

The Structure of Scientific Evolution

Peter R. Killeen
Arizona State University

Science is the construction and testing of systems that bind symbols to sensations according to rules. Material implication is the primary rule, providing the structure of definition, elaboration, delimitation, prediction, explanation, and control. The goal of science is not to secure truth, which is a binary function of accuracy, but rather to increase the information about data communicated by theory. This process is symmetric and thus entails an increase in the information about theory communicated by data. Important components in this communication are the elevation of data to the status of facts, the descent of models under the guidance of theory, and their close alignment through the evolving retroductive process. The information mutual to theory and data may be measured as the reduction in the entropy, or complexity, of the field of data given the model. It may also be measured as the reduction in the entropy of the field of models given the data. This symmetry explains the important status of parsimony (how thoroughly the data exploit what the model can say) alongside accuracy (how thoroughly the model represents what can be said about the data). Mutual information is increased by increasing model accuracy and parsimony, and by enlarging and refining the data field under purview.

Key words: epistemology, explanation, mapping, material implication, model, theory, truth

My contribution is more philosophical than scientific. This is risky. “Scientists need philosophers,” it has been said, “like birds need ornithologists.” Philosophy is necessary, however, because not all simple questions have simple answers; indeed, not all simple questions are well posed (Machado & Silva, 2007) and may frame an investigation in a suboptimal way. The enduring questions in philosophy often endure because they are the wrong questions. To ask, as Skinner (1950) did, “Are theories of learning necessary?” begs a more difficult question: “Necessary for what?” As behaviorists, what are our scientific goals, in light of which theories of learning may be deemed necessary, unnecessary, or even counterproductive to their achievement? Any credible answer, such as “to describe the conditions under which learning occurs,” requires a clear definition of *learning*, an ability to determine what a *condition* is, how to measure both, and how best to

determine their relation. Most important, it requires us to understand what a theory is, and that in turn requires embedding it in a coherent scientific structure. Scientists need philosophy, as we shall see, like ornithologists need birds.

Aristotle gave us the structure to frame the answers to these questions. He identified the four (or five) kinds of information needed to comprehend phenomena (Alvarez, 2009; Killeen, 2001). These kinds of information, mistranslated as “causes” (Hocutt, 1974), are shown in Figure 1 for the operant response. These causes may be parsed into immediate or molecular causes (the inner circle) and molar or long-term causes. Circling through the molecular causes, the three-term contingency defines the *operant* as a movement of the organism that satisfies some criterion (e.g., switch closure), triggered by a discriminative stimulus, and characteristically followed and maintained by a reinforcing stimulus. The final cause, or function, of the response is its instrumentality in obtaining the reinforcer. The neurophysiology of the process is currently uncertain but is some variant of Hebb’s law.

Address correspondence to the author at the Department of Psychology, Arizona State University, Tempe, Arizona 85287 (e-mail: killeen@asu.edu).

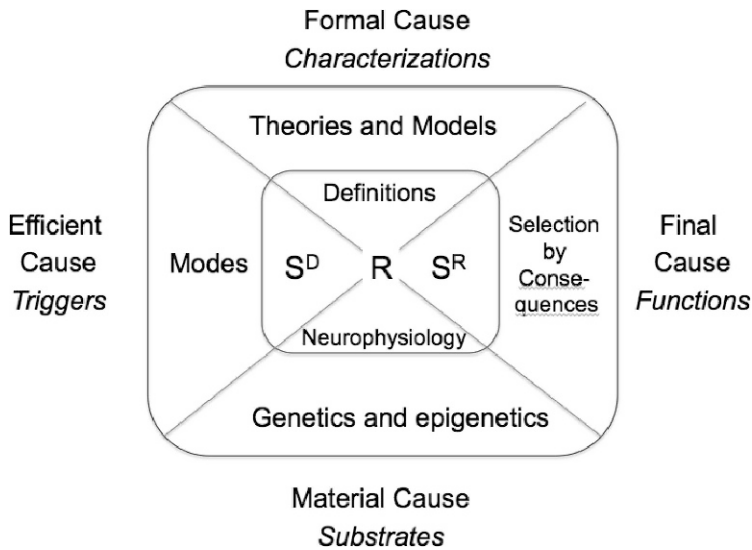


Figure 1. Aristotle's four causes applied to the analysis of the operant response. These causes answer *when*, *what*, *why*, and *how* questions. Formal causes, the focus of this paper, concern the form of the object of inquiry, given by reference to a formula: sentences, symbols, maps, cumulative records, or equations.

At a molar level, the efficient cause describes the context in which learning occurs, as exemplified in, for example, Timberlake's behavior systems theory (1993, 1994; Timberlake & Lucas, 1989). The molar material cause tells us how organisms are configured to learn, typically in the language of genetics and epigenetics. The final causes tell us why organisms are endowed with abilities to learn and perform an operant response, and that answer lies in the selection by consequences of organisms that learn, that is, by evolutionary theory. The molar formal causes, the theories and models of behavior, comprise the central topic of this paper.

The Aristotelian framework is a principled way to approach any important phenomenon. It is useful for inquiries as divergent as comprehending the nature of operant behavior (Killeen, 2001), embodied cognition (Killeen & Glenberg, 2010), developmental disabilities (Killeen, Tannock, & Sagvolden, 2012), and exceptional abilities (Killeen & Nash, 2003). It reminds us that explanations

in terms of reinforcement (consequentialism), or substrate (reductionism), or triggers (mechanistic), or equations (formalistic) are each important but incomplete parts of a comprehensive treatment and not improvements on, or substitutes for, the other causes. There is substantial debate within this scientific community on the precedence of one explanatory mode over another (e.g., Schlinger, 2011, and his references) but general recognition of their importance in toto. In subsequent writing, Aristotle added a fifth cause: an effect's ability to function as a trigger in the future. This is central to behavioral theory, exemplified by reinforcement in the context of a particular discriminative stimulus having the ability to make that stimulus increasingly effective as a trigger of the response. But this paper focuses on formal causes, because they provide the structure of science.

PARTS AS MODELS OF WHOLES

One of the objections to theories of learning, residing at the molar level of

formal causes, is that they often build too much into the brain (or into our accounts of what the brain does), giving those parts characteristics of the whole that they are proposed to explain. This is a kind of *metonymy*: a figure of speech in which a part stands for the whole or vice versa. People associate events and compute outcomes; should we be permitted to say that brains do the same things? How much is permitted of a theory? May we treat thinking as subvocal speech? Not only can the part stand for the whole, as above, but the whole may be embodied in the part. The *homunculus* is a small straw man, a minuscule version of us, who lives in our skull, monitors synapses for input, and pulls tendons for output. He associates stimuli and computes responses. There are some things that we can do that our homunculus cannot (e.g., hit a pop fly to left field or dress a turkey). The central question in considering the homunculus as a theory is what, if any, things the homunculus can do as well or better than we, his shell plus him, can do. Does he add enough value to reify him? Our bodies cannot time travel, other than by moving ahead stolidly, one day at a time, but fantasy takes our homunculus elsewhere. Might this create enough evolutionary pressure to evolve such an incubus to help us with planning? How are the radical behaviorist positions (that thinking is covert behavior, that imagining is seeing without the thing seen, and the like) anything other than shrinking a human to homuncular size and kicking him upstairs? What constraints need to be placed on those assertions, and what constraints do they need to place on data, to be proper? To get answers beyond the conventional arguments (which are numerous and sophisticated; see again the volume introduced by Schlinger, 2011), we have to understand the nature and constraints on science and the proper place of analogies and their proper responsibilities.

