Validation in Simulation: Various Positions in the Philosophy of Science

George B. Kleindorfer • Liam O'Neill • Ram Ganeshan
303 Beam Building, The Pennsylvania State University, University Park, Pennsylvania 16802-1913
Health Management and Policy, 2700 Steindler Building, University of Iowa, Iowa City, Iowa 52242-1008
QAOM Department, University of Cincinnati, Cincinnati, Ohio 45221-0130

"I don't see how finite human beings can have any opinion whether they have struck bottom or are on some transitory platform."
—Oliver Wendell Holmes, Jr.

There is still considerable doubt and even anxiety among simulation modelers as to what the methodologically correct guidelines or procedures for validating simulation models should be. Epistemically, the approaches one finds in the simulation literature run the gamut from objectivist to relativist with shades in between. At present in the philosophy of science, there appears to be a convergence toward a nonalgorithmic but discursive and nonrelativistic view of the argumentation involved in warranting scientific theorizing. The present paper attempts to give a description of the various philosophical positions as well as to summarize their problems and the kinds of evidentiary arguments they would each allow in arriving at defensible simulation models. From the debate, we attempt to set out a perspective that frees the practioner to pursue a varied set of approaches to validation with a diminished burden of methodological anxiety. Reciprocally this perspective does not let the modeler off of the hook but rather converts the validation problem into an ethical problem in which the practitioner must responsibly and professionally argue for the warrant of the model.
(*Simulation; Validation; Philosophy of Science; Hermeneutics*)

1. Introduction

The word "simulate" means to build a likeness, and, as such, the question of the accuracy of that likeness is never far behind. The validation problem in simulation is an explicit recognition that simulation models are like miniature scientific theories. Each of them is a set of propositions about how a particular manufacturing or service system works. As such, the warrant we give for these models can be discussed in the same terms that we use in scientific theorizing in general.

Naylor and Finger (1967) outlined three different philosophical positions on which they based a procedure for validating simulation models. Their multistage, empirically-oriented validation procedure is now included in most simulation textbooks (e.g., Law and Kelton 1991), and their oft-cited positions have come to be called the "historical" approaches to validation (Sargent 1992). We would like to revisit this subject. In our view, the limitations of the positions outlined by Naylor and Finger are the source of some of the current problems regarding validation in simulation.

We know from geology that earthquakes and volcanoes are caused by the movement of massive continental plates beneath the earth's surface. One cannot make sense of these surface eruptions without an understanding of plate tectonics. Similarly, one cannot make sense of the current debate regarding validation in simulation without a basic understanding of the weighty

0025-1909/98/4408/1087$05.00

arguments in the philosophy of science that underlie this discourse. Since the publication of Naylor and Finger's seminal article, the dialogue on validation in simulation has played out as a microcosm of the history of the philosophy of science. In recent years, however, there has been a shift in these philosophic plates, and thinking on the validation problem has lagged behind these important developments. Our main focus in this article is not on particular surface manifestations of these plates but rather on their underlying tensions and fault lines. As such, our objective here is not to prescribe a particular technique or algorithm for effecting model validation. While there are many such procedures outlined in the simulation literature (e.g., Sargent 1992), confusion and anxiety regarding the validation question remains. Instead, our purpose is to suggest how main currents in this literature can be used to open our eyes, argue for, and defend suitable stances we might take in holding that our simulation models meet appropriate scientific norms.

This paper is organized as follows: In §2, we discuss current problems regarding validation in simulation and trace the evolution of thinking on this issue. In §3, we provide an overview of the philosophical positions that underlie this debate, while paying particular attention to the relevance of these arguments to simulation modelers. In §4, we summarize the objectivist versus relativist debate which we view as the underlying plate tectonics in the simulation validation problem. In §5, we outline a new direction in the philosophy of science, known as *hermeneutics*, and in the final section, we describe the implications of this philosophical movement and the surrounding debate to the validation problem in simulation.

