its simplicity and strength. This of course implies that the distinction is made on some basis that is, at least in part, independent of considerations of simplicity and strength. For example, as suggested above, the strength of T is naturally understood as a matter of the extent to which the candidate laws of T can be combined with accidental claims about initial and boundary conditions to derive other claims about what happens—a conception of strength that seems to require that we are already in possession of a law/accident distinction. And as suggested below, simplicity considerations seem to have different implications depending on whether they are applied to initial and boundary conditions or to laws.

At this point, it will be useful to step back and address a more general line of argument that one frequently hears in support of the idea that something like the BSA ‘must’ be correct. The argument is that it is uncontroversial that scientists value both simplicity and strength, and that they also attach a high value to the discovery of laws. Moreover, it is also uncontroversial that simplicity and strength can sometimes conflict. So (the argument continues) even if extant versions of the BSA perhaps do not get all the details right and some further tweaking is required, it seems highly plausible that lawhood must somehow be bound up with ‘balancing’ simplicity and strength. One is thus led to the conclusion that the BSA must be on the right track, even if various refinements and complications need to be added.

I agree that nothing said above rules out the possibility that a more plausible version of the BSA might be developed around reconceptualizations of simplicity, strength, and how they trade off that better fit scientific practice. (Needless to say, it would be desirable to actually produce these reconceptualizations, rather than merely insisting they must be possible.) On the other hand, it is worth noting that even if there is something right about the general and vague thought that judgements about which generalizations are laws are in some way bound up with judgements of simplicity and strength, it does not automatically follow that it must be possible to use these notions (just by themselves) to specify what a law is in anything like the way that the BSA envisions. Here is an alternative possibility, which allows simplicity and strength to be connected to lawfulness, but not in the way the BSA supposes.¹⁷ Suppose, picking up on remarks on the previous page, the distinction between

laws and initial conditions (including non-lawful regularities involving initial conditions) is made at least in part, on some other, prior basis—call it *C*—that is independent of considerations having to do with balancing simplicity and strength.¹⁸ Suppose we then apply simplicity and strength to theory-choice in a way that presupposes this law/initial condition distinction made in part on the basis of *C*. In particular, we treat considerations of simplicity and strength as having a different import for claims about laws and claims about initial conditions: it is part of good scientific methodology to attach a great deal of value to simplicity in laws but, at least in many cases, simplicity in initial conditions is regarded as counting for less. Instead, if the world is such that a theory that is non-simple in some respects is required, we try to off-load whatever complexity is required onto the initial conditions, if by doing so we can keep the laws simple.¹⁹ (Keeping both simple is perhaps even better, but this may not always be possible.) With respect to strength, one important component of the kind of strength that laws are expected to possess is that we should be able to combine them with different assumptions about initial conditions to deduce a wide range of different (actually occurring) phenomena. On the other hand, perhaps strength in statements of initial conditions is, in itself, no virtue (or vice) at all; what matters instead is whether whatever detail is in the initial conditions can be combined with assumptions about laws to deduce a range of different phenomena. Detail in initial conditions that cannot be so used may be regarded as pointless, at least in many areas of science. (Consider a list of exact descriptions of the dimensions and mass of every grain of sand on a beach.) On the assumptions just described, assessments of simplicity and strength still interact with the law/initial conditions distinction (so that it will seem there is ‘something right’ about the project of

¹⁷ What follows is not just the observation that there may be more to the notion of law than is captured by the notions of simplicity and strength, a point that has been made by many authors (for example, [van Fraassen \[1989\]](#)). My point is rather that simplicity and strength may have a different relevance or significance for laws and for initial conditions—for example, simplicity in connection with laws may be different from simplicity in connection with initial conditions or may be differently valued. To the extent this is so, one has to already know whether one is dealing with a law-claim or an initial condition claim before one can discuss how these are connected with the notion of simplicity.

¹⁸ One candidate for this independent, prior basis for distinguishing laws from claims about initial conditions can be found in the notion of invariance. See [Woodward \(\[2013\]\)](#) for details.

¹⁹ Eugene Wigner invokes essentially this idea in his ([1970], p. 41ff.), and somewhat similar suggestion is made in [Hall \(\[forthcoming\]\)](#). Another possibility, also gestured at by [Wigner \(\[1970\], p.41ff\)](#) is that simplicity in initial conditions as well as in laws is highly valued but simplicity is understood differently depending on whether it is applied to laws or initial conditions. For example, randomly distributed or relatively structureless initial conditions (as opposed to those exhibiting some highly organized pattern) are often regarded as simple, where a corresponding assumption about the simplicity of laws would be regarded as misguided. (To get a sense of some of the subtleties surrounding the notion of simplicity, note that the specification of a particular set of random-looking initial conditions involves a great deal of information; informational content is often regarded as inversely related to simplicity, but nonetheless, random-looking initial conditions are often counted as simple.)

connecting lawfulness to simplicity and strength) but not necessarily in the way envisaged in the BSA. Among other things, on the assumptions just described, we apply considerations of simplicity and strength to a pre-existing distinction between laws and initial conditions, established in part on the basis of *C*, rather than using simplicity and strength to characterize this distinction.