Skinner's objections to theory echo his objections to mentalisms. He draws a fine line by abjuring mentalistic explanations such as "expectation" and "purpose," but he embraced sanitized versions of them; personalities as multiple interacting repertoires, and behavior as organized "with respect to" reinforcers (Skinner, 1953). Theories, Skinner said, are "any explanation of an observed fact which appeals to events taking place somewhere else, at some other level of observation, described in different terms, and measured, if at all, in different dimensions" (1950, p. 193). Such theories, many behaviorists infer, are off limits to our discipline. Consider, however, how you might explain thunder to your child. It starts with cloud formation through the sun's evaporation of ocean water, with the segregation of electric charges in the clouds producing an electric field. Eventually the field exceeds the air's dielectric constant, causing lightning. Air rushes in to fill the vacuum left by the stroke, causing thunder. This account appeals to events that take place somewhere else, at different levels of observation, described in different terms, and measured, if at all, in different dimensions (certainly not in the decibels of thunder). But just how would one give an explanation of thunder in terms of decibels? The common explanatory recourse of behaviorists to a "history of reinforcement" refers to events somewhere else, at some other level of observation. Theories generally invoke phenomena that are different from the things that they explain. They are part of a layered formal framework (the top wedge of Figure 1) that provides the substance of this essay.

THE EPISTEMOLOGICAL LAYER CAKE

Figure 2 shows a conventional separation of phenomena into two realms, the rational and the empiri-

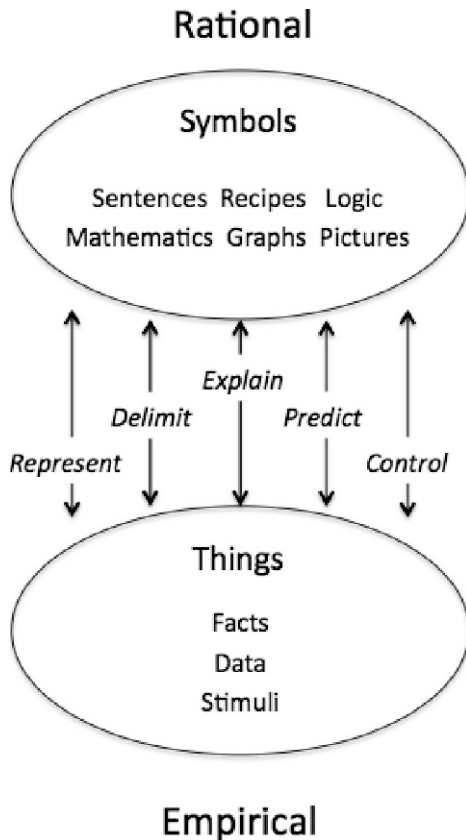


Figure 2. Science as mapping. Naming, numbering, measuring, circumscribing, elaborating, explaining, and predicting are all parts of the job descriptions of scientists. Engineers and technicians control. In all cases these activities involve finding, creating, or using a correspondence between formal structures and empirical structures. The logical rule of implication plays a key role in all of these functions.

cal. Inhabitants of the former are things we call structures and symbols, equations and formulas, that is, formal causes. They are generated and refined by formalists. Inhabitants of the latter are things we call data. They are generated and refined by empiricists. The business of science is to construct links between these realms. Scientists are the gluons that bind data to symbols, symbols to models. They use theories to construct models that share some of the character of the data.

Formal Systems

Formal systems are collections of abstract structures, including definitions of their elements and rules of their interaction. The interactions may be static or dynamic. Euclidean geometry is a classic formal system. An angle is the structure formed by the intersection of two straight lines. That the sum of three angles of a triangle equals 180° is a static rule of interaction. Given a triangular structure, specification of two angles determines the size of the third angle and dictates its distance. Rational systems such as geometry help us to cope with the empirical world because there is order in the latter that rational models can emulate. By *order* is meant that the empirical world is not as random as possible; its entropy is less than it might be. It is this order that models abstract and clarify. We triangulate distances, whether to mountains or houses, using the same geometric model. This mapping of formal structures onto empirical structures is the fundamental act of making meaning. All meaning is metaphorical in that sense: It is a mapping into one system (one we believe we have some understanding of) to parts of another (that we seek to understand). Metaphors are allusions, colored by connotations on a canvas of historical associations. Models, whether logical or mathematical, put a finer point on metaphor's broad brush. They do so by reducing the degrees of freedom for interpretation that is inherent in metaphor. They trade the metaphor's evocativeness for the model's precision.

Empirical Systems

In its lowest form, this is the ur-world of primitive sensations, the world that the introspectionist psychologists Wilhelm Wundt and E. B. Titchener hoped to recover by studied regression to the naive perceptual state of a child (Boring, 1929). But by the time a child is old enough to be a

subject in the laboratory, his or her earliest “buzzing-booming confusion” (James, 1890/1983) has congealed ineradicably into “things.” Had they caught the child any earlier, he or she could not have spoken coherently about the buzz. The treatment of primitive sensations as data, and data’s elevation into facts, are but the first steps up a many-tiered layer cake. Rest for a moment on these two levels, model and facts, and consider the mappings that science draws between them according to the task that the individual assays. The first question to address is the process by which the maps are constructed.

MAPPINGS

Implication

One rational structure stands out as a universal tool for the operation of all five bridging operations between the formal and the empirical domains. The *material implication*, also called the *material conditional*, is a logical connective written $A \rightarrow C$. Although the symbols A and C refer to antecedent and consequent, the implicit “before” and “after” in those labels refer to their location in the logical phrase, as written from left to right. We do not require that the antecedent precede the consequent in time, when those are instantiated as particular events. Material implication plays a key role in all of the endeavors of a scientist. For instance, a term is typically defined by listing the Conditions A under which we will call an Event C : If the rat depresses the lever with at least 0.1-N force, then that is a response. Operational definitions list the criteria for measurement required to assert that a particular event should be included in a class of that name. Such definitions clarify how we should talk, not necessarily how nature is. When out of touch with viable models of the phenomena, operational definitions can mislead; and when used in ways too divergent from common usage (a

tactic called “shifting the referent”), they often provide a quick solution to an intellectual quandary that is like the fast food from a vending machine: somewhat surprising, briefly appealing, and ultimately bad for you.

Sufficiency. Material implications are statements of sufficiency: If the antecedent occurs, the consequent must follow: If A , then C . Nothing else is required. In particular, the presence of C does not entail the presence of A : C may appear for other reasons, as a sneeze may be due to a cold, or an allergy, or bright sunlight. The form of argument called *modus ponens* involves the presentation of A and invocation of a rule $A \rightarrow C$, to conclude C . We may also infer from that rule that if C does not occur, A could not have been present: $\sim C \rightarrow \sim A$. The form of argument called *modus tollens* involves the demonstration of the absence or failure of C , $\sim C$, to conclude $\sim A$. This is a statement of necessity: Without C , no A ; C is necessary for A . Thus the one rule, $A \rightarrow C$, supports both the positive argument of *modus ponens*, and, through its dual, $\sim C \rightarrow \sim A$, the negative argument of *modus tollens*. The first defines what it means for one thing to be sufficient for another; the second what it means for one thing to be necessary for another.