2. Perspectives on the Problem of Validation in Simulation

If one culls out the sections on validation from any sample of simulation papers, one is immediately struck by the wide variation to be found. There will be descriptions about model behavior, success in application, reservations and restrictions, personal experiences, descriptions of success in the field or lab—in short anything that the experimenter deems relevant to the experience of formulating and applying the model. This diversity, it seems to us, is an indication that at a fundamental level there is still confusion about what "validation" involves or in some cases if it is even feasible to talk about it.

In their classic article on validation, Naylor and Finger (1967) outlined three different philosophical positions, which they called *rationalism*, *empiricism*, and *positive economics*. Their multi-stage procedure had a profound influence on simulation practice and formed the basis of what we have termed the "objectivist" approach to validation in simulation.

Nevertheless, several practical difficulties emerged from the procedures proposed by Naylor and Finger. For example, nowhere in their original article do they give any advice on what to do in the common situation in which the real system does not exist. Instead we find:

> Although the construction and analysis of a simulation model, the validity of which has not been ascertained by empirical observation, may prove to be of interest for expository or pedagogical purposes (e.g., to illustrate particular simulation techniques), such a model contributes nothing to the understanding of the system being simulated.

Given the empirical orientation of these authors, it is hardly surprising that many simulation modelers apparently believe that model validation is an "*either / or*" proposition: either we have unambiguous empirical evidence for warranting the validity of the model or our model is a completely speculative device from which no defensible answers can be obtained. For example, consider this summary of the validation process of a simulation model of emergency vehicle base locations:

> Another relevant issue involved how much our model output could deviate from system output and still remain valid . . . Unfortunately, we were building a model, so some errors and approximations were unavoidable. Model validation here was thus effected by means of a group decision between the client and analysts; when both groups were satisfied, the model was simply considered valid. (Goldberg et al. 1990, p. 133).

The authors apparently believe that any deviation from the real system output are the result of "errors," which diminish the model's validity. The passage also reveals an undercurrent of anxiety, that having given up on the idea of empirical validation, then there are no longer any justifiable criteria with which to argue for the warrant of the model.

While Naylor and Finger argued primarily for a foundationalist stance in simulation validation (we will define this more carefully below), later simulation modelers would argue against what was seen as the intransigence of the foundationalist position. For example, Richardson and Pugh (1981) criticized the then current thinking on validation by likening it to a process of "mathematical inoculation:"

> There is a tendency to think of validation as a process similar to warding off the measles: a model, susceptible to contagious criticism, gets validated and becomes immune to further attack. Often some measure of immunity is sought—a goodness of fit test, a multiple-correlation coefficient—much like a count of antibodies in a mathematical bloodstream. (p. 311)

Barlas and Carpenter (1990) went even further in proposing an alternative "relativist" paradigm of simulation validation, which they offered as the polar antithesis of the empiricist paradigm proposed by Naylor and Finger. In arguing their case, they appealed to philosopher Thomas Kuhn. Recently, Homer (1996) argued against "conceptual" approaches to validation and reiterated the need for empirical confidence building techniques within the systems dynamics tradition: "We should resist the notion that systems dynamics is mainly conceptual rather than empirical." (p. 17). (For a detailed account of empirical validation, see Law and Kelton 1991, and Sargent 1992.)

3. The Validation Problem in General

Some might hold that the term *validation*, that is, "to make valid," is already loaded with a philosophical commitment that a satisfactory model be rendered absolutely "true." The term conveys a sense that a scientific effort must be justified in some logical, objective, and algorithmic way. Hence, it makes assumptions or carries with it the "prejudices" of certain positions in the philosophy of science. We want it recognized that we are using "validation" because it is an accepted term in the simulation literature, and not because we have bought into the rationale for the philosophical positions that some may see behind it. Instead it is specifically our intent here to examine these positions *seriatim* and to question and criticize their relevance and value as perspectives on simulation modeling and practice.