5 Two Roles for Simplicity

I turn now to some other issues having to do with the role of simplicity in the BSA. In particular, I want to explore the relationship between simplicity as it figures in the BSA and accounts of simplicity of the sort found elsewhere in the philosophical literature.²⁰ Much of the existing literature (and certainly the literature that is most illuminating) in philosophy, machine learning, and statistics adopts a means/ends strategy for understanding simplicity: one identifies some goal of inquiry such as finding hypotheses that are true, or have high posterior probability, or are predictively accurate (to mention just a few possibilities) and then argues (in some cases proves) that picking the simplest hypothesis according to some specified criterion of simplicity is a good means for achieving this end. In other words, we both characterize simplicity

²⁰ Two further points about the argument that follows, which I relegate to a footnote so as not to disrupt the flow of my discussion: First, a number of writers have complained that the notion of simplicity that figures in the BSA is unclear and requires further elucidation. In what follows, I do not just make this general complaint. Rather, I attempt to explore in detail several different notions of, or rationales for, the use of simplicity, connected to the distinction between descriptive and inductive simplicity, and to different accounts of curve-fitting such as Bayesian and AIC-based accounts. I attempt to show that none of these accounts yields a treatment of simplicity that captures the role it is assigned in the BSA. This discussion is meant to raise in a pointed way the issue of what the positive story is about the role played by simplicity in the BSA, and how this role relates to other roles for simplicity discussed in the philosophical and statistical literature. Second, one of the referees comments that because simplicity in the BSA is meant to track lawfulness rather than truth, any discussion of the role of simplicity in connection with traditional curve-fitting problems (which have to do with competing contenders for, for example, the true curve or the most predictively accurate one) is irrelevant to the BSA and hence unnecessary. I do not know to what extent this assessment is shared by other defenders of the BSA but, to the extent that it is, this seems to me to simply illustrate how unclear the role or meaning of simplicity is in the BSA. The philosophical and statistical literature contains worked-out theories of simplicity that attempt to connect that notion to concepts like truth, predictive accuracy, and high posterior probability. Curve-fitting and the other contexts considered below are standard contexts in which such appeals to simplicity are made. Moreover, one of the central motivations for the BSA is that it appeals to a notion of simplicity that is widely used in science, which suggests the appropriateness of exploring the connection with curve-fitting and other standard uses of simplicity. If, as a law-tracking notion, simplicity in the BSA is completely distinct from all of this, what, positively, does it involve? If the claim is that the law-tracking notion of (or role for) simplicity in science is *sui generis* and completely independent of all other roles for simplicity, what reason is there to believe there is such a notion of simplicity? If the law-tracking notion is not independent of these other notions/roles, how is it related to them?

and attempt to justify the preference for simple hypotheses by showing that a preference for simpler hypotheses, so characterized, conduces to the goal in question. A paradigm of this sort of means/ends approach is the treatment of simplicity in the Akaike information criterion (AIC) framework, discussed in Section 6. Here one attempts to justify the use of a criterion for hypothesis choice that can be interpreted as trading off a particular measure of simplicity (number of free parameters) against a measure of fit or strength on the grounds that (roughly) adoption of this criterion will maximize predictive accuracy (measured in a certain way) regarding future data. Thus, the use of this particular criterion for simplicity in hypothesis choice is justified in terms of the goal of maximizing predictive accuracy. Many other treatments of simplicity, such as the use of the so-called Bayesian Ockham's razor (also Section 6), proceed in a broadly similar means/ends way, although here the goal is finding hypotheses with a high posterior probability. Yet another possibility, which also uses a means/ends strategy of justification, is to attempt to justify a preference for simplicity by appeal to considerations having to do with efficiency of search—for example, by showing that beginning with the simplest hypothesis minimizes expected number of retractions in a worst-case scenario, as in the formal learning literature (Kelly [2011]). One of the many attractions of these means/ends strategies is that they force one to be precise about what is meant by 'simplicity' and how it connects to various goals of inquiry. Within the approaches described above, 'simplicity' is not, as is in standard formulations of the BSA, treated as an unanalysed primitive notion that is postulated to be valuable in hypothesis choice for reasons that are not further explained (other than that scientists seem to value simplicity).

In what follows, I attempt to use these various alternative treatments of simplicity to explore the role assigned to simplicity in the BSA and how might this role be connected to the identification of laws. For example, one obvious question is whether there are means/ends arguments paralleling those elsewhere in the literature on simplicity that might be used to link simplicity (balanced against strength) to laws. Another is whether any of the arguments in the literature that are used to justify the adoption of a particular trade-off between simplicity and other values, such as fit (for example, those supporting the use of the AIC in hypothesis choice), transfer to the BSA context.

Lewis and many of the other philosophers who have defended the BSA have little or nothing to say about the connection between simplicity, as it figures in the BSA, and other rationales for the use of simplicity in science. They also have little that is illuminating to say about why we should prefer simpler hypotheses, in the sense of providing some argument that such a preference leads to the successful identification of laws or successful

achievement of some other scientific purpose. In other words, they do not provide anything like a means/end justification for reliance on simplicity of the sort described above. Instead, defenses of the BSA tend to treat the preference for simplicity as well as the notion of simplicity itself as a kind of primitive, and build these into the characterization of laws (laws just are generalizations that figure in systemizations best combining simplicity and strength) so that questions about the relationship between simplicity and lawfulness, and why we should value simplicity are not given non-trivial answers. It would be much more satisfying if defenders of the BSA were able to provide a characterization of the notion of simplicity on which they are relying that relates this to notions of simplicity used elsewhere in the literature, and also if they were to provide some reason to suppose that trading this notion off against strength tracks features that we antecedently think are possessed by laws. It is against this background that my discussion will proceed.

Turning now to a more detailed look at simplicity, recall, as a point of departure, Hans Reichenbach's well-known distinction between descriptive and inductive simplicity (Reichenbach [1958]). For our purposes, we may think of descriptive simplicity as a consideration that has to do with choosing between representations or hypotheses that are 'equivalent' with respect to some hypothesis/world-connection such as truth, but where one of these hypotheses—the descriptively simpler one—is more economical or easier to comprehend or use. For example, a hypothesis expressed in polar coordinates may be easier to work with for some purposes than a mathematically equivalent hypothesis expressed in Cartesian coordinates, and a hypothesis that some relationship is linear may be regarded as descriptively simpler for many purposes when the hypothesis is expressed as a functional relationship involving a polynomial rather than a Fourier sum. By contrast, inductive simplicity, in Reichenbach's characterization, concerns simplicity as a ground for choosing among non-equivalent hypotheses, at most one of which is true. For example, one is appealing to an inductive notion of simplicity if one claims that, given several data points that lie on a straight line, we should prefer the hypothesis that the true relationship is linear to the hypothesis that the relationship is quadratic of form ' $y = ax^2 + bx + c$ ' with $a \neq 0$, $b \neq 0$, on the grounds that the former hypothesis is 'simpler' and for this reason more likely to be true. In Reichenbach's formulation, inductive simplicity has to do with the choice between inequivalent hypotheses at most one of which is true. But in what follows, I will broaden the notion of inductive simplicity to include various other possibilities, such as non-equivalence with respect to predictive accuracy (where it is assumed that all of the competing hypotheses are literally false, but some are more predictively accurate than others).