Truth tabled. These relations may also be couched as truth values: If both A and $A \rightarrow C$ are true, then C must be true. Mathematicians and rhetoricians will often use the truth value interpretation for two important types of inference, *induction* and *reductio ad absurdum*. In the former, they show that *if* an equation is true for some number n , *then* it is true for $n + 1$. Next they show that it is true for some number, say $n = 1$. They may then conclude that it is true for all n . It is clear that this use of implication succeeds because it takes the consequent, $n + 1$, and uses it as the new antecedent n ; the argument is

iterated. This is the process used in one of the most beautiful inductive arguments of all time, constructed near the beginning of all recorded time: Euclid's demonstration that there is no greatest prime number. Scientific induction is less trustworthy, because it is impossible to establish the crucial $f(n) \rightarrow f(n + 1)$. On the $n + 1$ st morning the farmer comes, not to feed the turkey, but to feed on him. Less trustworthy still is its use in politics, where an induction called the *domino effect* has provided pretext for wars: If the n th country falls, then too must the $n + 1$ st. Used this way, it is an instance of the *slippery slope fallacy*.

If the mathematician wishes to prove A , and can show that (a) $(\sim A) \rightarrow C$, and (b) C is impossible (i.e., show $\sim C$ either directly, or through an iterated chain of argument as in induction), then (c) by *modus tollens* he or she can then assert the falsity of the antecedent: $(\sim A)$ is false. Finally (d) because $\sim(\sim A) \rightarrow A$, the point is proven. A classic use of this *proof by contradiction* is the proof that $\sqrt{2}$ is irrational. In more general terms, it is the form of argument called *reductio ad absurdum*: using the implication of an absurd consequent to discredit the antecedent. This is at the core of null hypothesis tests, which involve both induction and contradiction. If a sample has ensemble the character M (e.g., a mean = M), then the set of all samples probably has ensemble the character $\mu \approx M$. We assume $\mu = 0$ (the fact that we hope to disprove). $(\mu = 0) \rightarrow (M \approx 0)$. We want to show that the consequent is false so we can reject the antecedent. If data show that M is in fact sufficiently different from 0, so that expecting μ to = 0 is, in retrospect, somewhat absurd [i.e., unlikely: $p(M | \mu \approx 0) < .05$], then we reject the null hypothesis $\sim(\mu = 0)$. There is much to dislike about this argument, not least of which is that it lets us conclude nothing about what we hope to prove (Killeen, 2005).

In vulgar parlance, a *reductio ad absurdum* is heard in the elocution: "If that's the case, then I'm a monkey's uncle!" Some strange consequents, being less absurd to some ears than to others, may leave the auditor believing the implication rather than rejecting the premise. It is unwise to practice the logic of *reductio ad absurdum* in a bar.

The truth value interpretation, along with the defining truth table for implication, reaches limits in the case in which the antecedent has not been observed or is false. According to traditional logic, in this case of counterfactual conditionals, anything follows. That is because the rule itself is given a "true" truth value whatever the consequent. It does this because implication is only a relation of sufficiency, not necessity. There are other ways C may come about or not come about. If it rains hard then the sidewalks will be wet. It doesn't rain. Then the sidewalks may be wet or not, depending whether sprinklers come on, a carton of liquid is dropped, and so on. From a false premise anything follows: "If wishes were horses, then beggars would ride"; but equally: "If wishes were haystacks, then beggars would ride." The mapping of propositional logic to scientific inference fails here; it is itself driven to absurdity. A solution is to never assign truth values to rules. A better role will be found for truth values in a subsequent section of this paper. Of facts we may say "observed" or "not observed"; of rules we may say "holds"; "fails"; or "not tested". Then $A \rightarrow C$ holds if A and C are observed; fails if A and $\sim C$ are observed; and was *not tested* if $\sim A$. Doing otherwise is to assign a truth value to a rule on the basis of no evidence.

The above exercises with implication are important, because all of the roles of a scientist (representation, delimitation, explanation, prediction and control) involve that logical structure. This includes one of the

scientist's chief tasks, inferring causal relations. If temporal precedence is required, along with a few other conditions, then material implication is part of a model of efficient causality. It may also be adduced as a model of selection by consequences: If a stimulus or response A has been followed reliably by a reinforcing consequence C, then the rule $A \rightarrow C$ may be induced. Its manifestation is called a *conditioned response*, because the response C is conditional on the presentation of t A.

Representation

A subset of the elements in the empirical field is associated with a subset of the elements in the rational field. In the simplest case, the symbol may be a name or number; or it could be a more dynamic structure, such as a proposition, rule, or equation. Such rational structures are also the class name (or property or function) of other empirical phenomena that are similar to the target. Pointing to an object and calling it a *pig*, a *triangle*, or a *Latin square* grounds the symbols in the rational realm in a corresponding set of objects in the empirical realm. Grounding in the empirical realm is always eventually ostensive: pointing. If asked for meaning of a term, we often take the shortcut of defining it, which relates it to other terms, one of which, at some time, the inquirer has had pointed out to him or her.

The tension between real and ideal, between things and the stories that attempt to characterize them, is part of an ancient drama. The prisoners in Plato's cave attempted to reconstruct the "true" generating model, the unseen puppets, from the shadows on the wall of the cave. Plato's cave was a seminal model of representation. The ideal circle represents many mundane exemplars. Do ideals, the elements in the formal domain, exist? The ideals are the simplest thing that we can say about reals that captures a

major portion of our uncertainty concerning them. The ideal is not real, even though our discussion of it must be conducted in real symbols such as $(x - h)^2 + (y - k)^2 = r^2$, and real pictures such as O. The ideal is not real, but is typically all that we can easily understand about the real. The residuum is noise, or error variance, and remains that until a better model, one still parsimonious enough to comprehend, is evolved. Because we can understand the simpler ideal more readily than the multifarious real, we tend to reify it.

The left column of Table 1 represents simple models or approaches to them. The right column is what we experience. It may reflect the data used to ground the symbols, or the chaos of unfiltered reality, or the residuum after our models have spoken. Good models share much of their information in mutuality with the elements in their domain. Because ideals are simple, they are quickly learned; the residual may be appreciated as variations on a theme, or deprecated as rogue data or noise.

There is inevitable ambiguity in both the signified and the signifier. In the case of the former, we might be understood to be pointing at a sow rather than a generic member of the litter, or come to believe that all triangles must be right, or not see the crucial pattern in the Latin square. Conversely, we may use the signs for other classes of objects: *pigs* for sloppy or greedy humans, *triangles* for mating arrangements, and *Latin squares* for ancient forums. This slippage, in both the set of objects in the empirical domain that was referred to and the set of referents in the formal domain that was intended, is both inevitable and sometimes useful. Scientific progress involves a continual process of refinement of these maps by limiting or expanding or redefining the range of empirical phenomena to which a label or proposition applies. Equally important is the restriction, refinement, or

TABLE 1

The inhabitants of the two realms

Model	Data
Ideal	Real
Noumenon	Phenomenon
Apollonian	Dionysian
Linear	Nonlinear/Chaotic
Yang	Yin
Conceptual	Empirical
Theoretical	Factual
Logic	Data
Prototype	Exemplar
Type	Token
Template	Target
Signal	Signal + Noise
Parameter	Statistic
Rules	Exceptions
Laws	Anomalies
Nomothetic	Ideographic
Naturwissenschaft	Geisteswissenschaft

creation of new symbolic structures to improve or broaden the correspondence.