The validation problem has often been discussed more actively in other fields than in management science. In economics, for example, there is a tradition of argument and a vast literature on the justification of economic methodology (e.g., Caldwell 1991, Friedman 1953, and Wible 1982). The same problem has been raised in most areas of the social sciences. (There are over 3000 references in the Social Sciences Citation Index in a recent five year period (January 1990–June 1995) to Kuhn, Popper, and Lakatos.)

3.1. Validation and the Problem of Induction

The fundamental difficulty in warranting both simulation models and scientific theories has to do with the problem of induction. Since an observer has direct access only to his or her own peculiar and limited set of experiences, how can one justify generalizations beyond that particular and personal empirical domain? An analogous problem arises in simulation. How can we infer from our observations (experience) of a system that the model we produce captures its essential structure and parameters?

Historically, two broad and opposing positions have emerged in the philosophy of science which we will call here *foundationalism* and *anti-foundationalism*, or alternatively, *objectivism* and *relativism*. Because the historical development helps in understanding the *raison d' etre* for the particular philosophical positions we will discuss, we will repeatedly refer back to this fundamental distinction. Nevertheless, almost every important philosopher of science, while professedly taking one stance or the other, ameliorates or incorporates into his or her position some way of overcoming the extremes of the polarity. Later we will return to what the amelioration of this opposition might mean for validation in simulation.

Table 1 provides a summary of various positions in the philosophy of science and the related validation approaches in simulation (the classification of the various positions in this table is similar to Wible (1982, p. 355)). We note here that the abbreviated nature of the table can potentially damage the various positions it attempts to describe. (The locus classicus for most of the terminology we will use here is Lakatos 1970; see also Wible 1982, and Bernstein 1983.) *Objectivism* and *justificationism* are terms and positions that are closely related to *foundationalism*, *anti-justificationism*, *conventionalism*, or *relativism* are closely related to what we are calling *anti-foundationalism*.

Table 1 Various Positions in the Philosophy of Science

Positions in the Philosophy of Science	General Epistemological Focus	Criterion of the Philosophy	Representative Philosophers	Validation Approaches
Rationalism		Logical reduction	Descartes	Derived from rational foundation
Classical Empiricism	Logical justification of knowledge claims	Inductive generalization	J. S. Mill J. N. Keynes	Induced from empirical data
Logical Positivism		Empirical verification	Carnap, Russell Wittgenstein	Derived from empirical foundation
Instrumentalism		Predictive success, simplicity, or other aesthetic value	Pierce Friedman	Shown by predictive accuracy, simplicity, or other value
Dogmatic Falsificationism	Theories as frameworks for prediction and testing	"theory-free" observations to test theories	Popper	Continued testing to eliminate faulty models
Methodological Falsificationism		Survival of testing and criticism	Lakatos' version of Popper	Shown by testing and criticism
Bayesianism	Consistent treatment of probabilistic induction	Increase subjective probability	Howson Urbach	Empirical success increasing belief
Kuhnianism	Progressive historical growth of knowledge	Growth of knowledge through Paradigm shifts	Kuhn, Polyani Bohm Weimer	Accordance with expert opinion, professional acceptance
Lakatos' MSRP		Growth of knowledge through Research Programmes	Popper, Lakatos Bartley, Agassi	Increase empirical and theoretical content without ad hoc adjustment
Hermeneutics	Interpretation and understanding through dialog and practice	Knowledge growth by application with participation	Bernstein Gadamer	Participation by all interested in the outcome

Historically, most economists would have classified themselves as foundationalists or objectivists. Two of Naylor and Finger's three philosophical positions on validation (rationalism and empiricism) are foundationalist positions. The third, positive economics, is a version of instrumentalism that can be viewed as either foundationalist or anti-foundationalist.

