**STANFORD ENCYCLOPEDIA OF PHILOSOPHY**

# Simplicity

*First published Fri Oct 29, 2004; substantive revision Thu Feb 25, 2010*

Most philosophers believe that, other things being equal, simpler theories are better. But what exactly does theoretical simplicity amount to? Syntactic simplicity, or elegance, measures the number and conciseness of the theory's basic principles. Ontological simplicity, or parsimony, measures the number of kinds of entities postulated by the theory. One issue concerns how these two forms of simplicity relate to one another. There is also an issue concerning the justification of principles, such as Occam's Razor, which favor simple theories. The history of philosophy has seen many approaches to defending Occam's Razor, from the theological justifications of the Early Modern period, to contemporary justifications employing results from probability theory and statistics.

# 1. Introduction

There is a widespread philosophical presumption that simplicity is a theoretical virtue. This presumption that simpler theories are preferable appears in many guises. Often it remains implicit; sometimes it is invoked as a primitive, self-evident proposition; other times it is elevated to the status of a 'Principle' and labeled as such (for example, the 'Principle of Parsimony'). However, it is perhaps best known by the name 'Occam's (or Ockham's) Razor.' Simplicity principles have been proposed in various forms by theologians, philosophers, and scientists, from ancient through medieval to modern times. Thus Aristotle writes in his *Posterior Analytics*,

> We may assume the superiority *ceteris paribus* of the demonstration which derives from fewer postulates or hypotheses.[1]

Moving to the medieval period, Aquinas writes

> If a thing can be done adequately by means of one, it is superfluous to do it by means of several; for we observe that nature does not employ two instruments where one suffices (Aquinas 1945, p. 129).

Kant—in the *Critique of Pure Reason*—supports the maxim that "rudiments or principles must not be unnecessarily multiplied (*entia praeter necessitatem non esse multiplicanda*)" and argues that this is a regulative idea of pure reason which underlies scientists' theorizing about nature (Kant 1950, pp. 538–9). Both Galileo and Newton accepted versions of Occam's Razor. Indeed Newton includes a principle of parsimony as one of his three 'Rules of Reasoning in Philosophy' at the beginning of Book III of *Principia Mathematica*.

> Rule I: We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearances.

Newton goes on to remark that "Nature is pleased with simplicity, and affects not the pomp of superfluous causes" (Newton 1972, p. 398). Galileo, in the course of making a detailed comparison of the Ptolemaic and Copernican models of the solar system, maintains that "Nature does not multiply things unnecessarily; that she makes use of the easiest and simplest means for producing her effects; that she does nothing in vain, and the like" (Galileo 1962, p. 397). Nor are scientific advocates of simplicity principles restricted to the ranks of physicists and astronomers. Here is the chemist Lavoisier writing in the late 18th Century

> If all of chemistry can be explained in a satisfactory manner without the help of phlogiston, that is enough to render it infinitely likely that the principle does not exist, that it is a hypothetical substance, a gratuitous supposition. It is, after all, a principle of logic not to multiply entities unnecessarily (Lavoisier 1862, pp. 623–4).

Compare this to the following passage from Einstein, writing 150 years later.

> [T]he grand aim of all science…is to cover the greatest possible number of empirical facts by logical deductions from the smallest possible number of hypotheses or axioms (Einstein, quoted in Nash 1963, p. 173).

Editors of a recent volume on simplicity sent out surveys to 25 recent Nobel laureates in economics. Almost all replied that simplicity played a role in their research, and that simplicity is a desirable feature of economic theories (Zellner et al. 2001, p.2).

Within philosophy, Occam's Razor (OR) is often wielded against metaphysical theories which involve allegedly superfluous ontological apparatus. Thus materialists about the mind may use OR against dualism, on the grounds that dualism postulates an extra ontological category for mental phenomena. Similarly, nominalists about abstract objects may use OR against their platonist opponents, taking them to task for committing to an uncountably vast realm of abstract mathematical entities. The aim of appeals to simplicity in such contexts seem to be more about shifting the burden of proof, and less about refuting the less simple theory outright.

The philosophical issues surrounding the notion of simplicity are numerous and somewhat tangled. The topic has been studied in piecemeal fashion by scientists, philosophers, and statisticians. The apparent familiarity of the notion of simplicity means that it is often left unanalyzed, while its vagueness and multiplicity of meanings contributes to the challenge of pinning the notion down precisely.[2] A distinction is often made between two fundamentally distinct senses of simplicity: syntactic simplicity (roughly, the number and complexity of hypotheses), and ontological simplicity (roughly, the number and complexity of things postulated).[3] These two facets of simplicity are often referred to as *elegance* and *parsimony* respectively. For the purposes of the present overview we shall follow this usage and reserve 'parsimony' specifically for simplicity in the ontological sense. It should be noted, however, that the terms 'parsimony' and 'simplicity' are used virtually interchangeably in much of the philosophical literature.

Philosophical interest in these two notions of simplicity may be organized around answers to three basic questions;

(i) How is simplicity to be defined? **[Definition]**

(ii) What is the role of simplicity principles in different areas of inquiry? **[Usage]**

(iii) Is there a rational justification for such simplicity principles? **[Justification]**

As we shall see, answering the definitional question, (i), is more straightforward for parsimony than for elegance. Conversely, more progress on the issue, (iii), of rational justification has been made for elegance than for parsimony. It should also be noted that the above questions can be raised for simplicity principles both within philosophy itself and in application to other areas of theorizing, especially empirical science.

With respect to question (ii), there is an important distinction to be made between two sorts of simplicity principle. Occam's Razor may be formulated as an *epistemic* principle: if theory $T$ is simpler than theory $T^*$, then it is rational (other things being equal) to believe $T$ rather than $T^*$. Or it may be formulated as a *methodological* principle: if $T$ is simpler than $T^*$ then it is rational to adopt $T$ as one's working theory for scientific purposes. These two conceptions of Occam's Razor require different sorts of justification in answer to question (iii).

In analyzing simplicity, it can be difficult to keep its two facets—elegance and parsimony—apart. Principles such as Occam's Razor are frequently stated in a way which is ambiguous between the two notions, for example, "Don't multiply postulations beyond necessity." Here it is unclear whether 'postulation' refers to the entities being postulated, or the hypotheses which are doing the postulating, or both. The first reading corresponds to parsimony, the second to elegance. Examples of both sorts of simplicity principle can be found in the quotations given earlier in this section.

While these two facets of simplicity are frequently conflated, it is important to treat them as distinct. One reason for doing so is that considerations of parsimony and of elegance typically pull in different directions. Postulating extra entities may allow a theory to be formulated more simply, while reducing the ontology of a theory may only be possible at the price of making it syntactically more complex. For example the postulation of Neptune, at the time not directly observable, allowed the perturbations in the orbits of other observed planets to be explained without complicating the laws of celestial mechanics. There is typically a trade-off between ontology and ideology—to use the terminology favored by Quine—in which contraction in one domain requires expansion in the other. This points to another way of characterizing the elegance/parsimony distinction, in terms of simplicity of theory versus simplicity of world respectively.[4] Sober [2001] argues that both these facets of simplicity can be interpreted in terms of minimization. In the (atypical) case of theoretically idle entities, both forms of minimization pull in the same direction; postulating the existence of such entities makes both our theories (of the world) and the world (as represented by our theories) less simple than they might be.