Delimitation and Elaboration

Definitions may be couched as material implications. *If* a closed form has three sides consisting of noncollinear straight lines, *then* it is a triangle. *If* an individual manifests at least five symptoms from Category A and at least three from Category B, *then* he or she is categorized as *x*. Material implication identifies sufficient conditions for placing a phenomenon in a category. These are not necessary conditions, because there may be other events that elude the operational definition that belong in the category. *If* it has three legs and a round top that you sit on, *then* it is a stool. But bar stools have four legs. A general treatment will be given below for quantifying this slippage between constraints and constrained, a slippage that Wittgenstein famously characterized as “family resemblance.” This is important in delimitation, because many events may seem at first to lie outside the bourne of a particular definition. Either these are leftovers that the model doesn’t address, or they are inconsistent with the

operational definitions. Not all triangles are right, and neither squares nor circles are triangles. No problem. Rules can still be said to hold when things are left out. Newton’s laws of motion don’t predict color. It is the job of theoretical statements in a model to determine what is “on the table” and what is “off the table.” This delimitation reduces the entropy (the degrees of freedom) in the data field under the purview of the model, an important step in improving the communication between model and data.

The range and domain of models coevolve. Through learning we are able to assimilate new exemplars and distinguish exceptions; at the same time, we adapt the models to accommodate the idiosyncrasies of the new family members.

[Association is] undergoing important evolutionary changes. ... The notion of an association has adapted with changes in the elements that it takes as its arguments, in the conditions under which it is formed, and the way in which it is exhibited in behavior. Moreover, the association has survived by increasingly constraining the range of psychological phenomena it claims to explain. (Rescorla, 1998, p. 1)

After such constraints, a new expansion may occur.

The “law of the hammer” states that if a child is given a good hammer, then he or she will discover that everything needs its application. the extended law of the hammer states that if a scientist is given a good tool, then he or she will discover that everything needs its application. This can often be a good thing, because getting additional mileage out of tools such as microscopes and Skinner boxes and models extends their range and thus power. It amortizes the cost in learning to use them. Seeing a square as comprising two triangles gives an immediate formula for the area of the constituent right triangles: the side squared over two. Seeing a circle as comprising a multitude of triangles

gave Archimedes the best estimate of π available to antiquity. Increasing the purview of a model further increases the information mutual to model and data. Once a complicated tool is mastered, however, it is often used in preference to simpler or more appropriate tools. Its use has acquired behavioral momentum (Nevin, 1996).

Formal systems also require delimitation. Their flexibility must be abridged when the data are easily overpredicted: A seven-parameter model may tell us more about the hiccups in the data than about the beast that is making them. Metrics such as the Akaike information criterion help to address this potential for overpowering data. A classic example of constraining formal systems occurred when S. S. Stevens (1946) established the distinction between scale types (nominal, ordinal, interval, and ratio) and urged that scientists use a type appropriate for their data. If the data are the numbers on football jerseys, it makes little sense to average them. They are nominal measurements, and at best one can ensure that in no case is the same number handed out twice. Conversely, using only nominal or ordinal analyses may waste information when the data can support interval or ratio scales, as is the case for measures of elapsed time. In effect, Stevens's scale distinction is the formalists' version of Morgan's canon (Newbury, 1954; but see Thomas, 2006): "Do not map to stronger scale types if your measurements support only weaker ones."

Formal systems evolve, raising questions of which parts should be carried over to analyses of empirical domains. For millennia, Euclidean geometry was seen as a truth about the world. Non-Euclidean geometries were viewed by many at the time of their invention as aberrations, even by mathematicians such as Charles Dodgson, whose fables parodied

non-Euclidean geometries. Now the preferred geometry of relativistic physics is non-Euclidean, with Euclidean models sufficing for mundane geometry. Similarly, Newtonian mechanics was retained for ordinary earth physics, special relativity was required for particle physics, and general relativity was required for astrophysics.

Explanation

Phenomenon C is noticed, and it begs explanation. Perhaps sheets of green light are seen descending on the tundra. One then casts about for a formal model, a material implication that has "sheets of light in the northern sky" in its consequent. Many may be found and many ruled out. The best kind of potential explanation was in place before the question arose. Others, such as "Your ancestors did it to bless your marriage," fails in the face of the many divorcees to have also witnessed the phenomenon. After finding a relevant rule, a search must be made in the empirical world to determine whether those antecedents were in fact in place as necessary to cause the phenomenon. One of these, an interaction of energetic particles from the sun with the magnetosphere of the earth, will work. It also permits further predictions of correlations of aurora with solar activity. *Explanation*, then, means the identification of a model that is consistent with both the relevant facts that are observed to be present and with other models that are pertinent. The facts that are deemed relevant are a matter of negotiation. The more facts that are consistent with a model, the better. The current model of DNA conveys deep understanding, even though every model, including it, has its limits. As Kuhn (1970) noted, most new models not only incorporate additional facts but also sacrifice some explanations that their predecessors were able to make.

Sometimes no preexisting model adequately maps the facts, or we do not have access to such a model. Then explanations either (a) elicit the creation of testable hypotheses (Peirce, 1903, Rescher, 1978, and Upshur, 1997, called the creation and testing of such hypotheses *abduction*); (b) are held in abeyance (Wittgenstein, 1965, understood the universe of data not captured by language models; of those he said, "What we cannot speak of, we must pass over in silence"); or (c) are made up from a general, or ad hoc, even if unpredictable, model: "God," "survival of the fittest," "academic freedom," "men are just like that." Subsequently dislodging these explanations to see the phenomena with fresh eyes is often difficult; this is why advances in a field often occur to those not yet habituated to anomalies and to their ad hoc interpretations.

All explanations may be couched in the form of material implications. "Why did all the buildings collapse during the earthquake except these?" "You see those triangular struts? If triangles are placed under load, then they can support greater forces than the rectangles that were used in the collapsed buildings." Material implication identifies sufficient conditions for having found a particular outcome. These are not necessary conditions, because there may be other ways to explain the same results. If underlying ground liquefies, then structures built on it will be unstable. Oftentimes, many converging factors are jointly responsible for an outcome. The drive for simplicity often miscarries; many are the scientists who are pleased to find just one sufficient explanation for a phenomenon and reluctant to share credit with the promulgator of another sufficient explanation.

Prediction

Once a map has been constructed between sign and significate, it may

be exploited to make predictions. Now the explanations are tested, and the representations and delimitations begin to pay dividends. Contemporary scientists disliked some of the assumptions of Newton's system of the world, such as his postulate of gravity's action at a distance (Cohen, 1990). His method triumphed thanks to those who could suspend disbelief long enough to see the strong predictions that were possible with that system. (Newton did not require his audience to further suspend disbelief over the calculus: All of the proofs in the *Principia* were geometric.) Good predictions can follow from peccable models. Triangulation has been the primary means of mensurating the earth for hundreds of years, even if plane geometry fails at large distances on this sphere, whose great circle routes are geodesics. But a surveyor is not wrong on those counts. His or her criteria for models is that they be simple, reliable, and deliver the necessary degree of accuracy. Modern physicists generally accept relativity theory while they use Newton's mathematics to calculate orbits for satellites. Scientists forgo even face validity (think quantum mechanics), if the model can provide adequate pragmatic return for their patience. This was James's case for pragmatism; judge statements by their fruits as much as by their roots (James, 1907).