6 Simplicity in the Best Systems Account: Curve-fitting

Do either of these two ways of thinking about simplicity—descriptive and inductive—correspond to the role played by this notion in the BSA?²¹ To explore this question, consider a simple version of a traditional curve-fitting problem. Suppose a researcher is interested in the relationship between two variables, x and y , and believes, correctly, that y is some unknown function $y = f(x)$ of x alone. Suppose also that the researcher is able to measure both x and y without error, but so far has observed only a limited number of data points, all of which happen to lie on a straight line. The researcher also believes all of her/his data are generated by some single underlying mechanism or structure that operates in accordance with the relationship she is trying to discover, where ‘underlying mechanism’ carries with it the implication that any additional data will continue to be generated by this mechanism, in accord with $y = f(x)$. The researcher is faced with the problem of which curve to fit to the data: a linear relation, a quadratic or higher order curve, some complex sum of trigonometric functions, or what? Many writers suggest, on a variety of different grounds, that in this sort of case, when the data so far conform to a linear relationship and there is no relevant countervailing information, one should select a linear relationship to fit the data, and one should do so on the grounds that linear relationships are ‘simpler’ than relationships described by other functions such as higher-degree polynomials. (Recall Lewis’s remark to this effect, quoted in Section 2.) Typically, this is accompanied by the advice that if additional data are inconsistent with a linear relationship, one should then select the next lowest degree polynomial that fits the data well: first try a quadratic relationship, then a cubic if new data require this, and so on. In what follows, my interest is not primarily in evaluating whether this is good advice, but rather in drawing attention to some of the distinctive features that are present when one is faced with a problem of this kind and advocates this strategy.

First, we have assumed the case to be one in which the researcher has available, at the time she fits the curve, only some of the data required to determine which curve is correct and not all possible relevant data. Relatedly, the different curves among which the researcher chooses are thought of as alternative and non-equivalent empirical hypotheses about where any additional data points, should they be produced, would lie. A second researcher who instead fits a quadratic relationship to the so-far-observed data (with a non-zero coefficient on the quadratic term, of course) is thus making predictions about what additional data will look like that are different from the prediction made by the first researcher. Thus, at least in principle, there are

²¹ To avoid some irrelevant complications, I make no claims about there being two different concepts of simplicity and speak instead just of there being different roles for simplicity.

possible additional data such that, if it should occur, could be used to distinguish between these two proposed curves. The notion or use of simplicity that is used to guide the choice between these two curves is thus (broadly) inductive rather than descriptive, in Reichenbach's terminology.

What sort of considerations might be used to support a preference for a linear relationship in this case? Although nothing in what follows depends on the adoption of any particular rationale for curve-fitting preferences, and although some philosophers, including perhaps Lewis,²² may hold a different view, most if not all of the standard treatments of simplicity in the literature appeal (at least in part) to broadly empirical considerations of one sort or another.²³ One such set of empirical considerations, illustrated by the Bayesian treatment of curve-fitting described below, is information about the frequency distribution of particular sorts of systems among systems of the general kind *K* under investigation—for example, if linear systems are particularly common among systems of kind *K*, this can help to justify giving them preference in a simplicity ordering. As another illustration, attempts in formal learning theory to justify simplicity in terms of the efficient pursuit of truth (minimizing number of retractions in the worse case) also rest on empirical assumptions, although different and weaker ones.²⁴ Indeed, given that, in the sort of case under discussion, simplicity considerations are being used to guide choice among empirically non-equivalent hypotheses, it is hard to see how they could justifiably do so unless those simplicity

²² As I note, both Lewis and many of the other philosophers who defend the BSA have little to say about why, or on what basis, it is justifiable to prefer simpler hypotheses, *ceteris paribus*. For this reason, it is difficult to say whether or to what extent they think of simplicity preferences as resting on empirical assumptions. One possible interpretation of Lewis is that the choice of a 'natural' language for making simplicity comparisons will rest on or encode empirical assumptions, but as far as I know Lewis does not explicitly discuss this issue.

²³ For more general arguments that when simplicity considerations in science are used to justify choice among empirically non-equivalent hypotheses, these considerations must rest (at least in part) on broadly empirical considerations; see (Sober [1988]) and (Norton [manuscript]). As a specific illustration of this general point, Sober argues that while appeals to simplicity considerations play a role in phylogenetic inference, the considerations in question rest on domain-specific empirical assumptions about the processes underlying speciation, rather than on purely formal considerations that apply globally to all domains. Woodward ([2003], [2013]) makes similar claims about appeals to simplicity in causal inference. Of course, different appeals to simplicity will rest on empirical assumptions that may be either strong and specific or, in other cases, comparatively weak and generic.

²⁴ See the extremely interesting series of articles by Kevin Kelly [2011] (for example), which attempt to justify a preference for simplicity by appealing to the idea that (roughly) this minimizes the number of retractions in worst-case scenarios for specified complexity classes. The particular simplicity ordering that is justified by this line of argument depends (among other considerations) on the information states available to the inquirer, which is an empirical matter (Kelly, personal communication). I will add that even if it were true that this learning-theoretic justification for simplicity in terms of efficient pursuit of truth has no empirical presuppositions, this framework seems in any event to be inapplicable in the context of the BSA because (as noted below) one already has all possible data and issues about retracting as new data arrives do not arise.

considerations themselves are understood as carrying empirical commitments of some kind.