# 2. Ontological Parsimony

Perhaps the most common formulation of the ontological form of Occam's Razor is the following:

> (OR) Entities are not to be multiplied beyond necessity.

It should be noted that modern formulations of Occam's Razor are connected only very tenuously to the 14th-century figure William of Ockham. We are not here interested in the exegetical question of how Ockham intended his 'Razor' to function, nor in the uses to which it was put in the context of medieval metaphysics.[5] Contemporary philosophers have tended to reinterpret OR as a principle of theory choice: OR implies that—other things being equal—it is rational to prefer theories which commit us to smaller ontologies. This suggests the following paraphrase of OR:

> (OR$_1$) Other things being equal, if $T_1$ is more ontologically parsimonious than $T_2$ then it is

rational to prefer $T_1$ to $T_2$.

What does it mean to say that one theory is more ontologically parsimonious than another? The basic notion of ontological parsimony is quite straightforward, and is standardly cashed out in terms of Quine's concept of *ontological commitment*. A theory, $T$, is ontologically committed to $F$s if and only if $T$ entails that $F$'s exist (Quine 1981, pp. 144–4). If two theories, $T_1$ and $T_2$, have the same ontological commitments except that $T_2$ is ontologically committed to $F$s and $T_1$ is not, then $T_1$ is more parsimonious than $T_2$. More generally, a sufficient condition for $T_1$ being more parsimonious than $T_2$ is for the ontological commitments of $T_1$ to be a proper subset of those of $T_2$. Note that $OR_1$ is considerably weaker than the informal version of Occam's Razor, OR, with which we started. OR stipulates only that entities should not be multiplied *beyond necessity*. $OR_1$, by contrast, states that entities should not be multiplied *other things being equal*, and this is compatible with parsimony being a comparatively weak theoretical virtue.

One 'easy' case where $OR_1$ can be straightforwardly applied is when a theory, $T$, postulates entities which are explanatorily idle. Excising these entities from $T$ produces a second theory, $T^*$, which has the same theoretical virtues as $T$ but a smaller set of ontological commitments. Hence, according to $OR_1$, it is rational to pick $T^*$ over $T$. (As previously noted, terminology such as 'pick' and 'prefer' is crucially ambiguous between epistemic and methodological versions of Occam's Razor. For the purposes of defining ontological parsimony, it is not necessary to resolve this ambiguity.) However, such cases are presumably rare, and this points to a more general worry concerning the narrowness of application of $OR_1$. First, how often does it actually happen that we have two (or more) competing theories for which 'other things are equal'? As biologist Kent Holsinger remarks,

> Since Occam's Razor ought to be invoked only when several hypotheses explain the same set of facts equally well, in practice its domain will be very limited…[C]ases where competing hypotheses explain a phenomenon equally well are comparatively rare (Holsinger 1980, pp. 144–5).

Second, how often are one candidate theory's ontological commitments a proper subset of another's? Much more common are situations where ontologies of competing theories overlap, but each theory has postulates which are not made by the other. Straightforward comparisons of ontological parsimony are not possible in such cases.

Before setting aside the definitional question for ontological parsimony, one further distinction should be mentioned. This distinction is between *qualitative* parsimony (roughly, the number of types (or kinds) of thing postulated) and *quantitative* parsimony (roughly, the number of individual things postulated).[6] The default reading of Occam's Razor in the bulk of the philosophical literature is as a principle of qualitative parsimony. Thus Cartesian dualism, for example, is less qualitatively parsimonious than materialism because it is committed to two broad kinds of entity (mental and physical) rather than one. Section 6.1 contains a brief discussion of quantitative parsimony; apart from this the focus will be on the qualitative notion. It should be noted that interpreting Occam's Razor in terms of *kinds* of entity brings with it some extra philosophical baggage of its own. In particular, judgments of parsimony become dependent on how the world is sliced up into kinds. Nor is guidance from extra-philosophical usage—and in particular from science—always clearcut. For example, is a previously undiscovered subatomic particle made up of a novel rearrangement of already discovered sub-particles a new 'kind'? What about a biological species, which presumably does not contain any novel basic constituents? Also, ought more weight to be given to broad and seemingly fundamental divisions of kind—for example between the mental and physical—than between more parochial divisions? Intuitively, the postulation of a new kind of matter would seem to require much more extensive and solid justification than the postulation of a new sub-species of spider.[7]

The third and final question from Section 1 concerns potential justifications for principles of ontological parsimony such as Occam's Razor. The demand for justification of such principles can be understood in two

importantly distinct ways, corresponding to the distinction between epistemic principles and methodological principles made at the end of Section 1. Justifying an epistemic principle requires answering an epistemic question: why are parsimonious theories more likely to be true? Justifying a methodological principle requires answering a pragmatic question: why does it make practical sense for theorists to adopt parsimonious theories?[8] Most attention in the literature has centered on the first, epistemic question. It is easy to see how syntactic elegance in a theory can bring with it pragmatic advantages such as being more perspicuous, being easier to use and manipulate, and so on. But the case is more difficult to make for ontological parsimony.[9] It is unclear what particular pragmatic disadvantages accrue to theories which postulate extra kinds of entities; indeed—as was mentioned in the previous section—such postulations can often bring with them striking syntactic simplification.

Before looking at approaches to answering the epistemic justification question, mention should be made of two positions in the literature which do not fall squarely into either the pragmatic or epistemic camp. The first position, associated primarily with Quine, argues that parsimony carries with it pragmatic advantages and that pragmatic considerations themselves provide rational grounds for discriminating between competing theories (Quine 1966, Walsh 1979). The Quinean position bases an answer to the second question on the answer to the first, thus blurring the boundary between pragmatic and epistemic justification. The second position, due to Sober, rejects the implicit assumption in both the above questions that some global justification of parsimony can be found (Sober 1988, 1994). Instead Sober argues that appeals to parsimony always depend on local background assumptions for their rational justification. Thus Sober writes:

> The legitimacy of parsimony stands or falls, in a particular research context, on subject matter specific (and *a posteriori*) considerations. […] What makes parsimony reasonable in one context may have nothing in common with why it matters in another (Sober 1994).

Philosophers who reject these arguments of Quine and Sober, and thus take the demand for a global, epistemic justification seriously, have developed a variety of approaches to justifying parsimony. Most of these approaches can be collected under two broad headings:

> (A) *A priori* philosophical, metaphysical, or theological justifications.

> (B) Naturalistic justifications, based on appeal to scientific practice.

As we shall see, the contrast between these two sorts of approach mirrors a broader divide between the rival traditions of rationalism and empiricism in philosophy as a whole.

As well as parsimony, the question of rational justification can also be raised for principles based on elegance, the second facet of simplicity distinguished in Section 1. Approaches to justifying elegance along the lines of (A) and (B) are possible, but much of the recent work falls under a third category;

> (C) Justifications based on results from probability theory and/or statistics.

The next three sections examine these three modes of justification of simplicity principles. The *a priori* justifications in category (A) concern simplicity in both its parsimony and elegance forms. The justifications falling under category (B) pertain mostly to parsimony, while those falling under category (C) pertain mostly to elegance.