All predictions may be couched in the form of material implications. If a triangle is placed under load, then it will support a greater force than any other form placed under similar load. If an individual has mental disorder x , then he or she will be at substantially greater risk for social maladjustment, drug abuse, and incarceration. Material implication identifies sufficient conditions for expecting to find a particular outcome. These are not necessary conditions, as there may be other ways to achieve the same results. If a stimulus is a reinforcer, then responses that precede it will

increase in frequency. But they may increase in frequency because of a change in drive level, or because they are correlated with other events that are reinforced, or because other incompatible behaviors are punished. Successful predictions strengthen the rule inductively, but because the rules state only sufficient conditions, they cannot prove the rule: False or absent antecedents are consistent with true consequents. A failed prediction will invalidate the rule (*modus tollens*), but even this is typically amended to the status of “an exception to the rule,” and mentioned, if at all, in a footnote. In the social sciences, ability to predict 80% of the variance in the data is most impressive, even though the other 20% constitutes a failure of the rule.

Mathematical predictions. Even mathematical models have the semblance of material implication. The equation for a sine wave is $y(t) = A \sin(\omega t + \phi)$. For Amplitude $A = 1$, angular frequency $\omega = 1$ and phase $\phi = 0$, we may predict that if $t = 1$, then $y(t) = 0.841 \dots$. If we set these conditions (the antecedents, or initial conditions, A , ω , and ϕ) and find that $y(t)$ is some other number, that particular sine function would not hold as a rule for that phenomenon. In vulgar parlance, “The theory would be falsified.” In our way of speaking, that model would not hold for that datum.

Such mathematical models are only sufficient conditions for their predictions. We may find that $y(t) = 0.841 \dots$ at other values of t : In fact, for all values of $t = 1 + 2\pi$ we would find the same value. It is permissible for different antecedents to have the same consequent. It is not permissible, in logic or mathematics (statistics is a different story), for the same antecedent to have different consequents.

Model tuning. An irksome use of the term *prediction* is common in scientific parlance. When the free parameters of a model are adjusted

so as to make the model consonant with observed data, as, say, in adjusting ω in the above paragraph to get 0.841 out of the sine function, the investigator may say “the model predicts the data.” *Retrodicts* or *postdicts* would be more accurate. Just as there is an overvaluation of the role of true prediction in science, there is a correlated arrogation of predictive ability to merely compliant models. It is as though you claimed that your watch can predict sunrise, even though you have to keep adjusting the predictions as the seasons evolve. Such adjustment of models in light of data, in which information flows from the empirical to the rational, is better called *tuning*, or aligning the model to the data. In this case we can say, “the model conforms to the data.” Will the model hold the tune, unmodified, from one set of data to the next? If so, we finally have license to say *predicts*.

Control

Skinner emphasized the role of control in science, influenced perhaps by William James’s pragmatism, which took ability to control as a test of the truth of a model. Certainly the inability to use a model, such as reinforcement theory, to control behavior should raise eyebrows; unruly children are bad witness to a behavioral creed of their parents. The model for control is, again, material implication. Find a rule that has the desired result as the consequent, and instantiate the antecedent. If you want good behavior, catch them being good and reinforce that. Or arrange the antecedent (elicit good behavior) and reinforce it. If that doesn’t work, either we have misidentified the elements (what we used wasn’t a reinforcer, or it was applied in a way that was ineffective, or it reinforced a different aspect of behavior than we had envisaged; exercises in post hoc delimiting) or the rule doesn’t hold. There is, of course,

much art in the application of such rules, because many rules govern behavior, and they do so under strong stimulus control. The same response that is reinforced in a public house is punished in a public school. The same reinforcer that works in each may become a punisher in each if seen as a manifestation of unseemly control.

Newton's celestial mechanics is not faulted because we are unable to use it to move the earth. Ability to control is not necessary to accept a model as provisionally good; our model of aurora borealis is countenanced despite our inability to make an aurora to bless every wedding. Skinner should have argued for "prediction *or* control" as goals of science. Representation, delimitation, and validation then become instruments of those goals. Others would prefer to identify the generation of good explanatory models as the goal of science, equating that with scientific understanding, and using prediction or control as the means of their test and validation. Because the material implication establishes sufficient but not necessary conditions, there are potentially many ways to control a phenomenon. Good technicians own many tools.

MORE LAYERS

The Layers of Inference

Two layers do not suffice for an epistemological cake. Figure 3 expands the model. The top layer is explanatory propensity or predilection. Some individuals prefer astrological explanations to astronomical ones; some prefer Christian orthodoxy to Darwinian biology. Many scientists hold a *mechanistic* worldview, some a *contextualist* worldview; some psychologists are cognitivists; others are behaviorists. A colleague has dedicated his career to demonstrating the error of all linear models of psychological phenomena; this motivates his theoretical undertak-

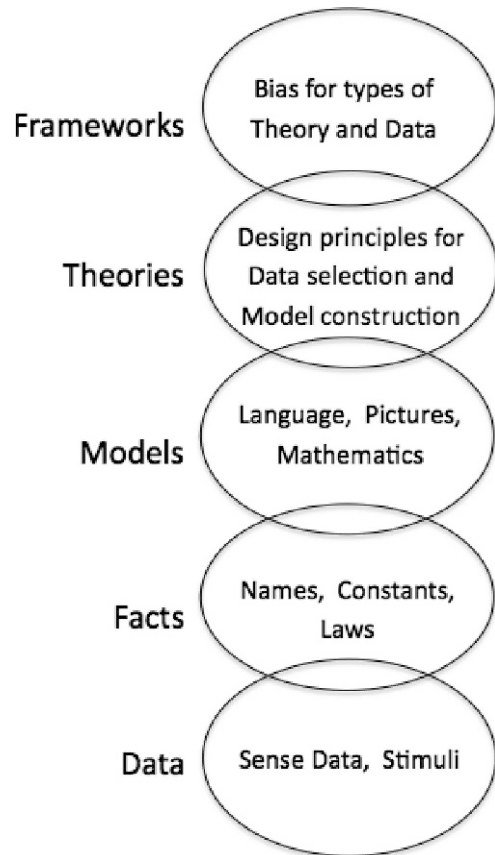


Figure 3. The epistemological layer cake. The kinds of theories that are invoked to make sense of the world are determined by worldviews or themata. Theories specify the design principles for models and delimit the data for which they are responsible. Models, subsets of symbol systems, take antecedent facts or values as input and entail consequent facts. Facts are defined by experts based on empirical analysis of specific sense data. At all levels, material implication plays a central role (based on conversations with David Hestenes).

ings and choice of experiments and is refractory to argument. Einstein's repugnance at quantum indeterminacy is legendary. This is the realm of *worldview*, *framework*, or *themata* (Killeen & Glenberg, 2010). It is an instance of Bacon's *idols of the theatre*, some of which "immigrated into men's minds from the various dogmas of philosophy" (1620, Aphorism XLIV).

Theory. Below frameworks or worldviews lies the realm of theory. The word *theory* is often mistakenly used when hypothesis (a tentative model) is meant. Theories dictate what sort of data constitutes the elements, what sorts of models are appropriate, and what operations with the models are valid. Newton's mechanics provides the classic example of a theory. It defined the elements to be point particles; his three principles of motion gave rules for defining forces in terms of motion and rules for adding sets of forces (Hestenes, 1992). Some general rules for model construction concern coherence, comprehensibility, and parsimony. In behavioral theory, some of the key elements are stimulus, response, and reinforcer; some of the processes are conditioning and extinction. These may be studied in themselves or used to explain other observed phenomena. Some processes, such as induction and context and the reflex reserve, are of ambiguous status and should be invoked very cautiously. Desires, interpretations, and intentions are better not used at all, if a place at the behaviorist table is desired. Some of the rules for data selection are clarity, consistency reproducibility, and simplicity; "smooth curves" in Skinner's parlance.