To briefly develop the Bayesian possibility mentioned in the preceding paragraph, suppose (cf. MacKay [2003], Henderson *et al* [2010]) a researcher has reason to believe that, considering families of polynomials of every degree, the domain under investigation is such that each family is equally likely to contain the true curve and that within each family the probability density for which particular curve is correct is spread in a fairly uniform way over each member of the family. The effect of these assumptions is that close-by individual curves of higher degree receive diminishing probability mass because they must share it with more and more competitors; families of close-by linear relationships are regarded as more probable than families of close-by quadratic relationships, and so on. Updating via Bayes's theorem in this 'hierarchical Bayesian' framework, the researcher will initially (when only limited data are available) assign a higher posterior to linear relationships, although as additional data come in, this may lead to higher posteriors for higher-order relationships. In effect, the researcher is adopting a simplicity ordering over her/his hypothesis space, which introduces an inductive bias into her/his search for the best-fitting curve, an ordering that tells her/him to first consider linear hypotheses, move on to quadratic hypotheses if and only if data come in that force this move, and so on. The choice of the framework of families of polynomials, so that a curve receives lower probability if it is represented by a higher-degree polynomial, encodes or reflects this particular ordering. Within this polynomial framework, curves with more free parameters (that is, when the curves are represented by polynomials) receive lower prior probabilities—an effect described as the Bayesian Ockham's razor (Henderson *et al.* [2010]). As noted above, if linear relationships are extremely common for systems of that sort and higher-degree polynomial relationships less common, this would provide one obvious way of justifying the adoption of such an ordering. If, on the contrary, the systems under investigation frequently exhibited various forms of periodic motion, then it might be more appropriate to adopt a 'Fourier framework' in which hypotheses like $y = \sin(x)$ are regarded as simplest and functions represented by sums of products of trigonometric functions as more complex.

Note that so far nothing in these curve-fitting examples warrants talk of a 'trade-off' between simplicity and strength, as this is understood in the BSA. We have assumed that the data are measured without error and the task is to find the true curve that fits these data exactly. Simplicity considerations establish an ordering among the hypotheses we consider (for example, adopting a polynomial representations means 'try linear relationships first'), with hypothesis choice being adjusted as necessary as additional data come in. But once a curve fails to fit the data, it is disqualified and we move on to the next

simplest curve. We do not retain the simpler curve on the grounds that even though it does not fit the data exactly, this is made up for by its greater simplicity. This role for simplicity looks (with an important qualification) more like the threshold view described in Section 2 than the trade-off described in the BSA.²⁵

In addition to this consideration, there are other important disanalogies between the role assigned to simplicity in the BSA and the role assigned in the curve-fitting problem described above. In the latter case, simplicity functioned to guide hypothesis choice in a context in which not all of the data were in. But within the BSA, all of the data that will ever be available are in because everything that happens everywhere and for all time is represented in the HB, which we are attempting to systematize. Unlike the curve-fitting examples, we cannot in this case think of ourselves as using simplicity to guide choices among alternative curves that make empirically inconsistent predictions about what will be observed, in the sense that one set of predictions about what is in the HB will turn out to be true and the other false. This is because only curves that make true claims about the HB are even candidates for the best systemization; thus, different candidates for the best systemization cannot imply claims about the HB that are inconsistent with each other. This in itself seems to make the use of simplicity in the context of the BSA different from the use in ordinary curve-fitting problems of the sort so far considered. Indeed, insofar as the rationale for choosing the simplest curve is based on the idea that this is more likely than competitors to yield true or correct predictions about future data, this rationale seems otiose in the context of the BSA. Why bother with this rationale for simplicity if all of the data that will ever occur are already incorporated in the HB?²⁶

It might be argued in response that this dismissal of a role for simplicity is too quick and that there is an important respect in which alternative candidates for the best systemization can be thought of as making inconsistent predictions in the context of the BSA.²⁷ In particular, different systemizations

²⁵ The qualification is that in the standard treatment of the curve-fitting problem without error, it is assumed that data will eventually come in that will be inconsistent with all curves but one. By contrast, as noted above, in the BSA we choose among candidate systemizations all of which are consistent with the HB, although some of these systemizations will fail to imply data in the HB that are implied by other systemizations. Here, a threshold view is one according to which a theory that fails to imply data cannot make up for this by being simpler.

²⁶ On most of the standard treatments of simplicity in curve-fitting problems such as Bayesian accounts or the Akaike framework discussed below, the term in the criterion for hypothesis choice that reflects the influence of simplicity (the prior in the Bayesian framework or the number of parameters term in the Akaike framework) becomes less important as more data come in, and the term that reflects fit to the data (the likelihood terms) or strength becomes more important. Arguably, this is normatively appropriate: as more data come in, it should have a greater influence on the choice of the correct or most predictively accurate curve. Simplicity in the context of the BSA does not behave in this way; it continues to be important even when all of the data are in.

²⁷ This response was suggested by a referee in comments on an earlier draft.

can be thought of as non-equivalent in the sense that they make different predictions about what would be observed were various contrary-to-fact conditions realized. For example, if generalization G is a law according to systemization S_1 and a non-lawful regularity according to systemization S_2 , then (it might be thought) S_2 predicts that there are non-actual circumstances in which G would be false (call this prediction P), whereas S_1 denies this. It might be urged that this gives us a sense in which choosing between two different systemizations of the HB is analogous to an ordinary curve-fitting problem in which we choose between two competing hypotheses, with simplicity playing essentially the same role in both cases.

I think that, on reflection, this line of argument is unpersuasive. The way in which the two ‘predictions’, P and not- P , described above, are in conflict differs in a fundamental way from the way in which two non-equivalent curves making different predictions in an ordinary curve-fitting problem are in conflict. In contrast to an ordinary curve-fitting problem, in the context of the BSA we know in advance that no evidence will ever show that one of P or not- P is correct and the other mistaken.²⁸ Indeed, with respect to the different predictions they make, each of the systemizations, S_1 and S_2 , is self-certifying, in the sense that S_1 judges according to its own lights that violation of G is impossible, and S_2 judges that such violation is possible, again according to its own lights. The real content of any disagreement between an adherent of S_1 and an adherent of S_2 is simply that they disagree about which of S_1 and S_2 is the best systemization of the data in the HB; any disagreement they have about P is completely derivative from this.