3.2. Foundationalism in General

A foundationalist believes that there is a unique ultimate basis, either in experience or rational thought, into which a model or theory must be resolvable if one is to validate it (Lakatos 1970). That basis is either to be found in direct experience (empiricism) or in clear, self-evident ideas from one's own mind (rationalism). Empiricist foundationalists like to say that the model must be "verified." By this they mean that it must be tied to an empirical basis. Rationalists do not use the term "verification," but they differ only in claiming another kind of foundation, that of self-evident rational principles. For rigorous empiricists and rationalists, validation is an absolute: no equivocation is allowed by stopping at a half way mark or allowing human judgment or decision to enter the process. The reduction to the foundation may be viewed as a logical process in which one takes higher level propositions of a theory, and reduces them to the constituent elementary propositions of the foundation. The flavor of this approach to empiricism is caught in Naylor and Finger's approving repetition of phrases taken from Reichenbach: "A sentence the truth of which cannot be determined from possible observations is meaningless" (Naylor and Finger 1967, Reichenbach 1951).

3.2.1. Problems with the Classical Foundationalist Positions. At present, empiricist based foundationalism is no longer in vogue, and the problems that have been raised in connection with it are important for simulationists to understand. We mention only two. First, the argument that there is in principle a unique and neutral empirical basis has been seriously questioned and thrown out by most modern philosophers of science. This difficulty has been described in various ways: see for example Weinberg (1960) and Popper's discussion of Fries Trilemma (Popper 1959, Lakatos 1970). Basically, no one has been able to demonstrate through a noncircular logical argument the existence in principle of an empirical or rational basis. Yet even if we did have access to a theory-free empirical foundation and we restricted ourselves to holding only those propositions that had been built up from it, we would still have to deal with the problem of induction. Most would hold that the terms and propositions of a theory are general terms and propositions and as such are always beyond our direct experience. As such, they are analogous to propositions about a population as opposed to a finite sample of observations taken from that population.

On the rationalist side of foundationalism, similar difficulties arise. The "self-evident" basic propositions of one generation of scientists and philosophers are no longer self evident and basic to the next generation. The expansion from Euclidean to non-Euclidean geometry is often cited as an example (Churchman 1973). Despite these problems, we may still find it unavoidable to use rationalist-like locutions in validation. For example, consider the ordinary inventory equation for a nonperishable good:

$$I(t + 1) = I(t) + P(t) - S(t),$$

which simply states that next month's inventory equals this month's inventory plus what is produced minus what is sold. Assuming that there is no other admission or leakage from this inventory other than from production and from sales, the equation is simply a restatement of the principle of the conservation of mass. A pre-Einsteinian rationalist would have argued that this is a presupposition that is not empirically learned. Without assuming it, he or she would have argued, we would not be able to organize our experience of material objects. Many of the relationships in our models may be of this type. It also seems that many of the structures that we associate with mathematical spatial and temporal ordering of our world must be claimed along rationalistic argumentative lines. Kant remains the preeminent philosopher as a source for these arguments. In spite of the difficulties with rationalism, it does seem to have a definite place in warranting some of the basic relationships we include in simulation models.

3.3. Instrumentalism

If one does not accept that general propositions can be justifiably and inductively constructed from direct, empirical observations, then one possible response is that of instrumentalism, (Popper 1959, Boland 1979, Wible 1982), which is to trace the problem to an over-emphasis on the meaning and status of general propositions themselves. In this view, the general propositions of a scientific theory or a simulation model are demoted to the role of convenient arrangements (instruments) that we use to order our observations. For an instrumentalist, the general structures of a model or theory have an unavoidable conventionalistic element. Direct theory-free empirical observations are assumed to be available, and general propositions take on the status of conventionally received organizers of these experiences. The instrumentalist may argue that we choose them for utilitarian, or aesthetic reasons or reasons of convenience more than anything else.

The "positive economics" of Milton Friedman (1953) is an instrumentalism of this type. According to Friedman:

> . .theory has no substantive content . . . its function is to serve as a filing system for organizing empirical material. . . . Only factual evidence can show whether [a theory] is 'right' or 'wrong,' or better, tentatively 'accepted' as valid or 'rejected' . . . the only relevant test of the validity of a hypothesis is comparison of its predictions with experience. (pp. 7–9)

Thus, for Friedman, theory and facts are clearly separable, an idea which he illustrates with the well-known example of the expert billiard player: "It seems not at all unreasonable that excellent predictions would be yielded by the hypothesis that the billiard player made shots as if he knew the complicated mathematical formulas that would give the optimum directions of travel, could estimate accurately by eye the angles, etc., describing the location of the balls, could make lightning calculations from the formulas, and then could make the balls travel in the direction indicated by the formulas."