# 3. *A Priori* Justifications of Simplicity

The role of simplicity as a theoretical virtue seems so widespread, fundamental, and implicit that many philosophers, scientists, and theologians have sought a justification for principles such as Occam's Razor on similarly broad and basic grounds. This rationalist approach is connected to the view that making *a priori* simplicity assumptions is the only way to get around the underdetermination of theory by data. Until the

second half of the 20<sup>th</sup> Century this was probably the predominant approach to the issue of simplicity. More recently, the rise of empiricism within analytic philosophy led many philosophers to argue disparagingly that *a priori* justifications keep simplicity in the realm of metaphysics (see Zellner et al. 2001, p.1). Despite its changing fortunes, the rationalist approach to simplicity still has its adherents. For example, Richard Swinburne writes:

> I seek…to show that—other things being equal—the simplest hypothesis proposed as an explanation of phenomena is more likely to be the true one than is any other available hypothesis, that its predictions are more likely to be true than those of any other available hypothesis, and that it is an ultimate *a priori* epistemic principle that simplicity is evidence for truth (Swinburne 1997, p. 1).

### (i) Theological Justifications

The post-medieval period coincided with a gradual transition from theology to science as the predominant means of revealing the workings of nature. In many cases, espoused principles of parsimony continued to wear their theological origins on their sleeves, as with Leibniz's thesis that God has created the best and most complete of all possible worlds, and his linking of this thesis to simplifying principles such as light always taking the (time-wise) shortest path. A similar attitude—and rhetoric—is shared by scientists through the early modern and modern period, including Kepler, Newton, and Maxwell.

Some of this rhetoric has survived to the present day, especially among theoretical physicists and cosmologists such as Einstein and Hawking.[10] Yet there are clear dangers with relying on a theological justification of simplicity principles. Firstly, many—probably most—contemporary scientists are reluctant to link methodological principles to religious belief in this way. Secondly, even those scientists who do talk of 'God' often turn out to be using the term metaphorically, and not necessarily as referring to the personal and intentional Being of monotheistic religions. Thirdly, even if there is a tendency to justify simplicity principles via some literal belief in the existence of God, such justification is only rational to the extent that rational arguments can be given for the existence of God.[11]

For these reasons, few philosophers today are content to rest with a theological justification of simplicity principles. Yet there is no doubting the influence such justifications have had on past and present attitudes to simplicity. As Smart (1994) writes:

> There is a tendency…for us to take simplicity…as a guide to metaphysical truth. Perhaps this tendency derives from earlier theological notions: we expect God to have created a beautiful universe (Smart 1984, p. 121).

### (ii) Metaphysical Justifications

One approach to justifying simplicity principles is to embed such principles in some more general metaphysical framework. Perhaps the clearest historical example of systematic metaphysics of this sort is the work of Leibniz. The leading contemporary example of this approach—and in one sense a direct descendent of Leibniz's methodology—is the possible worlds framework of David Lewis. In one of his earlier works, Lewis writes,

> I subscribe to the general view that qualitative parsimony is good in a philosophical or empirical hypothesis (Lewis 1973, p. 87).

Lewis has been attacked for not saying more about what exactly he takes simplicity to be (see Woodward 2003). However, what is clear is that simplicity plays a key role in underpinning his metaphysical framework, and is also taken to be a *prima facie* theoretical virtue.

### (iii) 'Intrinsic Value' Justifications

Some philosophers have approached the issue of justifying simplicity principles by arguing that simplicity has intrinsic value as a theoretical goal. Sober, for example, writes:

> Just as the question 'why be rational?' may have no non-circular answer, the same may be true of the question 'why should simplicity be considered in evaluating the plausibility of hypotheses?' (Sober 2001, p. 19).

Such intrinsic value may be 'primitive' in some sense, or it may be analyzable as one aspect of some broader value. For those who favor the second approach, a popular candidate for this broader value is aesthetic. Derkse [1992] is a book-length development of this idea, and echoes can be found in Quine's remarks—in connection with his defense of Occam's Razor—concerning his taste for "clear skies" and "desert landscapes." In general, forging a connection between aesthetic virtue and simplicity principles seems better suited to defending methodological rather than epistemic principles.

### (iv) Justifications via Principles of Rationality

Another approach is to try to show how simplicity principles follow from other better established or better understood principles of rationality.[12] For example, some philosophers just stipulate that they will take 'simplicity' as shorthand for whatever package of theoretical virtues is (or ought to be) characteristic of rational inquiry. A more substantive alternative is to link simplicity to some particular theoretical goal, for example unification (see Friedman 1983). While this approach might work for elegance, it is less clear how it can be maintained for ontological parsimony. Conversely, a line of argument which seems better suited to defending parsimony than to defending elegance is to appeal to a principle of epistemological conservatism. Parsimony in a theory can be viewed as minimizing the number of 'new' kinds of entities and mechanisms which are postulated. This preference for old mechanisms may in turn be justified by a more general epistemological caution, or conservatism, which is characteristic of rational inquiry.

Note that the above style of approach can be given both a rationalist and an empiricist gloss. If unification, or epistemological conservatism, are themselves *a priori* rational principles, then simplicity principles stand to inherit this feature if this approach can be carried out successfully. However, philosophers with empiricist sympathies may also pursue analysis of this sort, and then justify the base principles either inductively from past success or naturalistically from the fact that such principles are in fact used in science.

To summarize, the main problem with *a priori* justifications of simplicity principles is that it can be difficult to distinguish between an *a priori* defense and *no* defense(!). Sometimes the theoretical virtue of simplicity is invoked as a primitive, self-evident proposition that cannot be further justified or elaborated upon. (One example is the beginning of Goodman and Quine's 1947 paper, where they state that their refusal to admit abstract objects into their ontology is "based on a philosophical intuition that cannot be justified by appeal to anything more ultimate.") (Goodman & Quine 1947, p. 174). It is unclear where leverage for persuading skeptics of the validity of such principles can come from, especially if the grounds provided are not themselves to beg further questions. Misgivings of this sort have led to a shift away from justifications rooted in 'first philosophy' towards approaches which engage to a greater degree with the details of actual practice, both scientific and statistical. These other approaches will be discussed in the next two sections.

## 4. Naturalistic Justifications of Simplicity

The rise of naturalized epistemology as a movement within analytic philosophy in the second half of the 20$^{th}$ Century has largely sidelined the rationalist style of approach. From the naturalistic perspective, philosophy is conceived of as continuous with science, and not as having some independently privileged status. The perspective of the naturalistic philosopher may be broader, but her concerns and methods are not fundamentally different from those of the scientist. The conclusion is that science neither needs—nor can

legitimately be given—external philosophical justification. It is against this broadly naturalistic background that some philosophers have sought to provide an epistemic justification of simplicity principles, and in particular principles of ontological parsimony such as Occam's Razor.

The main empirical evidence bearing on this issue consists of the patterns of acceptance and rejection of competing theories by working scientists. Einstein's development of Special Relativity—and its impact on the hypothesis of the existence of the electromagnetic ether—is one of the episodes most often cited (by both philosophers and scientists) as an example of Occam's Razor in action (see Sober 1981, p. 153). The ether is by hypothesis a fixed medium and reference frame for the propagation of light (and other electromagnetic waves). The Special Theory of Relativity includes the radical postulate that the speed of a light ray through a vacuum is constant relative to an observer no matter what the state of motion of the observer. Given this assumption, the notion of a universal reference frame is incoherent. Hence Special Relativity implies that the ether does not exist.