Models. The modeling game creates and tests new models using rules allowed by theory to represent and explain facts. Friction is introduced as a force opposing motion; air resistance is introduced as a force that can vary in nonlinear ways with the speed of an object. Analysis into constituents and the summation of those constituents is a standard technique of the Newtonian approach, one that is greatly facilitated by use of the calculus (another model system). Kin selection is a simple extension of Darwinian selection, sociobiology a more complicated one. Low-level models, called laws, can be found in this domain, settled to the bottom of the layer close to

facts, or, when universally acknowledged, at the top of the empirical domain. Descartes's laws of reflection and refraction are basic descriptive models, as is Ohm's law relating voltage, current, and resistance. Within limits, they are valid independent of theories that seek to derive them as consequences. Such is also the case for the matching law in operant psychology.

Metamodels. Models are often used to operate, not directly on facts, but on other models. Maxwell's equations originally described the operations of small vortices until those were discarded as unnecessary. One cannot write the equation for a complex atom or molecule without Bohr's billiard-ball model of the atom, because it is to the Bohr model, not directly to nature, to which the various terms in that equation refer (Kuhn, 1979). Ohm's law is derivable from basic physics with suitable assumptions. Any theory in physics that has electrical conduction as its domain would be severely embarrassed if it could not be tuned to deliver Ohm's law. This is the case for other empirical laws, from the laws of reflection and refraction to the matching law. The simple regularities of Ohm's law (or Descartes's law, or Herrnstein's law) were the first targets, not the facts on which the law was based. Many of these derivations, like those of optimal foraging theory, are simply exercises to see just what assumptions have to be invoked to get where you know you want to go. Sometimes they establish values for constants that become a part of the structure of the model and no longer count as free parameters. The Avagadro constant N_A , c , the fine structure constant α , G , and so on, provide examples.

Models may thus be stacked, much in the manner that object-oriented computer languages take as operands other operators. The lower parts of each layer blend with the upper parts of the layer below. Given our limited

working memories, this ability to “take for granted” the outputs of lower level models is essential (if also sometimes fatal: Conditioning to the word can blind us to the thing). Bacon’s “Truth arises more readily from error than from confusion” may be glossed as “An accurate model is more likely to arise from a less accurate one than from raw sense data.” Stacking models compresses them, diminishing their apparent complexity. This is what MacKay (2003) meant by “I believe that one of the main aims of learning is to end up knowing less. ... Brains are the ultimate compression and communication systems” (p. v).

Definitions. Facts are not data. They are consensually agreed-upon labels for data. Some phenomena in the empirical realm cohere as entities with little coaxing; baseballs and hats and sunny days may be taken as given, with only umpires and haberdashers and weathermen reserving opinion. In general, however, decisions of what attributes are sufficient for class inclusion comes about through deliberations of committees and through conditioning of individuals. Whether a dive scores 9 or a wine 90; whether a kiss is a sin, a death a murder, or a manuscript a dissertation; these all depend on the deliberations of finders of fact appropriate to each judgment. Our parents stipulated our rules of etiquette; our clerics our morals. Graduate students spend hours “cleaning up their data.” In this process, some data are dismissed as artifacts or transcription errors or experimental errors. They become nonfacts. Absent an ability to verbalize their criteria, judges may rely on their own emotional reactions to define goods and evils. Such criteria are not easily explained or exported, even though feelings such as disgust that may constitute them can be strongly conditioned in the young.

Rules stipulate the way words are used to nominate data to categories.

The rules of who may run for the U.S. Congress, for instance, and the criteria for electing him or her, are well defined. Certain attributes are required of the candidate, the electors, and the process. Occasionally, courts intervene to change the process, or to “clean up the data,” as in recent presidential elections. Most other nominations are more ambiguous, and in many cases the finders of fact have different theories concerning the process itself. Three current views about nature of nomination are revealed by the dialogue of three umpires after the game. The first, a naive realist: “Some are balls, some are strikes, and I call ‘em as they are.” The second umpire, an idealistic phenomenologist: “Some are balls, some are strikes, and I call ‘em as I see ‘em.” The third umpire, a constructivist: “There ain’t no balls or strikes until I call ‘em.” The first umpire would view instant replay with potential embarrassment, the second nonchalantly, and the third with animosity.

When categorical decisions affect the public good, police, lawyers, judges, and juries of peers are called on to elevate acts to facts. It is a critical question of who has standing on such panels. Many hours are spent in *voir dire*, the process of questioning and often excluding potential jurors from a trial. Many would prefer to exclude congressmen from committees that define the value of π and fundamentalists from school boards that approve biology textbooks. Reading the signatures of muons is best left to physicists, lumps to radiologists, and old bones to archaeologists. When it comes to the categories *good* and *bad*, the experts are first parents and priests and rabbis, later peers, and eventually ethical philosophers. Greek tragedies such as *Antigone* explore the conflict between what is good in the eyes of family, state, and gods; different experts with different priorities. Like *Antigone*, citizens and

juries everywhere are immersed in the conflict among various experts with different priorities in evaluating theories, models, and facts.

EVALUATING MODELS

Estimating the value of a model (evaluating it) involves verification and validation. Verification is the processes of determining whether the model is internally consistent and well formulated. Validation requires the specification of what phenomena a model is responsible for and how good it is at capturing them.

Verification

Model verification involves scrutinizing a model to see if it is what it is claimed to be. This set of operations exists primarily in the rational realm and is often accomplished at the stage of review before publication or product release. Is the math correct? Does the algorithm or analogue function consistently? Is its output unambiguous? Are there unacknowledged steps or assumptions in the model? Will it break down? It may take a team of experts weeks or, through sporadic efforts, even years to verify a model (e.g., Yamaguchi, 2006); complicated models are sometimes impossibly difficult to verify.

Validation

Model validation addresses whether the model succeeds in capturing aspects of the empirical field it was designed for (Krause, 2012). There are two issues in model validation: determining what the model supposed to account for (the delimitation of parts of the empirical domain) and determining how well it does account for those parts. If it is an operational definition, does it leave obvious members of the category outside it? Does it include ringers? How true to the data is it? Is it possible for a model to be false?

Truth. Truth is not a property of the world, of the empirical realm; nor

is it a property of the symbolic representational realm. It is a property of a map between those realms. For Thomas Aquinas, truth was the equation of thing and mind: A judgment is said to be true when it conforms to the external reality. For Descartes, *truth* denoted the conformity of thought with its object. For Russell, a belief is true when there is a corresponding fact and false when there is no corresponding fact. For James, it was a consilience of the statement with both empirical consequences and with other beliefs that are held to be true. One of the best, for its concreteness and thoroughness in referring to all four cells of the matrix that relates state of the world to state of the model was one of the earliest, Aristotle's: "To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, and of what is not that it is not, is true" (Glanzberg, 2009).

Most of these venerable philosophers used a binary model of truth and falsity. But not all did: Eight centuries after Aristotle, Philoponus argued that truth is neither in things or events nor in the statements about them, but rather in the relation between the former and the latter. He gave the simile of truth being like the fit of a foot to a shoe (Wildeberg, 2008). We also treat truthfulness as the fit of a model to data, because binary models of truth shortchange the rich continuum of truthfulness. How many drops of rain must one see to say "It is raining," and be called truthful? How warm the water before saying "It is hot"? How warm the heart before saying "I love you"? There are always degrees of consistency between description and the thing described. This is why judges are given discretion on sentencing: Not all thefts, nor all murders, are created equal. Nor are all fits of feet to shoes, nor of models to data.