One consequence of this is that the kind of means/ends justification that is available for choosing the simplest curve in the curve-fitting context (that the simplest curve is more likely to be true or predictively accurate) cannot be used (at least in any obvious way) to provide a rationale for the use of simplicity in the context of the BSA. That is, there is no basis for an argument that might be used to show that the systemization that best balances simplicity and strength according to BSA standards is more likely to make true or accurate predictions regarding counterfactual claims like P , because the truth about P is not an independently accessible fact against which we can compare different systemizations. Put differently, within the BSA framework it is the best systemization that tells us which are the true counterfactuals, rather than the true counterfactuals acting as an independent constraint on which is the best systemization. This is a consequence of the fact that, according to the BSA, there is no independent nomological order or independent facts about counterfactual dependence (apart from whatever is implied by the best systemization of

²⁸ That is, ‘no evidence’ in any ordinary or usual sense of ‘evidence’. Any ‘evidence’ that supports P or not- P just consists of whatever considerations support the claim that S_1 or S_2 is the better systemization of the evidence in the HB.

the HB) against which competing systemizations may be judged. This contrasts sharply with the situation in which criteria for curve-fitting like the Bayesian Ockham's razor or AIC are applicable. Here there are independent facts about which predictions about future data are correct or more accurate, and arguments which appeal to such facts to motivate specific conceptions of simplicity.

7 Simplicity as a Corrective for Overfitting

It might seem that we can get closer to something that looks like a justification for a BSA-style trade-off involving simplicity and strength if we modify the curve-fitting example so that the observed data are measured with error; or reflect noise in addition to whatever it is that we are interested in measuring; and/or the data are conceptualized as a sample from some underlying population that is really what is of interest, but where any individual sample may not be fully representative of this underlying population because of 'chance' effects. These assumptions introduce the possibility of a kind of 'overfitting'—a possibility that is not present when we imagine that the data are measured without error—and this in turn points to a role for simplicity considerations in curve-fitting about which there is an extensive literature: simplicity as a corrective for overfitting. To illustrate this idea, consider a context in which the true relationship is linear but, because the values of x and y are measured with error or reflect sampling variability due to chance effects, what we in fact observe is a scatter of data points that do not exactly fall on a straight line. In this case, if we try to fit a higher-degree polynomial to this data, estimating the coefficients of the various terms from the data, we may be able to achieve a better fit to the data so far observed than is provided by the linear relationship simply because there are more free parameters in the higher-degree polynomial. However, in doing this we will be (in part) fitting noise or sample variability due to chance effects, and this will show up in the fact that when we attempt to use the higher degree polynomial to predict new data, it will predict less well than the linear relationship because the exact noise distribution will be different with new data.

One common solution to this problem is to assess candidate curves by some measure that reflects both fit to the data so far received and that discounts for the number of free parameters in the fitted curve. A well-known example is the AIC, mentioned above, which measures fit by means of the log of the likelihood function, $\ln(L)$, for the data given the fitted curve and then discounts this by the number of free parameters, k : $\text{AIC} = 2k - 2\ln(L)$. If we think of fit to the data so far observed as connected to strength and the number of free parameters as inversely associated with simplicity, it might seem that in this

case we have a ‘trade-off’ between simplicity and something that looks like strength, and that this resembles the trade-off in the BSA.

On closer examination, however, this analogy with the BSA looks highly dubious. First, the difficulty identified at the end of Section 6 continues to apply: the respect in which different candidates for the best systemization of the HB make different ‘predictions’ (about, say, what would happen in counterfactual circumstances that will never be realized) is fundamentally different from the way in which different curves make different predictions in the context in which AIC is used. In the context of the BSA, one is not, as with AIC, attempting to predict future data on the basis of a body of data that is incomplete in the sense that this is only a portion of the data that eventually will be seen. In addition, it is unclear how to make sense of the notion of measurement error or non-representative sampling in the context of the HB because the HB is understood as recording, without error, all that happens. Because of this, it is hard to see how the particular rationale that motivates the use of AIC—the need to avoid overfitting—applies in the case of the BSA. In the case of ordinary curve-fitting on noisy and incomplete data, it is clear how overfitting leads to less empirically accurate predictions on new data. Again, it is unclear what is analogous to this in the case of the BSA.

We can provide further support for this assessment by considering the justification for using AIC in ordinary curve-fitting problems in a bit more detail. Although AIC is sometimes described (as I have above) as involving a trade-off between simplicity and fit (or strength), it is important to be clear about just what this amounts to. In particular, in the derivation of AIC, there is no adoption of a postulate or criterion for simplicity as a primitive assumption that is then traded off on some intuitive basis against strength. Instead, one proceeds by assuming that the data to be fitted are generated by some unknown relationship (the true curve) that is contained within the family of curves to be fitted. Because the data contain noise or error, any curve, C , fitted to them will almost certainly differ from the true curve, T , and will be at least somewhat predictively inaccurate on future data. However, given certain assumptions, one can estimate how far various candidates for the fitted curve may be expected to differ on average from the true curve, as indicated by a particular measure (the Kullback–Leibler or K-L divergence) of the ‘distance’ between C and T . One can then choose the fitted curve that minimizes this distance. (Note that the fitted curve that is chosen is expected to be literally false, but is expected to perform better than alternatives on future predictions. This is another difference with the BSA, where it is assumed that the axioms make only literally true predictions.) It turns out that the uncorrected maximum likelihood estimator, captured just by $\ln(L)$, yields a biased estimate for the K-L divergence; but an asymptotically unbiased estimator of the K-L divergence is provided by correcting the $\ln(L)$ term by the factor $2k$, and this is

taken to justify the use of AIC. Thus, the presence of the term involving k in the formula for AIC arises because the correction provided by this term leads to a better estimator for the expected predictive accuracy of the fitted curve than just the $\ln(L)$ term itself. Put more directly, the simplicity term does not enter into the formula for AIC because simplicity is being accorded some intrinsic value that trades off against strength (understood as minimizing expected predictive inaccuracy), but rather because the goal of maximizing expected predictive accuracy in itself requires inclusion of this term. In other words, the justification for the simplicity term is entirely in terms of the fact that it is a means to the (strength-like) goal of predictive accuracy. In these respects, the role of 'simplicity' in AIC and the justification for its use is different from the role it is assigned in the BSA.