This episode can be viewed as the replacement of an empirically adequate theory (the Lorentz-Poincaré theory) by a more ontologically parsimonious alternative (Special Relativity). Hence it is often taken to be an example of Occam's Razor in action. The problem with using this example as evidence for Occam's Razor is that Special Relativity (SR) has several other theoretical advantages over the Lorentz-Poincaré (LP) theory in addition to being more ontologically parsimonious. Firstly, SR is a simpler and more unified theory than LP, since in order to 'save the phenomena' a number of *ad hoc* and physically unmotivated patches had been added to LP. Secondly, LP raises doubts about the physical meaning of distance measurements. According to LP, a rod moving with velocity, $v$, contracts by a factor of $(1 - v^2/c^2)^{1/2}$. Thus only distance measurements that are made in a frame at rest relative to the ether are valid without modification by a correction factor. However, LP also implies that motion relative to the ether is in principle undetectable. So how is distance to be measured? In other words, the issue here is complicated by the fact that—according to LP—the ether is not just an extra piece of ontology but an *undetectable* extra piece. Given these advantages of SR over LP, it seems clear that the ether example is not merely a case of ontological parsimony making up for an otherwise inferior theory.

A genuine test-case for Occam's Razor must involve an ontologically parsimonious theory which is not clearly superior to its rivals in other respects. An instructive example is the following historical episode from biogeography, a scientific subdiscipline which originated towards the end of the 18th Century, and whose central purpose was to explain the geographical distribution of plant and animal species.[13] In 1761, the French naturalist Buffon proposed the following law;

> (BL) Areas separated by natural barriers have distinct species.

There were also known exceptions to Buffon's Law, for example remote islands which share (so-called) 'cosmopolitan' species with continental regions a large distance away.

Two rival theories were developed to explain Buffon's Law and its occasional exceptions. According to the first theory, due to Darwin and Wallace, both facts can be explained by the combined effects of two causal mechanisms—dispersal, and evolution by natural selection. The explanation for Buffon's Law is as follows. Species gradually migrate into new areas, a process which Darwin calls "dispersal." As natural selection acts over time on the contingent initial distribution of species in different areas, completely distinct species eventually evolve. The existence of cosmopolitan species is explained by "improbable dispersal," Darwin's term for dispersal across seemingly impenetrable barriers by "occasional means of transport" such as ocean currents, winds, and floating ice. Cosmopolitan species are explained as the result of improbable dispersal in the relatively recent past.

In the 1950's, Croizat proposed an alternative to the Darwin-Wallace theory which rejects their presupposition of geographical stability. Croizat argues that tectonic change, not dispersal, is the principal causal mechanism which underlies Buffon's Law. Forces such as continental drift, the submerging of ocean

floors, and the formation of mountain ranges have acted within the time frame of evolutionary history to create natural barriers between species where at previous times there were none. Croizat's theory was the sophisticated culmination of a theoretical tradition which stretched back to the late 17th Century. Followers of this so-called "extensionist" tradition had postulated the existence of ancient land bridges to account for anomalies in the geographical distribution of plants and animals.[14]

Extensionist theories are clearly less ontologically parsimonious than Dispersal Theories, since the former are committed to extra entities such as land bridges or movable tectonic plates. Moreover, Extensionist theories were (given the evidence then available) not manifestly superior in other respects. Darwin was an early critic of Extensionist theories, arguing that they went beyond the "legitimate deductions of science." Another critic of Extensionist theories pointed to their "dependence on ad hoc hypotheses, such as land bridges and continental extensions of vast extent, to meet each new distributional anomaly" (Fichman 1977, p. 62) The debate over the more parsimonious Dispersal theories centered on whether the mechanism of dispersal is sufficient on its own to explain the known facts about species distribution, without postulating any extra geographical or tectonic entities.

The criticisms leveled at the Extensionist and Dispersal theories follow a pattern that is characteristic of situations in which one theory is more ontologically parsimonious than its rivals. In such situations the debate is typically over whether the extra ontology is really necessary in order to explain the observed phenomena. The less parsimonious theories are condemned for profligacy, and lack of direct evidential support. The more parsimonious theories are condemned for their inadequacy to explain the observed facts. This illustrates a recurring theme in discussions of simplicity—both inside and outside philosophy—namely, how the correct balance between simplicity and goodness of fit ought to be struck. This theme takes center stage in the statistical approaches to simplicity discussed in Section 5.

Less work has been done on describing episodes in science where elegance—as opposed to parsimony—has been (or may have been) the crucial factor. This may just reflect the fact that considerations linked to elegance are so pervasive in scientific theory choice as to be unremarkable as a topic for special study. A notable exception to this general neglect is the area of celestial mechanics, where the transition from Ptolemy to Copernicus to Kepler to Newton is an oft-cited example of simplicity considerations in action, and a case study which makes much more sense when seen through the lens of elegance rather than of parsimony.[15]

Naturalism depends on a number of presuppositions which are open to debate. But even if these presuppositions are granted, the naturalistic project of looking to science for methodological guidance within philosophy faces a major difficulty, namely how to 'read off' from actual scientific practice what the underlying methodological principles are supposed to be. Burgess, for example, argues that what the patterns of scientific behavior show is not a concern with multiplying entities *per se*, but a concern more specifically with multiplying 'causal mechanisms' (Burgess 1998). And Sober considers the debate in psychology over psychological egoism versus motivational pluralism, arguing that the former theory postulates fewer types of ultimate desire but a larger number of causal beliefs, and hence that comparing the parsimony of these two theories depends on what is counted and how (Sober 2001, pp. 14–5). Some of the concerns raised in Sections 1 and 2 also reappear in this context; for example, how the world is sliced up into kinds effects the extent to which a given theory 'multiplies' kinds of entity. Justifying a particular way of slicing becomes more difficult once the epistemological naturalist leaves behind the *a priori*, metaphysical presuppositions of the rationalist approach.

One philosophical debate where these worries over naturalism become particularly acute is the issue of the application of parsimony principles to *abstract* objects. The scientific data is—in an important sense—ambiguous. Applications of Occam's Razor in science are always to concrete, causally efficacious entities, whether land-bridges, unicorns, or the luminiferous ether. Perhaps scientists apply an unrestricted version of Occam's Razor to that portion of reality in which they are interested, namely the concrete, causal, spatiotemporal world. Or perhaps scientists apply a 'concretized' version of Occam's Razor unrestrictedly.

Which is the case? The answer determines which general philosophical principle we end up with: ought we to avoid the multiplication of objects of whatever kind, or merely the multiplication of concrete objects? The distinction here is crucial for a number of central philosophical debates. Unrestricted Occam's Razor favors monism over dualism, and nominalism over platonism. By contrast, 'concretized' Occam's Razor has no bearing on these debates, since the extra entities in each case are not concrete.