Friedman argues that the hypothesis that expert billiard players do mathematical calculations is, in effect, a valid "instrument" because it is likely to yield accurate predictions. In effect, who cares if the "assumptions" of the theory are incorrect (i.e. that expert billiard players have mathematical training); the relevant point for Friedman is that expert billiard players act "as if" they do. He argues that we should only be concerned with the predictions of a theory, not its assumptions.

> . . . truly important and significant hypotheses will be found to have "assumptions" that are wildly inaccurate descriptive representations of reality, and in general, the more significant the theory, the more unrealistic the assumptions (in this sense). (p. 14)

Cyert and Grunberg (1963) use this same analogy to argue against Friedman's position. The hypothesis that "expert billiard players solve math-problems" yields two sets of predictions: (1) that expert billiard players can direct the balls as they wish, and (2) that since they have had mathematical training, they should be able to exhibit their mathematical skills in other ways (e.g., by explicitly solving complex mathematics problems). Whereas the hypothesis yields correct predictions in the first instance, it yields wrong predictions with respect to the second.

If the validity of the theory is judged only by the success of its "predictions," then *even by Friedman's own criteria* the theory fails. Yet Friedman argues that the hypothesis is indeed valid in the sense of making reliable predictions—if only we restrict ourselves to the domain where he intends to use it, that is, economic policy making. It is this restriction to the boundaries of economic analysis that leads to the dissatisfaction of Friedman's critics with his instrumentalism. While instrumentalism may be a self-sufficient position in its own terms (Boland 1979), its critics would argue that these seemingly arbitrary exemptions of some subclasses of its predictions from testing make it an unsatisfactory position in the philosophy of science.

So far we have been describing instrumentalism, specifically Friedman's version of it, as conventionalist with respect to theory but as objectivist with respect to the empirical domain in which the theory is required to make predictions. Although most economists would likely put it in that category, we should note that it is possible for an instrumentalist to hold that the empirical basis is also held by convention. In fact, there are places in Friedman's essay where he sways in this direction.

Despite the dissatisfaction one might feel about instrumentalism as a philosophical position, Naylor and Finger were correct in recognizing that the approach does have a place in warranting various structures we might put in simulation models. Most of the systems we simulate have an unavoidable conventionalist element. For example, when simulating the costing side of a manufacturing system, we may want to model the accounting relationships involved in monetary flows. We might choose to aggregate or charge them to certain centers in the model. How we decide to do this is often a matter of a set of assumptions that we should argue for on the grounds of convenience, simplicity, success in prediction, or some other instrumentalist-like criterion.

3.4. Anti-foundationalism in General

As a practical matter, the failure of foundationalism is implicitly recognized by every practitioner in the simulation trade. One of the most common mistakes of beginners is the inclusion of too much detail in their models. The old hands always tell novices to make their models as simple as possible by including only first order effects. But this raises the validation problem. What are the first order effects? How do we justify what we have included as such an effect? How do we make that choice credible? If we accept that the absolute validation required by an empirically oriented foundationalist is out of the window (there are still those who would disagree), then we realize that judgment and decision making cannot be avoided.

Historically, the controversies associated with foundationalist positions have brought about a reappraisal of the entire foundationalist project. The main trend has moved away from arguing that the warrant for scientific theories can be established as a logical activity involving verification or neutral empirical testing. Part of the debate has stemmed from a reappraisal of past successful episodes in science, such as the Newtonian and the Copernican revolutions (Bernstein 1983, p. 13). Other approaches have involved the inclusion of normative, prescriptive, or value oriented considerations into theory legitimation. Still others include an appraisal and a borrowing of ideas from other philosophical traditions outside of Anglo-American philosophy of science.