8 Descriptive Simplicity in the Best Systems Account?

So far I have focussed on the use of simplicity to choose among hypotheses that are in some sense empirically non-equivalent—that is, on inductive simplicity in a broad sense, having to do with hypotheses making different claims about what is true or what predictions about not yet seen data will be most accurate. What about descriptive simplicity? Might this provide us with a better model of how simplicity figures in the BSA?

This suggestion seems to make most sense if, following Lewis, we suppose that there is a privileged language, L , in terms of which simplicity comparisons among competing generalizations are to be made. L might be chosen on the basis of naturalness and/or on the basis of descriptive simplicity-based considerations to do with ease of comprehension and use. Simplicity comparisons among different candidates for the best systemization will then concern how descriptively simple these are when expressed in L . (That the candidates might be assessed differently with respect to simplicity if equivalent systemizations were expressed in some alternative language is regarded as irrelevant because L has a privileged status.)

My discussion of this alternative will be brief because many of the points I wish to make have been made by others. First, it is hard to see how to interpret descriptive simplicity and notions like ease of use as meaning anything other than 'simple and useful for us'. If so, and if the only notion of simplicity that the BSA uses is this descriptive notion, we seem led to a notion of law that is strongly relativized to what we human beings (perhaps now, given our contingent history) find easiest to use and reason with. This seems at odds with the intent of at least many defenders of the BSA who seek a notion of law that is not so completely relativized in this way. Thus, Lewis, in a passage quoted above, says that the standards of simplicity 'are only partly a matter of psychology'. Moreover, once the relevant notion of simplicity in the BSA is taken to

be descriptive simplicity, the claim that there is some single privileged representation or language in science that is relevant to simplicity assessments looks dubious. If something like ease of use is the relevant consideration underlying simplicity judgements, why should we not expect different languages or representational schemes to be used in different contexts or for different purposes, with each being judged as ‘simplest’, depending on which is most useful for the case at hand? In fact, it looks as though this is exactly what one finds in scientific practice. Possible cases include representations of gravity in terms of metric structure in general relativity, and (arguably equivalent) representations in terms of a fixed background space-time and a separate gravitational potential; both of which are used in different contexts, and the use of both Hamiltonian and Lagrangian formalisms to represent many phenomena in mechanics. These sorts of practices seem contrary to any picture according to which there is a single privileged language or representational scheme that is used in science.

9 Simplicity as Due to Human Intellectual Limitations

Is there some alternative way of thinking about the role of simplicity within the BSA besides the two possibilities explored above? Marc Lange introduces the following dialogue to elucidate the BSA:

You: Describe the Universe please, Lord.

God: Right now there’s a particle in state ψ_1 and another particle in state ψ_2 and I’ll get to the other particles in a moment, but in exactly 150 million years and three seconds, there will be a particle in state ψ_3 and . . .

You: Lord, I have an appointment in a few minutes.

God: All right, I’ll describe the universe in the manner that is as brief and informative as it is possible simultaneously to be—by giving you the members of the ‘Best System’. ([2009], pp. 101–2)²⁹

Lange’s purpose in introducing this dialogue is to raise some questions about the immutability of laws, but another way of interpreting the dialogue is to take it as suggesting that the role of simplicity within the BSA derives from human limitations, intellectual and otherwise, and/or facts about restricted human interests. (In this respect, the motivation bears some resemblance to the role for descriptive simplicity explored in Section 8, but here the emphasis

²⁹ Several readers have drawn my attention to the fact that a similar metaphor is used by Albert ([manuscript]). In both Lange’s and Albert’s versions there is no explicit reference to human cognitive limitations. As a consequence, it is left unclear why God’s account is subject to just the constraints described—that is, why God is requested to provide an account that balances simplicity and informativeness in the manner suggested. Limitations on time and processing capacity available to humans would explain why such constraints are imposed. Without some story about why God’s account is subject to just these constraints, the need for compression seems mysterious and unmotivated.

is more on what is beyond our power to comprehend than on what is easiest for us to use and on the resulting need to omit information rather than just present it in a user-friendly format.) The idea is that a being with unlimited capacities for information processing and calculation, and unlimited time could make use of a specification of the full HB. However, this is too complex and informationally rich to be used in its entirety by human beings (and many of the details are not likely to be of interest to us anyway). Instead, to be used by us, the information in the HB must be greatly compressed. Generally speaking, compressed descriptions will be simpler and the best such description will be one that achieves an optimal trade-off between compression and informativeness. As a rough illustration, imagine that quantity x is distributed across space and time in an extremely complex way, so that a full specification of its distribution would be so informationally overwhelming for us that we would not be able to make use of it. In such a case, we might find it more useful to be given a summary, or a compressed version, of aspects of this distribution—for example, the mean value of x within some space-time region. Notice this compression or summary need not contain anything false; it leaves out information that is contained in the full specification of the HB, but it need not introduce inaccuracies because what it says about the average value of x may be exactly correct. Now add to the story that there is a second quantity, y , also distributed in a complex way, and that there is some extremely intricate function, F , describing a pattern or regularity relating x and y . F itself is so complex (that is, when assessed by reference to our information processing abilities) that even if it were described to us, we would be unable to comprehend or use it. However, it may be the case that there is some function G that relates a summary of x to a summary of y —for example, G might give the mean value of y as a function of the mean value of x . This function G might be one that we could use and comprehend because it makes fewer demands on human processing resources; in this sense it would be simpler than F . Moreover, G also need not be false or inaccurate; it may be that the average value of y is always linked to the average value of x in exactly the way claimed. The suggestion we are considering is that the BSA should be interpreted in such a way that it is scenarios of this sort that lead us to regard the regularity described by G as a law of nature.³⁰