# 5. Probabilistic/Statistical Justifications of Simplicity

The two approaches discussed in Sections 3 and 4—*a priori* rationalism and naturalized empiricism—are both in some sense extreme. Simplicity principles are taken either to have no empirical grounding, or to have solely empirical grounding. Perhaps as a result, both these approaches yield vague answers to certain key questions about simplicity. In particular, neither seems equipped to answer how exactly simplicity ought to be balanced against empirical adequacy. Simple but wildly inaccurate theories are not hard to come up with. Nor are accurate theories which are highly complex. But how much accuracy should be sacrificed for a gain in simplicity? The black-and-white boundaries of the rationalism/empiricism divide may not provide appropriate tools for analyzing this question. In response, philosophers have recently turned to the mathematical framework of probability theory and statistics, hoping in the process to combine sensitivity to actual practice with the 'trans-empirical' strength of mathematics.

Philosophically influential early work in this direction was done by Jeffreys and by Popper, both of whom tried to analyze simplicity in probabilistic terms. Jeffreys argued that "the simpler laws have the greater prior probability," and went on to provide an operational measure of simplicity, according to which the prior probability of a law is $2^{-k}$, where $k$ = order + degree + absolute values of the coefficients, when the law is expressed as a differential equation (Jeffreys 1961, p. 47). A generalization of Jeffreys' approach is to look not at specific equations, but at *families* of equations. For example, one might compare the family, LIN, of linear equations (of the form $y = a + bx$) with the family, PAR, of parabolic equations (of the form $y = a + bx + cx^2$). Since PAR is of higher degree than LIN, Jeffreys' proposal assigns higher probability to LIN. Laws of this form are intuitively simpler (in the sense of being more elegant).

Popper [1959] pointed out that Jeffreys' proposal, as it stands, contradicts the axioms of probability. Every member of LIN is also a member of PAR, where the coefficient, $c$, is set to 0. Hence 'Law, $L$, is a member of LIN' entails 'Law, $L$, is a member of PAR.' Jeffreys' approach assigns higher probability to the former than the latter. But it follows from the axioms of probability that when $A$ entails $B$, the probability of $B$ is greater than or equal to the probability of $A$. Popper argues, in contrast to Jeffreys, that LIN has *lower* prior probability than PAR. Hence LIN is—in Popper's sense—more falsifiable, and hence should be preferred as the default hypothesis. One response to Popper's objection is to amend Jeffrey's proposal and restrict members of PAR to equations where $c \neq 0$.

More recent work on the issue of simplicity has borrowed tools from statistics as well as from probability theory. It should be noted that the literature on this topic tends to use the terms 'simplicity' and 'parsimony' more-or-less interchangeably (see Sober 2003). But, whichever term is preferred, there is general agreement among those working in this area that simplicity is to be cashed out in terms of the number of free (or 'adjustable') parameters of competing hypotheses. Thus the focus here is totally at the level of theory. Philosophers who have made important contributions to this approach include Forster and Sober [1994] and Lange [1995].

The standard case in the statistical literature on parsimony concerns *curve-fitting*.[16] We imagine a situation in which we have a set of discrete data points and are looking for the curve (i.e. function) which has generated them. The issue of what family of curves the answer belongs in (e.g. in LIN or in PAR) is often referred to as *model-selection*. The basic idea is that there are two competing criteria for model selection— parsimony and goodness of fit. The possibility of measurement error and 'noise' in the data means that the correct curve may not go through every data point. Indeed, if goodness of fit were the only criterion then

there would be a danger of 'overfitting' the model to accidental discrepancies unrepresentative of the broader regularity. Parsimony acts as a counterbalance to such overfitting, since a curve passing through every data point is likely to be very convoluted and hence have many adjusted parameters.

If proponents of the statistical approach are in general agreement that simplicity should be cashed out in terms of number of parameters, there is less unanimity over what the goal of simplicity principles ought to be. This is partly because the goal is often not made explicit. (An analogous issue arises in the case of Occam's Razor. 'Entities are not to be multiplied beyond necessity.' But necessity for *what*, exactly?) Forster [2001] distinguishes two potential goals of model selection, namely probable truth and predictive accuracy, and claims that these are importantly distinct (Forster 2001, p. 95). Forster argues that predictive accuracy tends to be what scientists care about most. They care less about the probability of an hypothesis being exactly right than they do about it having a high degree of accuracy.

One reason for investigating statistical approaches to simplicity is a dissatisfaction with the vagaries of the *a priori* and naturalistic approaches. Statisticians have come up with a variety of numerically specific proposals for the trade-off between simplicity and goodness of fit. However, these alternative proposals disagree about the 'cost' associated with more complex hypotheses. Two leading contenders in the recent literature on model selection are the Akaike Information Criterion [AIC] and the Bayesian Information Criterion [BIC]. AIC directs theorists to choose the model with the highest value of $\{\log L(\Theta_k)/n\} - k/n$, where $\Theta_k$ is the best-fitting member of the class of curves of polynomial degree $k$, $\log L$ is log-likelihood, and $n$ is the sample size. By contrast, BIC maximizes the value of $\{\log L(\Theta_k)/n\} - k\log[n]/2n$. In effect, BIC gives an extra positive weighting to simplicity by a factor of $\log[n]/2$ (where $n$ is the size of the sample).[17]

Extreme answers to the trade-off problem seem to be obviously inadequate. Always picking the model with the best fit to the data, regardless of its complexity, faces the prospect (mentioned earlier) of 'overfitting' error and noise in the data. Always picking the simplest model, regardless of its fit to the data, cuts the model free from any link to observation or experiment. Forster associates the 'Always Complex' and the 'Always Simple' rule with empiricism and rationalism respectively.[18] All the candidate rules that are seriously discussed by statisticians fall in between these two extremes. Yet they differ in their answers over how much weight to give simplicity in its trade-off against goodness of fit. In addition to AIC and BIC, other rules include Neyman-Pearson hypothesis testing, and the minimum description length (MDL) criterion.

There are at least three possible responses to the varying answers to the trade-off problem provided by different criteria. One response, favored by Forster and by Sober, is to argue that there is no genuine conflict here because the different criteria have different aims. Thus AIC and BIC might both be optimal criteria, if AIC is aiming to maximize predictive accuracy whereas BIC is aiming to maximize probable truth. Another difference that may influence the choice of criterion is whether the goal of the model is to extrapolate beyond given data or interpolate between known data points. A second response, typically favored by statisticians, is to argue that the conflict is genuine but that it has the potential to be resolved by analyzing (using both mathematical and empirical methods) which criterion performs best over the widest class of possible situations. A third, more pessimistic, response is to argue that the conflict is genuine but is unresolvable. Kuhn (1977) takes this line, claiming that how much weight individual scientists give a particular theoretical virtue, such as simplicity, is solely a matter of taste, and is not open to rational resolution. McAllister (2007) draws ontological morals from a similar conclusion, arguing that sets of data typically exhibit multiple patterns, and that different patterns may be highlighted by different quantitative techniques.