3.4.1. Methodological Falsificationism. Popper (1959) maintains that there is no solution to the problem of induction in the sense of there being a logic of induction that can be used to construct or justify theories. As such, scientific theories can come from meta-scientific or personal psychological sources. Whatever their origins, the point of scientific activity is still to test theories empirically. Conventionalism comes into this position in determining what is meant by an empirical test. Tests are carried out with reference to basic statements, which are empirically oriented and are agreed upon by the discussants of the theory being tested. In *The Logic of Scientific Discovery*, Popper (1959) uses the metaphor of "piles in a swamp" to explain what he means by these basic statements. He says that, like piles, basic statements are driven down deep enough into the empirical ground to support a theory and that the pounding is stopped when it is decided that they have gone far enough. These propositions are always open for discussion and are not considered to be neutral or theory-free. Thus, according to this reading of Popper, conventionalism enters not at the top of the theory where general propositions are chosen, but at the bottom where the empirically oriented propositions are agreed upon. The ability to test the theory at the bottom—where it meets the data—is crucial. And even though the data are open to interpretation, the agreed upon empirical "observations" are the deciding factor in determining whether the theory is rejected. The only demand that Popper made of a scientific theory was that it be falsifiable; that is, the theory must specify some class of basic statements that—if true—would invalidate it.

What this sophisticated version of Popper seems to value most in methodology is the ability to criticize theories. For him validation was the survival by the model or theory of an open and thorough-going criticism, either in comparison with other models or with respect to whatever agreed-upon data are available.

We should note that there is still a considerable controversy surrounding the works of Karl Popper, specifically whether he is a foundationalist or not. Depending on what is emphasized in his work, one can come out either way. Popper's student, Imre Lakatos (1970), has been the primary source for reading Popper as an anti-foundationalist.

Modern classical statistics is directly related to falsificationism. Classical statistical inference is, in effect, a set of conventions that have been adopted that allow those agreeing with them to make general statements about a population from finite samples taken from it. As such, the statistical methods in the simulation literature for validating models take the falsificationist route in deciding how to handle the problem of induction.

Something akin to the empirically oriented conventionalism that Popper described is a good way to understand what comes into play in the formulation of most simulation models. For example, when we decide to fit a distribution to a certain delay, such as a machining time, we are, in effect, choosing not to explore the logic of that process any further. Instead we agree with others that empirically we have gone far enough in modeling the process. We add the further convention of buttressing our decision with goodness of fit or other statistical tests. If someone disagrees with our decision, we can always model the machining process further, until there is an agreed upon stopping place. This process of forming basic statements in simulation is analogous to the process that Popper described as the questioning and criticism that goes on in establishing a conventionalist empirical basis for scientific activity.

For some simulationists this conventionalist move has long been recognized. For example, Forrester (1961) writes: "Basing decisions on judgment does not mean that the choices lack a foundation of fact or contact with reality. It means rather that a choice is made concerning what part of the available knowledge is to be relied upon" (p. 118). For further discussion of the relevance of Popper's falsificationism to validation, see Herskovitz (1991).

3.4.2. Kuhnianism. Most readers of T. S. Kuhn's book, *The Structure of Scientific Revolutions*, do not see it placed in the controversy between foundationalism and conventionalism that we have been describing. Instead it is usually read in isolation as a history or sociology of science. The normal foundationalist description of the origin of scientific theories is that they are induced from the accumulation of data obtained from empirical observations. Kuhn argued that this account of the genesis of scientific theories cannot be squared with what happened historically. Instead of the accumulation of neutral theory-free observations, scientific theories derive from a gestalt, a set of exemplars, or what he called a paradigm. As a pattern, the paradigm serves as an implicit guide in defining the relevant data and making them coherent with respect to the theories developed around them. This is also true with respect to the methodology in any given area of science; that is, the paradigm of the field serves as a guide to what is the correct methodology. As the field develops, the sense of the paradigm becomes explicit in methodological rules and models. Historically, Kuhn described paradigms as going through a life cycle of genesis, normal activity and finally rejection as anomalies arise and new paradigms develop. (For further reading, see Sterman 1985, and Wittenberg 1990.)