An additional problem with the binary truth model is the asymmetry it imposes on model evaluation. A

single false prediction, it is said, will invalidate a model, whereas a thousand true predictions will not make it true. A problem with this standard treatment is that, as Philoponus noted 1,500 years ago, models should never be considered true or false; it is the relation of their consequents to the things in the empirical realm that may prove true or false, and that to varying degrees. Models are tools; a hammer is not false because it cannot turn a screw. It is ill chosen for the task. The relation of a model's antecedents to things in the empirical domain makes it well applied or inappropriate. The relation of a model's consequents to things in the empirical domain makes it accurate, or inaccurate, or even misleading. A model may be inaccurate and yet still quite useful. I do not discard my bathroom scales because they tell me that I need no stamps on my letter, as it apparently weighs nothing (bad application), nor do I because they weigh me heavier than my doctor's scales (imperfect accuracy). We do not discard Newton's laws of motion because they make false predictions for objects at subluminal speeds or are incapable of predicting the trajectories of chaotic systems. They are still used to plan astronomical missions. Exceptions prove the rules; they test them. To pass the test, we must rethink and improve the model, or we must treat the exceptions as anomalies or blemishes; footnote them and hope that no more show up. When the list of anomalies gets too long, the model should be abandoned.

Our model for *truth*, then, is a graph with accuracy as the abscissa and the probability of saying "true" as the ordinate, with an ogive relating them. Executives must set a threshold for their binary actions, but scientists should deal with the abscissae, not ordinate, because it is always more informative to report the accuracy of a model over a particular domain of data than to assign a truth value.

Communication

Models speak to us about the world. They are stories about a set of things in need of a name or explanation or prediction. Just how much they communicate about those things may be calculated as information mutual to model and data. A model may be perfectly accurate about a local fact, one with little complexity. We can increase the amount of information communicated by a model by increasing the domain of the model, the things about which it makes predictions, as long as the prediction error does not increase proportionately. This is why scientists always push models to attempt to account for more data: It is the justification of the extended law of the hammer.

Some facts are equally well captured by different models. A circle is an ellipse with both foci at the same point. An ellipse is a conic section, and so on. Why do we find it more natural to say that that stain on the tablecloth is a circle rather than an ellipse or conic section? Because the latter accurate descriptions also describe infinities of other curves that are irrelevant. The more complicated descriptions of the stain made it possible (indeed, encouraged the listener) to infer things that weren't true. ("Why would he bother to say 'ellipse' if it were simply a circle? It must be elongated.") The default assumption in communication is minimization of complexity: inference to the simplest explanation. The proportion of variance *in the model* that was controlled by the data was much greater for the circle model than for the conic section model, even though both were equally true. The circle was a more parsimonious model for that datum than any other description of conic sections.

As models get more flexible, they become increasingly able to say things that are not true. Model selection requires stringent taming

by data. To say that an object will continue in motion unless opposed by other forces is simple. To specify those forces, such as friction, and how they change with velocity makes the model more powerful and at the same time more complicated. To say that animals match the allocation of their behavior in patches to equal the allocation of reinforcers in those patches is an extremely simple (i.e., parsimonious) model: *matching*. To say that they do so with bias adds both flexibility and a parameter, and thus the need for constraint by data increases. To say that they do so as a biased power function of the allocation of reinforcers (*generalized matching*) increases complexity once again. For that increase to be worthwhile, the prediction error must be very low; the data must unambiguously tell us whether the power is really different from 1 and the bias different from 0. A model that is very flexible in relation to the data it confronts is not parsimonious: Simpler ones would do well enough. The injunction to "Keep models as simple as possible, but no simpler" is an injunction to keep model flexibility as minimal as possible, until predictive accuracy becomes too compromised by the parsimony. Occam's razor should cut whiskers, not skin.

Parsimony is relative. A skeptic might dismiss a model with "Give me enough parameters and I'll draw an elephant." Give him a pencil. He will require at least 80 parameters (coordinates for cartoon ellipses). Any fewer might draw a cow or an egg. The question is how few you can use until enough people start calling your Dumbo "Bessie." This is a key comparison in the evaluation of models: All evaluation is, or should be, against competing models, that is, the *model comparison* approach. Even the coefficient of determination (the proportion of variance accounted for by a model) is a comparison of the candidate model against the mean of the data, a simple statistical model of

central tendency. Comparison depends on the range of models on the table and the domain of data under consideration. Your critic's 18-parameter model elephant might be adequate if the only data under consideration were cartoons of elephants and trees but not if the data included cows.

A central thesis of this paper is that the goal of science is to maximize the information mutual to theory and data. One may do this by increasing the accuracy of the model or by increasing the parsimony of the model. Corresponding operations in the empirical realm are to increase the amount of data it can emulate and to reduce the anomalous data. This is why scientists often give parsimony a standing tantamount to that of accuracy. A model that parsimoniously accounts for a small data set may elicit words such as "neat," "nifty," or "sweet." A model that parsimoniously accounts for a large data set—that shares large mutual information with data—may elicit a deep sense of beauty.

SCIENCE AS COMMUNICATION

The goals of science are best couched in the language of information theory, the language of communication. It is the goal of science to create models of events that report them both accurately and parsimoniously; that maximizes their shared information. How does one measure the information mutual to theory and data? The Kullback-Leibler measure called *divergence* gives the information lost when a model is used to approximate data. A problem with the estimation of divergence is that it requires knowledge of the actual model, the actual data-generation machine. This problem was brilliantly solved by Akaike when he showed that divergence may be estimated as a linear function of $AIC = 2k - 2\ln(L)$, where k is the number of parameters in the model, and L is the maximized value of the likelihood

function for the model (Bozdogan, 2000). The model that minimizes AIC minimizes prediction error. Burnham and Anderson (2001, 2002) provide an accessible introduction to AIC, its extension to small samples, and to the model-comparison philosophy of scientific inference. AIC is sometimes thought of as a correction for model nonparsimony because of two models that deliver comparable values of L , the one with the fewer parameters k gives the smaller AIC, and is thus preferred. The purpose of the AIC is not to enforce parsimony, however, but to provide an unbiased measure of divergence.

Inaccuracy may be measured as the uncertainty in the data given the model. It is best computed with AIC and best interpreted as how much more is left to be said about the data once the model has spoken. Unparsimony is the uncertainty in the model given the data. Its measurement is not yet perfected. It can be interpreted as how much the model can say that is not relevant to the data in its domain. The most promising approach is in the measure of computational complexity called *minimum description length* (Grünwald, 2005; Grünwald, Myung, & Pitt, 2005; Su, Myung, & Pitt, 2005; Wagenmakers & Waldorp, 2006).

The evolution of science is a record of speciation of models and data sets, each in closer accord with the other, under the competitive pressure to better account for more of the entropy in the fields of model and data. The evolutionary fitness of models involves not only accuracy (i.e., goodness of fit) but also parsimony and power. Power is measured by the information mutual to theory and data. Theories, like organisms, are subject to diverse selection pressures: Comparison against data is like natural selection; the prose, power, and novelty with which a model is presented are like sexual selection. Fitness, in this struggle, means more than goodness of fit.