Aside from this issue of conflicting criteria, there are other problems with the statistical approach to simplicity. One problem, which afflicts any approach emphasizing the elegance aspect of simplicity, is language relativity. Crudely put, hypotheses which are syntactically very complex in one language may be syntactically very simple in another. The traditional philosophical illustration of this problem is Goodman's

'grue' challenge to induction. Are statistical approaches to the measurement of simplicity similarly language relative, and—if so—what justifies choosing one language over another? It turns out that the statistical approach has the resources to at least partially deflect the charge of language relativity. Borrowing techniques from information theory, it can be shown that certain syntactic measures of simplicity are asymptotically independent of choice of measurement language.[19]

A second problem for the statistical approach is whether it can account not only for our preference for small numbers over large numbers (when it comes to picking values for coefficients or exponents in model equations), but also our preference for whole numbers and simple fractions over other values. In Gregor Mendel's original experiments on the hybridization of garden peas, he crossed pea varieties with different specific traits, such as tall versus short or green seeds versus yellow seeds, and then self-pollinated the hybrids for one or more generations.[20] In each case one trait was present in all the first-generation hybrids, but both traits were present in subsequent generations. Across his experiments with seven different such traits, the ratio of dominant trait to recessive trait averaged 2.98 : 1. On this basis, Mendel hypothesized that the true ratio is 3 : 1. This 'rounding' was made prior to the formulation of any explanatory model, hence it cannot have been driven by any theory-specific consideration. This raises two related questions. First, in what sense is the 3 : 1 ratio hypothesis *simpler* than the 2.98 : 1 ratio hypothesis? Second, can this choice be justified within the framework of the statistical approach to simplicity? The more general worry lying behind these questions is whether the statistical approach, in defining simplicity in terms of number of adjustable parameters, is replacing the broad issue of simplicity with a more narrowly—and perhaps arbitrarily—defined set of issues.

A third problem with the statistical approach concerns whether it can shed any light on the specific issue of ontological parsimony. At first glance, one might think that the postulation of extra entities can be attacked on probabilistic grounds. For example, quantum mechanics together with the postulation 'There exist unicorns' is less probable than quantum mechanics alone, since the former logically entails the latter. However, as Sober has pointed out, it is important here to distinguish between *agnostic* Occam's Razor and *atheistic* Occam's Razor. Atheistic OR directs theorists to claim that unicorns do *not* exist, in the absence of any compelling evidence in their favor. And there is no relation of logical entailment between {*QM* + there exist unicorns} and {*QM* + there do not exist unicorns}. This also links back to the terminological issue. Models involving circular orbits are more parsimonious—in the statisticians' sense of 'parsimonious'—than models involving elliptical orbits, but the latter models to not postulate the existence of any more things in the world.

# 6. Other Issues Concerning Simplicity

This section addresses three distinct issues concerning simplicity and its relation to other methodological issues. These issues concern quantitative parsimony, plenitude, and induction.

## 6.1 Quantitative Parsimony

Theorists tend to be frugal in their postulation of new entities. When a trace is observed in a cloud-chamber, physicists may seek to explain it in terms of the influence of a hitherto unobserved particle. But, if possible, they will postulate one such unobserved particle, not two, or twenty, or 207 of them. This desire to minimize the number of individual new entities postulated is often referred to as *quantitative parsimony*. David Lewis articulates the attitude of many philosophers when he writes:

> I subscribe to the general view that qualitative parsimony is good in a philosophical or empirical hypothesis; but I recognize no presumption whatever in favour of quantitative parsimony (Lewis 1973, p. 87).

Is the initial assumption that one particle is acting to cause the observed trace more rational than the

assumption that 207 particles are so acting? Or is it merely the product of wishful thinking, aesthetic bias, or some other non-rational influence?

Nolan [1997] examines these questions in the context of the discovery of the neutrino.[21] Physicists in the 1930's were puzzled by certain anomalies arising from experiments in which radioactive atoms emit electrons during so-called Beta decay. In these experiments the total spin of the particles in the system before decay exceeds by $1/2$ the total spin of the (observed) emitted particles. Physicists' response was to posit a 'new' fundamental particle, the neutrino, with spin $1/2$ and to hypothesize that exactly one neutrino is emitted by each electron during Beta decay.

Note that there is a wide range of very similar neutrino theories which can also account for the missing spin.

> $H_1$: 1 neutrino with a spin of $1/2$ is emitted in each case of Beta decay.

> $H_2$: 2 neutrinos, each with a spin of $1/4$ are emitted in each case of Beta decay.

and, more generally, for any positive integer $n$,

> $H_n$: $n$ neutrinos, each with a spin of $1/2n$ are emitted in each case of Beta decay.

Each of these hypotheses adequately explains the observation of a missing $1/2$-spin following Beta decay. Yet the most quantitatively parsimonious hypothesis, $H_1$, is the obvious default choice.[22]

One promising approach is to focus on the relative explanatory power of the alternative hypotheses, $H_1$, $H_2$, … $H_n$. When neutrinos were first postulated in the 1930's, numerous experimental set-ups were being devised to explore the products of various kinds of particle decay. In none of these experiments had cases of 'missing' $1/3$-spin, or $1/4$-spin, or $1/100$-spin been found. The absence of these smaller fractional spins was a phenomenon which competing neutrino hypotheses might potentially help to explain.

Consider the following two competing neutrino hypotheses:

> $H_1$: 1 neutrino with a spin of $1/2$ is emitted in each case of Beta decay.

> $H_{10}$: 10 neutrinos, each with a spin of $1/20$, are emitted in each case of Beta decay.

Why has no experimental set-up yielded a 'missing' spin-value of $1/20$? $H_1$ allows a better answer to this question than $H_{10}$ does, for $H_1$ is consistent with a simple and parsimonious explanation, namely that there exist no particles with spin $1/20$ (or less). In the case of $H_{10}$, this potential explanation is ruled out because $H_{10}$ explicitly postulates particles with spin $1/20$. Of course, $H_{10}$ is consistent with *other* hypotheses which explain the non-occurrence of missing $1/20$-spin. For example, one might conjoin to $H_{10}$ the law that neutrinos are always emitted in groups of ten. However, this would make the overall explanation less syntactically simple, and hence less virtuous in other respects. In this case, quantitative parsimony brings greater explanatory power. Less quantitatively parsimonious hypotheses can match this power only by adding auxiliary claims which decrease their syntactic simplicity. Thus the preference for quantitatively parsimonious hypotheses emerges as one facet of a more general preference for hypotheses with greater explanatory power.

One distinctive feature of the neutrino example is that it is 'additive.' It involves postulating the existence of a collection of qualitatively identical objects which collectively explain the observed phenomenon. The explanation is additive in the sense that the overall phenomenon is explained by summing the individual positive contributions of each object. Whether the above approach can be extended to non-additive cases involving quantitative parsimony (for example, postulating the existence of unseen planets to explain an observed perturbation of an orbit) is unclear.[23]

## 6.2 Principles of Plenitude

Ranged against the principles of parsimony discussed in previous sections is an equally firmly rooted (though less well-known) tradition of what might be termed "principles of explanatory sufficiency."[24] These principles have their origins in the same medieval controversies that spawned Occam's Razor. Ockham's contemporary, Walter of Chatton, proposed the following counter-principle to Occam's Razor:

> [I]f three things are not enough to verify an affirmative proposition about things, a fourth must be added, and so on (quoted in Maurer 1984, p. 464).