Note that in this picture (unlike the interpretations of the BSA considered previously), a trade-off between simplicity and strength arises in a natural

³⁰ Note how this role for simplicity differs from both of the two roles distinguished earlier. F and G are not competitors and do not make conflicting empirical predictions, so simplicity in this third role is not being used to distinguish among competing alternative hypotheses in the way that the inductive use of simplicity is. Moreover, F and G or x and the average of x do not differ only with respect to their descriptive simplicity because F is strictly stronger than G and similarly for x and mean (x).

way. Roughly speaking, simplicity comes into the story as a fix for a level of strength that overwhelms our information-processing capacities; simplicity involves discarding strength and information until we get to a point at which what we are left with is tractable from the point of view of our information-processing abilities. So it is obvious why we should expect simplicity, so conceived, to trade off with strength (with the trade-off taking the form envisioned in the BSA, rather than simplicity bearing a threshold relation to strength)—where the optimal trade-off is determined by facts about our intellectual capacities. Thus, this interpretation of the BSA seems well-motivated in a way that many of the alternative interpretations considered above are not.³¹

Despite this, I again find it doubtful that advocates of the BSA will wish, on reflection, to endorse this account of the role of simplicity. (Again, if I am wrong about this, those advocates should say so and embrace the consequences for this role for simplicity.) The problem is that if anything deserves to be called a fundamental law of nature in this scenario, it is surely the regularity described by *F* rather than the regularity described by *G*. (Or, at least, this is the case if, as I assume, the goal of the BSA is to provide an account of which regularities are laws, rather than which regularities are believed by us to be laws.) In the story I have provided, the preference for *G* over *F* is dictated entirely by human processing limitations, again violating the constraint that the relevant notion of simplicity in the BSA should be one that is at least in part independent of human psychology. Moreover, human processing limitations may well alter over time with advances in technology, changes in ways of representing and manipulating information, perhaps with changes in education and the social structure of science, and so on. Suppose that changes in information-processing capacities permit us to make use of some alternative function, *H*, which is not as detailed as *F* but considerably more detailed than *G*, to represent the relationship between *x* and *y*. How should we think about *G* at this point? On the face of things, it would appear that *G* no longer represents an optimum balance of simplicity and strength; instead, it is now *H* that is the new optimum. Does this mean that *G* is no longer a law of nature, although it once was? To avoid this conclusion, should

³¹ Although I lack the space for details, something like this picture seems to capture some features emphasized in the extensive literature on modelling trade-offs in complex systems (see, for example, [Levins \[1966\]](#)). Models that attempt to capture the full details of behaviour of such systems may be so complex, both computationally and in terms of the information they require as input, so as to be completely unusable (and not even constructable) by us. So we adopt a strategy of not even trying to predict or capture many of those details in our models, and aim instead at trying to predict or explain a few large-scale patterns via a tractable model. This strategy involves a kind of trade-off between simplicity and strength that resembles the trade-off described in the BSA, but I take it that no one thinks that the generalizations figuring in such models are ‘fundamental laws’—indeed the presence of such trade-offs is taken by many writers to be an indication that we are dealing with non-fundamental science ([Gelfert \[2013\]](#)).

we perhaps say that the laws are those regularities corresponding to the axioms and theorems in a systemization that would achieve an optimal balance of simplicity and strength if it were the case that human information processing abilities were developed to their fullest possible extent? If so, how do we tell what this counterfactual requires?

10 Summary

Let me try to draw together the various strands in this discussion. The notion of simplicity appealed to in the BSA must satisfy a number of constraints. It must be ‘thick’ or substantive enough that it can serve as a basis (along with strength) for choosing among systemizations that are genuine alternatives to one another with respect to their claimed status as laws. This suggests that it must involve something more than mere descriptive simplicity because descriptive simplicity is not a criterion for choosing among non-equivalent hypotheses. On the other hand, the role of simplicity within the BSA is not the usual role of inductive simplicity in guiding us to the choice of true hypotheses over false ones, or hypotheses that are more probable or have higher than expected empirical accuracy—rationales that are often appealed to in order to explain the methodological significance of simplicity in more ordinary curve-fitting problems. This is because all of the candidates for the best systemization (as well as the HB that they purport to systematize) contain only truths. Nor can we understand the role of simplicity within the BSA in terms of the efficient pursuit of the truth, minimizing retractions, and so on. We also need a notion of strength that reflects the idea that a strong systemization is one that captures a lot of the information in the HB, but this systemization must also be one that does not permit the derivation of accidental regularities. Are there notions of simplicity and best balance that possesses all of these features and that figure in scientific practice in the way that advocates of the BSA contend? I certainly do not claim to have surveyed all possibilities, but I do suggest that those who wish to answer this question in the affirmative need to provide more details about what they have in mind by these notions.

To drive these points home, let me put them in terms of a series of questions for advocates of the BSA:

- (1) What does ‘strength’ mean in the context of the BS? Is it a function of all of the logical consequences of the BS, regardless of whether they have anything to do with what is reported in the HB? Is it a function only of the logical consequences of the BS that are (in some way) informative about the HB? Is it a function of the extent to which the BS can be used, in conjunction with some of the information in the HB, to deduce other information in the HB? Relatedly, what

information, if any, about initial and boundary conditions makes it into the ‘best systemization’? If the answer is ‘little information’, what are the implications of this for how the notion of strength should be understood?