For Kuhn, validation was "a complex process with social, psychological and historical dimensions" (Sterman 1996, personal communication). Part of this process involves persuading others that one's model falls into a well-accepted way of seeing the problem. Again, as in Friedman's case, past success may be a part of the appeal. In the field of simulation, Kuhn's approach is the kind of defense that we invoke when we call an expert to check our simulation model. For Kuhn, an "expert" is a recognized adept in a reigning paradigm. Thus arguing for the "face validity" of a model (Sargent 1992) is a validation move that is consistent with the Kuhnian perspective, in that sense. The expert comparison of an actual system with the output of a simulation model, as in Turing's test (Schruben 1980), is also consistent with this perspective.

The criteria that may determine the adequacy of the model involve values, such as consistency, fruitfulness of the theory or model, and the like (Kuhn 1970, p. 199). The invocation of these values is, in effect, an appeal to others to make a value-based agreement on the adequacy of the model.

Kuhn's assertions regarding the nature of scientific development have come under heavy criticism, and these criticisms have important implications for the problem of validation in simulation. Some have seen his account of scientific development as including an arbitrary relativistic element that leaves out the question of scientific validity altogether. Kuhn refused to let himself be called a relativist. Instead he maintained that there can be reasoned and persuasive argument for a scientific position without committing oneself to some version of foundationalism.

3.4.3. The Methodology of Scientific Research Programs (MSRP). Imre Lakatos sought to put together a position that lies somewhere in between Kuhn's alleged relativism and Popper's empirically oriented falsificationism (Lakatos 1970). Like Kuhn, Lakatos thought of a scientific movement as possessing a central core or paradigm-like structure, which he called the hard core of the movement or theory. Surrounding this core is a belt of problems that are attacked and developed in a series of theoretical steps over a period of scientific growth, which he called the *research program*.

Lakatos described what he claimed were the decisions underlying the formulation and testing of scientific theories. These decisions were Lakatos's explicit recognition that scientific theorizing is not built on a so-called "objective" basis but rather involves a value-based set of criteria.

One of Lakatos's most interesting discussions deals with the scientist's decisions about *ceteris paribus* assumptions, the boundary conditions that surround a given theory. An important objection to falsificationism is that many long standing theories of science are not falsified by observational statements, although participants may hold these statements to be infallibly true. A researcher can always save a favored theory from an empirical result at odds with it by invoking an auxiliary hypothesis. Lakatos (1970) has given a detailed account of this kind of move and also of the so-called Duhem-Quine thesis with which it is usually associated.

As an example of this phenomenon, consider the following:

> Lorenz tried to save Newtonian cosmology in the face of the Michelson-Morley experiment by invoking the auxiliary hypothesis that measuring rods shrink in the direction of motion

through the ether. Sometimes the introduction of an auxiliary hypothesis is a rewarding path, as when the data suggested that Uranus did not follow the orbit predicted by Newton's law of gravity. Here the auxiliary hypothesis was to suggest that a more distant planet was perturbing the orbit, preserving the inverse square law. Of course, subsequent observation revealed the presence of Neptune. Other times, auxiliary hypotheses serve to insulate a theory from confrontation with unfavorable data. (Sterman 1996, personal communication)

As an example (adapted from Howson and Urbach 1989, pp. 106–107), Immanuel Velikovsky, in his controversial book *Worlds in Collision* (1950), suggests that a near collision with a massive comet was responsible for the ten plagues and the parting of the Red Sea. When asked why communities of the time failed to record of these tremendous happenings, he argued that the catastrophe was so terrifying that it led to a "collective amnesia."