A related counter-principle was later defended by Kant:

> The variety of entities should not be rashly diminished (Kant 1950, p. 541).
> *Entium varietates non temere esse minuendas.*

There is no inconsistency in the coexistence of these two families of principles, for they are not in direct conflict with each other. Considerations of parsimony and of explanatory sufficiency function as mutual counter-balances, penalizing theories which stray into explanatory inadequacy or ontological excess.[25] What we see here is an historical echo of the contemporary debate among statisticians concerning the proper trade-off between simplicity and goodness of fit.

There is, however, a second family of principles which do appear directly to conflict with Occam's Razor. These are so-called 'principles of plenitude.' Perhaps the best-known version is associated with Leibniz, according to whom God created the best of all possible worlds with the greatest number of possible entities. More generally, a principle of plenitude claims that if it is *possible* for an object to exist then that object *actually* exists. Principles of plenitude conflict with Occam's Razor over the existence of physically possible but explanatorily idle objects. Our best current theories presumably do not rule out the existence of unicorns, but nor do they provide any support for their existence. According to Occam's Razor we ought *not* to postulate the existence of unicorns. According to a principle of plenitude we *ought* to postulate their existence.

The rise of particle physics and quantum mechanics in the 20th Century led to various principles of plenitude being appealed to by scientists as an integral part of their theoretical framework. A particularly clear-cut example of such an appeal is the case of magnetic monopoles.[26] The 19th-century theory of electromagnetism postulated numerous analogies between electric charge and magnetic charge. One theoretical difference is that magnetic charges must always come in oppositely-charged pairs, called "dipoles" (as in the North and South poles of a bar magnet), whereas single electric charges, or "monopoles," can exist in isolation. However, no actual magnetic monopole had ever been observed. Physicists began to wonder whether there was some theoretical reason why monopoles could not exist. It was initially thought that the newly developed theory of quantum mechanics ruled out the possibility of magnetic monopoles, and this is why none had ever been detected. However, in 1931 the physicist Paul Dirac showed that the existence of monopoles is consistent with quantum mechanics, although it is not required by it. Dirac went on to assert the existence of monopoles, arguing that their existence is not ruled out by theory and that "under these circumstances one would be surprised if Nature had made no use of it" (Dirac 1930, p. 71, note 5). This appeal to plenitude was widely—though not universally—accepted by other physicists.

> One of the elementary rules of nature is that, in the absence of laws prohibiting an event or phenomenon it is bound to occur with some degree of probability. To put it simply and crudely: anything that *can* happen *does* happen. Hence physicists must assume that the magnetic monopole exists unless they can find a law barring its existence (Ford 1963, p. 122).

Others have been less impressed by Dirac's line of argument:

Dirac's…line of reasoning, when conjecturing the existence of magnetic monopoles, does not differ from 18th-century arguments in favour of mermaids…[A]s the notion of mermaids was neither intrinsically contradictory nor colliding with current biological laws, these creatures were assumed to exist.[27]

It is difficult to know how to interpret these principles of plenitude. Quantum mechanics diverges from classical physics by replacing of a deterministic model of the universe with a model based on objective probabilities. According to this probabilistic model, there are numerous ways the universe could have evolved from its initial state, each with a certain probability of occurring that is fixed by the laws of nature. Consider some kind of object, say unicorns, whose existence is not ruled out by the initial conditions plus the laws of nature. Then one can distinguish between a weak and a strong version of the principle of plenitude. According to the weak principle, if there is a small finite probability of unicorns existing then given enough time and space unicorns will exist. According to the strong principle, it follows from the theory of quantum mechanics that if it is possible for unicorns to exist then they do exist. One way in which this latter principle may be cashed out is in the 'many-worlds' interpretation of quantum mechanics, according to which reality has a branching structure in which every possible outcome is realized.

## 6.3 Simplicity and Induction

The problem of induction is closely linked to the issue of simplicity. One obvious link is between the curve-fitting problem and the inductive problem of predicting future outcomes from observed data. Less obviously, Schulte [1999] argues for a connection between induction and ontological parsimony. Schulte frames the problem of induction in information-theoretic terms: given a data-stream of observations of non-unicorns (for example), what general conclusion should be drawn? He argues for two constraints on potential rules. First, the rule should converge on the truth in the long run (so if no unicorns exist then it should yield this conclusion). Second, the rule should minimize the maximum number of changes of hypothesis, given different possible future observations. Schulte argues that the 'Occam Rule'—conjecture that $\Omega$ does not exist until it has been detected in an experiment—is optimal relative to these constraints. An alternative rule —for example, conjecturing that $\Omega$ exists until 1 million negative results have been obtained—may result in two changes of hypothesis if, say, $\Omega$'s are not detected until the 2 millionth experiment. The Occam Rule leads to at most one change of hypothesis (when an $\Omega$ is first detected). (See also Kelly [2004].)

With respect to the justification question, arguments have been made in both directions. Scientists are often inclined to justify simplicity principles on broadly inductive grounds. According to this argument, scientists select new hypotheses based partly on criteria that have been generated inductively from previous cases of theory choice. Choosing the most parsimonious of the acceptable alternative hypotheses has tended to work in the past. Hence scientists continue to use this as a rule of thumb, and are justified in so doing on inductive grounds. One might try to bolster this point of view by considering a counterfactual world in which all the fundamental constituents of the universe exist in pairs. In such a 'pairwise' world, scientists might well prefer pairwise hypotheses in general to their more parsimonious rivals. This line of argument has a couple of significant weaknesses. Firstly, one might legitimately wonder just how successful the choice of parsimonious hypotheses has been; examples from chemistry spring to mind, such as oxygen molecules containing two atoms rather than one. Secondly, and more importantly, there remains the issue of explaining *why* the preference for parsimonious hypotheses in science has been as successful as it has been.

Making the justificatory argument in the reverse direction, from simplicity to induction, has a strong historical precedent in philosophical approaches to the problem of induction, from Hume onwards. Justifying the 'straight rule' of induction by appeal to some general Principle of Uniformity is an initially appealing response to the skeptical challenge. However, in the absence of a defense of the underlying Principle itself (and one which does not, on pain of circularity, depend inductively on past success), it is unclear how much progress this represents. There have also been attempts (see e.g. Steel [2009]) to use simplicity considerations to respond to Nelson Goodman's 'new riddle of induction.'