SUMMARY

To comprehend a phenomenon we must understand how to define it, what triggers it, its components, and what makes it endure (Figure 1). The focus of this paper has been on the formal structures central to science (Figure 2), by which we represent, explain, predict, and control events. The formal structures coevolve with the facts that they address. At the lowest levels, they elevate data to the status of facts; at higher levels they constitute theories that give design principles for models and marshal the domain of data (Figure 3). These central activities of scientists and engineers may be represented as implementations of the logical structure called material implication. The source of most models is metaphor and analogy; verbal or physical structures that are noticed to have things in common with one another. Just as organic evolution is driven by two kinds of pressures (natural selection and sexual selection), the evolution of formal structures is driven by two kinds of pressures (rational selection and emotional selection). The former selects based on goodness of fit of model to data and parsimony of model. These are always best treated in a model-comparison framework, because only relative measures are interpretable. Emotional selection is driven by the worldviews, conditioning history, and political and ideological milieu of the scientists and by the clarity and cogency of the proposed model.

It is not the scientists' job to find truth (that binary function on the accuracy of a proposition) but to create representational structures that maximize the information mutual to propositions and the data that they represent. This is accomplished by maximizing the accuracy of a model while minimizing its complexity. Models with greater mutual information are more powerful than those with less, even though they may

be equally accurate in their respective domains of data. Science is an exercise in communication both between model and data and between the model–data system and an audience. Different audiences require different degrees of parsimony, complementary degrees of accuracy. The structure of science evolves by erecting more powerful theories on the shoulders of lesser precedents. Ability to communicate scientific constructs is a key selective pressure in that evolution.

REFERENCES

- Alvarez, M. P. (2009). The four causes of behavior: Aristotle and Skinner. *International Journal of Psychology and Psychological Therapy*, 9, 45–57.
- Bacon, F. (1620). *The New Organon: Or true directions concerning the interpretation of nature* (Vol. Book 1).
- Boring, E. G. (1929). *A history of experimental psychology*. New York, NY: Appleton-Century-Crofts.
- Bozdogan, H. (2000). Akaike's Information criterion and recent developments in information complexity. *Journal of Mathematical Psychology*, 44, 62–91.
- Burnham, K. P., & Anderson, D. R. (2001). Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research*, 28, 111–119.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). New York, NY: Springer-Verlag.
- Cohen, I. B. (1990). Newton's method and Newton's style. In F. Durham & R. D. Purrington (Eds.), *Some truer method: Reflections on the heritage of Newton* (pp. 15–57). New York, NY: Columbia University Press.
- Glanzberg, M. (2009). Truth. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from <http://plato.stanford.edu/archives/spr2009/entries/truth/>
- Grünwald, P. D. (2005). Introducing the minimum description principle. In P. D. Grünwald, I. J. Myung, & M. A. Pitt (Eds.), *Advances in minimum description length: Theory and applications* (pp. 3–21). Cambridge, MA: MIT Press.
- Grünwald, P. D., Myung, I. J., & Pitt, M. A. (2005). *Advances in minimum description length: Theory and applications*. Cambridge, MA: MIT Press.
- Hestenes, D. (1992). Modeling games in the Newtonian world. *American Journal of Physics*, 60, 732–748.
- Hocutt, M. (1974). Aristotle's four because. *Philosophy*, 49, 385–399.
- James, W. (1983). *Principles of psychology*. Cambridge, MA: Harvard University Press. (Original work published 1890)
- James, W. (1907). *Pragmatism: A new name for some old ways of thinking*. New York, NY: Longmans, Green.
- Killeen, P. R. (2001). The four causes of behavior. *Current Directions in Psychological Science*, 10, 136–140.
- Killeen, P. R. (2005). Tea-tests. *The General Psychologist*, 40, 16–19.
- Killeen, P. R., & Glenberg, A. M. (2010). Resituating cognition. *Comparative Cognition & Behavior Reviews*, 5, 59–77.
- Killeen, P. R., & Nash, M. (2003). The four causes of hypnosis. *The International Journal of Clinical and Experimental Hypnosis*, 51, 195–231.
- Killeen, P. R., Tannock, R., & Sagvolden, T. (2012). The four causes of ADHD: A framework. In S. C. Stanford & R. Tannock (Eds.), *Behavioral neuroscience of attention deficit hyperactivity disorder and its treatment* (Vol. 9, pp. 391–425). Berlin, Germany: Springer-Verlag.
- Krause, M. S. (2012). Measurement validity is fundamentally a matter of definition, not correlation. *Review of General Psychology*, 16, 391–400.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (Vol. 2). Chicago, IL: University of Chicago Press.
- Kuhn, T. (1979). Metaphor in science. In A. Ortony (Ed.), *Metaphor and thought* (1st ed., pp. 409–419). New York, NY: Cambridge University Press.
- Machado, A., & Silva, F. J. (2007). Toward a richer view of the scientific method: The role of conceptual analysis. *American Psychologist*, 62, 671–681.
- MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms*. Cambridge, UK: Cambridge University Press.
- Nevin, J. A. (1996). The momentum of compliance. *Journal of Applied Behavior Analysis*, 29, 535–547.
- Newbury, E. (1954). Current interpretation and significance of Lloyd Morgan's canon. *Psychological Bulletin*, 51, 70–74.
- Peirce, C. S. (1903). *Abduction and induction. Philosophical writings of Peirce*. New York, NY: Dover.
- Rescher, N. (1978). *Peirce's philosophy of science*. Notre Dame, IN: University of Notre Dame Press.
- Rescorla, R. A. (1998). *The survival of the association*. Paper presented at the Learning: Association or Computation, Rutgers, NJ.
- Schlinger, H. D. Jr (2011). Introduction: Private events in a natural science of behavior. *The Behavior Analyst*, 34, 181–184.
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, 57, 193–216.

- Skinner, B. F. (1953). *Science and human behavior*. New York, NY: The Free Press.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, *103*, 677–680.
- Su, Y., Myung, I. J., & Pitt, M. A. (2005). Minimum descriptive length and cognitive modeling. In P. D. Grünwald, I. J. Myung, & M. A. Pitt (Eds.), *Advances in minimum description length: Theory and applications* (pp. 411–433). Cambridge, MA: MIT Press.
- Thomas, R. K. (2001). Lloyd Morgan's canon: A history of misrepresentation. *History & Theory of Psychology*. Retrieved from <http://httpprints.yorku.ca/archive/00000017/>
- Timberlake, W. (1993). Behavior systems and reinforcement: An integrative approach. *Journal of the Experimental Analysis of Behavior*, *60*, 105–128.
- Timberlake, W. (1994). Behavior systems, associationism, and Pavlovian conditioning. *Psychonomic Bulletin & Review*, *1*, 405–420.
- Timberlake, W., & Lucas, G. A. (1989). Behavior systems and learning: From misbehavior to general principles. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Instrumental conditioning theory and the impact of constraints on learning* (pp. 237–275). Hillsdale, NJ: Erlbaum.
- Upshur, R. (1997). Certainty, probability and abduction: Why we should look to C. S. Peirce rather than Gödel for a theory of clinical reasoning. *Journal of Evaluation in Clinical Practice*, *3*, 201–206.
- Wagenmakers, E.-J., & Waldorp, L. E. (2006). Editors' introduction. *Journal of Mathematical Psychology*, *50*, 99–100.
- Wildeberg, C. (2008). John Philoponus. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from <http://plato.stanford.edu/archives/fall2008/entries/philoponus/>
- Wittgenstein, L. (1965). *Philosophical investigations 1965*. New York, NY: Macmillan.
- Yamaguchi, M. (2006). Complete solution of the Rescorla-Wagner model for relative validity. *Behavioural Processes*, *71*, 70–73.