- (2) Are there many examples in the practice of fundamental science in which simplicity and strength are traded off in the way envisioned in the BSA? Are there normative reasons, connected to generally accepted goals of science, for expecting such trade-offs to occur? If the trade-offs that occur are more like those described by the threshold or lexical priority view, what are the implications of this fact for the BSA?
- (3) What is the notion of simplicity that figures in the BSA and how does this connect to other notions of simplicity or rationales for appeals to simplicity in the scientific, statistical, and philosophical literature? Is the notion of simplicity used in the BSA linked to descriptive simplicity or to inductive simplicity? To some *sui generis* notion of (or role for) simplicity that is distinct from both of these? If so, what is this *sui generis* notion and why should we suppose that it is connected to the identification of laws in the way that the BSA claims?

Even if many of the arguments of this article are misguided, these are questions that would be desirable for advocates of the BSA to address.

11 Concluding Remarks

This article has argued that there are significant problems with the BSA as currently formulated. On the other hand, although I have not tried to argue for this view here, I share the misgivings of BSAers about the alternative metaphysical accounts that are on offer. If one is sceptical about the prospects for the BSA as well, how should one proceed? Here it is important that we not throw out the baby with the bathwater. Rejecting the various philosophical accounts of the ‘truth-makers’ for laws does not mean that we should reject the notion of a law of nature itself as incoherent or unfounded. The distinction between laws and initial conditions (including regularities involving initial conditions) may well have its limits (for example, in some cosmological contexts), but it remains central to a great deal of physical theorizing. It is connected to many other distinctions: to differences in the way that laws and initial conditions figure in physical reasoning; to the connection between laws and symmetry conditions (laws are expected to satisfy symmetry conditions that regularities in initial conditions are not); and to differences in the kinds of evidence that are required to establish that law-claims and claims about initial conditions are true. The alternative to trying to provide a

metaphysics for laws is a more descriptive or interpretive project that attempts to describe the notion or notions of law at work in different areas of science, and how these notions figure in scientific reasoning and are connected to evidence and to other aspects of scientific practice. It would involve, for example, empirical investigations of the various considerations in virtue of which scientists actually infer the existence of laws—investigations that take seriously the possibility that some of these considerations will be rather different than those on which the BSA focuses. In my opinion, this is a project that has some chance of success, while attempts to provide a metaphysical foundation for laws—whether of the BSA variety or any other—seem (so far at least) to have produced little that is useful or illuminating.

Acknowledgements

Thanks to Alexander Reutlinger and several anonymous referees for helpful and stimulating comments on earlier drafts.

*Department of History and Philosophy of Science
University of Pittsburgh, USA
jfw@pitt.edu*

References

- Albert, D. [2000]: *Time and Chance*, Cambridge, MA: Harvard University Press.
- Albert, D. [manuscript]: 'Physics and Chance', <http://philoscology.rutgers.edu/images/uploads/chapter_1.rtf>, accessed 18 September 2012.
- Armstrong, D. [1983]: *What Is a Law of Nature?* Cambridge: Cambridge University Press.
- Bird, A. [2007]: *Nature's Metaphysics: Laws and Properties*, Oxford: Oxford University Press.
- Dretske, F. [1977]: 'Laws of Nature', *Philosophy of Science*, **44**, pp. 248–68.
- Einstein, A. [1934]: 'On the Method of Theoretical Physics', *Philosophy of Science*, **1**, pp. 163–9.
- Gelfert, A. [2013]: 'Strategies of Model-Building in Condensed Matter Physics: Trade-Offs as a Demarcation Criterion between Physics and Biology?', *Synthese*, **190**, pp. 253–72.
- Hall, N. [forthcoming]: 'Humean Reductionism about Laws of Nature', <philpapers.org/rec/HALHRA>, accessed 20 August 2011.
- Henderson, L., Goodman, N., Tenenbaum, J. and Woodward, J. [2010]: 'The Structure and Dynamics of Scientific Theories: A Hierarchical Bayesian Perspective', *Philosophy of Science*, **77**, pp. 172–200.
- Kelly, K. [2011]: 'Simplicity, Truth, and Probability', in P. Bandyopadhyay and M. Forster (eds), *Philosophy of Statistics*, Amsterdam: Elsevier.
- Lange, M. [2000]: *Natural Laws in Scientific Practice*, Oxford: Oxford University Press.

- Levins, R. [1966]: 'The Strategy of Model Building in Population Biology', *American Scientist*, **54**, pp. 421–31.
- Lewis, D. [1980]: 'A Subjectivist's Guide to Objective Chance', in his *Philosophical Papers*, Volume 2, pp. 83–113.
- Lewis, D. [1983]: 'New Work for a Theory of Universals', *Australasian Journal of Philosophy*, **61**, pp. 343–77.
- Lewis, D. [1999]: *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press.
- Loewer, B. [2009]: 'Why Is There Anything but Physics?', *Synthese*, **170**, pp. 217–33.
- Maudlin, T. [2007]: *The Metaphysics within Physics*, Oxford: Oxford University Press.
- Newton, I. [1968]: 'The Rules of Reasoning in Philosophy', in A. Motte (trans.), *The Mathematical Principles of Natural Philosophy*, Volume 2, London: Dawson, pp. 202–5.
- Norton, J. [manuscript]: 'Simplicity in Model Selection'.
- Reichenbach, H. [1958]: *The Philosophy of Space and Time*, New York: Dover.
- Roberts, J. [2008]: *The Law-Governed Universe*, Oxford: Oxford University Press.
- Sober, E. [1988]: *Reconstructing the Past*, Cambridge, MA: MIT Press.
- Tooley, M. [1987]: *Causation: A Realist Approach*, Oxford: Oxford University Press.
- Van Fraassen, B. [1989]: *Laws and Symmetry*, Oxford: Oxford University Press.
- Wigner, E. [1970]: *Symmetries and Reflections*, Cambridge, MA: MIT Press.
- Woodward, J. [2003]: *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.
- Woodward, J. [2013]: 'Laws, Causes, and Invariance', in S. Mumford and M. Tugby (eds), *Metaphysics and Science*, Oxford: Oxford University Press.