# Bibliography

- Aquinas, T. (1945) *Basic Writings of St. Thomas Aquinas*, trans. A.C. Pegis, New York: Random House.
- Audi, R. (ed.) (1995) *The Cambridge Dictionary of Philosophy*, Cambridge: Cambridge University Press.
- Baker, A. (2001) "Mathematics, Indispensability and Scientific Practice," *Erkenntnis*, 55, 85–116.
- ——. (2003) "Quantitative Parsimony and Explanation," *British Journal for the Philosophy of Science*, 54, 245–259.
- ——. (2007) "Occam's Razor in Science: a Case Study from Biogeography," *Biology and Philosophy*, 22, 193–215.
- Barnes, E. (2000) "Ockham's Razor and the Anti-Superfluity Principle", *Erkenntnis*, 53, 353–74.
- Bennett, J. (ed.) (1965) *Experiments in Plant Hybridisation* (by Gregor Mendel), London: Oliver & Boyd.
- Bilaniuk, O.-M. & Sudarshan, E. (1969) "Particles Beyond the Light Barrier," *Physics Today*, 22, 43–52.
- Bunge, M. (1963) *The Myth of Simplicity: Problems in Scientific Philosophy*, Englewood Cliffs: Prentice Hall.
- Burgess, J. (1998) "Occam's Razor and Scientific Method," in Schirn (ed.) (1998), 195–214.
- Derkse, W. (1992) *On Simplicity and Elegance*, Delft: Eburon.
- Dirac, P. (1930) "The Proton," *Nature*, London, 126, 606.
- Fetzer, J. (ed.) (1984) *Principles of Philosophical Reasoning*, Totowa, NJ: Rowman & Allanheld.
- Fichman (1977) "Zoography and the Problem of Land Bridges," *Journal for the History of Biology*, 10(1), 45–63.
- Ford, K. (1963) "Magnetic Monopoles," *Scientific American*, 209, 122–31.
- Forster, M. (1995) "The Curve-Fitting Problem," in Audi (ed.) (1995).
- ——. (2001) "The New Science of Simplicity," in Zellner et al. (eds.) (2001), 83–119.
- Forster, M. & Sober, E. (1994) "How to Tell when Simpler, More Unified, or Less Ad Hoc Theories will Provide More Accurate Predictions," *British Journal for the Philosophy of Science*, 45, 1–35.
- Friedman, M. (1983) *Foundations of Space-Time Theories*, Princeton: Princeton University Press.
- Galileo, G. *Dialogue Concerning the Two Chief Wolrd System*, transl. Drake (1962), Berkeley.
- Gauch, H. (2003) *Scientific Method in Practice*, Cambridge: Cambridge University Press.
- Goodman, N. (1955) *Fact, Fiction, and Forecast*, Cambridge, MA: Harvard University Press.
- Goodman, N. & Quine, W. (1947) "Steps Towards a Reconstructive Nominalism," *Journal of Symbolic Logic*, 12, 105–122.
- Groarke, L. (1992) "Following in the Footsteps of Aristotle: the Chicago School, the Glue-stick and the Razor," *Journal of Speculative Philosophy*, 6(3), 190–205.
- Holsinger, K. (1981) "Comment: the Blunting of Occam's Razor," *Canadian Journal of Zoology*, 59, 144–6.
- Jeffreys, H. (1961) *Theory of Probability*, Oxford: Clarendon Press.
- Kant, I. *The Critique of Pure Reason*, transl. Kemp Smith (1950), London.
- Kelly, K. (2004) "Justification and Truth-Finding Efficiency: How Ockham's Razor Works," *Minds and Machines*, 14, 485–505.
- Kemeny, J. (1959) *A Philosopher Looks at Science*, New York: Van Nostrand.
- Kragh, H. (1981) "The Concept of the Monopole: a Historical and Analytical Case-Study," *Studies in History and Philosophy of Science*, 12(2), 141–72.
- Kuhn, T. (1977) "Objectivity, Value Judgment, and Theory Choice," in *The Essential Tension*, Chicago: University of Chicago Press, 320–39.
- Lange, M. (1995) "Spearman's Principle," *British Journal for the Philosophy of Science*, 46, 503–52.
- Lavoisier, A. (1862) "Réflexions sur le Phlogistique," in *Oeuvres*, vol. 2, 623–4, Paris: Imprimerie Impériale.

- Lewis, D. (1973) *Counterfactuals*, Oxford: Basil Blackwell.
- Maurer, A. (1984) "Ockham's Razor and Chatton's Anti-Razor," *Mediaeval Studies*, 46, 463-75.
- McAllister, J. (2007) "Model Selection and the Multiplicity of Patterns of Empirical Data," *Philosophy of Science*, 74, 884–894.
- McKeon, R. (1941) *The Basic Works of Aristotle*, New York: Random House.
- Nash, L. (1963) *The Nature of the Natural Sciences*, Boston: Little, Brown.
- Nelson, G. (1978) "From Candolle to Croizat: Comments on the History of Biogeography," *Journal of the History of Biology*, 11(2), 269–305.
- Newton, I. (1964) *The Mathematical Principles of Natural Philosophy*, New York: Citadel Press.
- Nolan, D. (1997) "Quantitative Parsimony", *British Journal for the Philosophy of Science*, 48, 329–43.
- –––. (1999) "Is Fertility Virtuous in its Own Right?", *British Journal for the Philosophy of Science*, 50, 265–82.
- Plutynski, A. (2005) "Parsimony and the Fisher-Wright Debate," *Biology and Philosophy*, 20, 697–713.
- Popper, K. (1959) *The Logic of Scientific Discovery*, London: Hutchinson.
- Quine, W. (1966) "On Simple Theories of a Complex World", in *The Ways of Paradox*, New York: Random House.
- –––. (1981) *Theories and Things*, Cambridge, MA: Harvard University Press.
- Rescher, N. (1998) *Complexity: a Philosophical Overview*, New Brunswick, NJ: Transaction.
- Sarkar, S. & Pfeifer, J. (eds.) (2006) *The Philosophy of Science: An Encyclopedia*, London: Routledge.
- Schirn, M. (ed.) (1998) *The Philosophy of Mathematics Today*, Oxford, New York: Clarendon Press.
- Schlesinger, G. (1963) *Method in the Physical Sciences*, London: Routledge.
- Schulte, O. (1999) "Means-End Epistemology," *British Journal for the Philosophy of Science*, 50, 1–31.
- Smart, J. (1984) "Ockham's Razor," in *Principles of Philosophical Reasoning*, ed. Fetzer, 118–28.
- Sober, E. (1981) "The Principle of Parsimony," *British Journal for the Philosophy of Science*, 32, 145–56.
- –––. (1988) *Reconstructing the Past: Parsimony, Evolution and Inference*, Cambridge, MA: MIT Press.
- –––. (1994) "Let's Razor Ockham's Razor", in *From A Biological Point of View*, Cambridge: Cambridge University Press, 136–57.
- –––. (2001) "What is the Problem of Simplicity?" in Zellner et al. (eds.) (2001), 13–31.
- –––. (2006) "Parsimony," in Sarkar & Pfeifer (eds.) (2006), Vol 2, pp. 531–538.
- Steel, D. (2009) "Testability and Ockham's Razor: How Formal and Statistical Learning Theory Converge in the New Riddle of Induction," *Journal of Philosophical Logic*, 38, 471–489.
- Swinburne, R. (1997) *Simplicity as Evidence for Truth*, Milwaukee: Marquette University Press.
- Thornburn, W. (1918) "The Myth of Occam's Razor," *Mind*, 27, 345–53.
- Vitányi, L. & Li, M. (2001) "Simplicity, Information, Kolmogorov Complexity and Prediction," in Zellner et al. (eds.) (2001), 135–55.
- Walsh, D. (1979) "Occam's Razor: A Principle of Intellectual Elegance", *American Philosophical Quarterly*, 16, 241–4.
- Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford University Press.
- Zellner, A., Keuzenkamp, H. & McAleer, M. (eds.) (2001) *Simplicity, Inference and Modelling: Keeping It Sophisticatedly Simple*, Cambridge: Cambridge University Press.

# Other Internet Resources

- David Dowe's Occam's Razor links (Computer Science, Monash University)

[Please contact author with other suggestions.]

# Related Entries

computability and complexity | induction: problem of | learning theory, formal | logic: inductive | naturalism | Ockham [Occam], William

---

---