Lakatos tried to lay out a prescriptive model to avoid these ad hoc dodges; this is his MSRP. He argued that progressive development avoids adjustments made purely in order to rationalize away empirical problems. Each development in a research program is successful if it includes greater empirical content in the domain that has been or will be explained. This appeal to an empirically oriented adequacy is one of the elements that allies him with Popper's falsificationism. Unfortunately, Lakatos's anti-foundationalism—the nonneutrality of empirical content—came back on his own prescriptions with a vicious kick. One person's empirically progressive development can be another's arbitrary and fictional myth. Paul Feyerabend (1988) was especially active in describing this shortcoming in Lakatos's position.

In simulation, we can find ourselves in the same situation when determining where to draw the boundaries around a model, that is, in deciding what are its surrounding conditions and how far to extend them. It is not uncommon for those unfamiliar with simulation to make extensive demands for the inclusion of factors in the model. A foundationalist bias on the part of a model evaluator can increase this pressure; for such an observer "every fact must be present." Without some sort of wider epistemological perspective such as we are advocating here, the simulationist may have difficulty handling these demands.

Lakatos' MSRP has evoked much discussion and support among economists (e.g., Blaug 1976, Caldwell 1991, Howson 1976, and Latsis 1976).

3.5. The Bayesian Approach.

The philosophies of science that we have discussed thus far largely restrict themselves to deterministic theories of science. But many scientific theories, including simulation models, are explicitly probabilistic in nature, and for this reason, are impossible to validate in an empiricist fashion. As an example (from Howson and Urbach 1989, p. 7), Mendel's theory of inheritance states the probabilities with which certain genes occur during reproduction. The theory therefore does not preclude or predict a specific gene configuration. Nevertheless, Mendel "confirmed" his theory by conducting pea plant experiments and obtaining results that his theory had predicted to be highly probable. The methods by which these statistical or nondeterministic theories are assessed are normally treated as a branch of statistics, although they belong more naturally to the philosophy of science.

Much of what is standard today in classical statistics was inspired by the work of Fisher, and later, Neyman and Pearson. Howson and Urbach (1989), in *Scientific Reasoning: The Bayesian Approach*, point out several inconsistencies in how these statistical methods deal with probabilistic induction. The Bayesian approach is concerned principally with dealing with the problem of probabilistic induction (of which the deterministic version is a special case) in a consistent way. While some opponents of Bayesianism see it as the last holdover of foundationalism (Giere 1984, p. 336), most Bayesians would probably not consider themselves foundationalists. Howson and Urbach, for example, plainly accept the thesis that empirical observations are not indubitable givens but rather fallible theoretically oriented posits (1989, p. 94). As constructed from observations, subjective probabilities may be accepted conventionalistically as measures of the credibility of general propositions and theories. Howson and Urbach have shown how some of the classical problems involved in the separation of theories from background knowledge may be worked out in Bayesian terms. In this sense, the belief (expressed as subjective probabilities) in the paradigm or research program of simulation is increased with every successful implementation of a simulation project.

4. Objectivism Versus Relativism in Recent Philosophy of Science

In §2, we argued that much of the current dialogue on validation in simulation can be understood in relation to two "poles": the objectivism advocated by Naylor and Finger (1967) and the relativism set forth by Barlas and Carpenter (1990). Yet while this debate may be new to the simulation arena, its origins in philosophy date as far back as ancient Greece, when Socrates argued against the alleged relativism of Protagoras (Plato, *The Protagoras;* see also Bernstein 1983, p. 8).

One of the subplots of the philosophy of science that the reader of this paper may have already picked up is the extent to which each of our later philosophers (Friedman, Popper, Kuhn, etc.) appears to wobble on both sides of the *foundationalist / anti-foundationalist* distinction. It is common for this *either / or* debate to beget a kind of epistemic anxiety that draws one magnetically into one pole or the other. Bernstein (1983) uses the label "Cartesian Anxiety" for the discomfort involved in this pull. He traces it back to Descartes's attempt to find a rational foundation, and the doubts arising from that attempt. In the last two decades, however, various philosophers have argued that there is something wrong with this *"either / or"* debate and have begun to hold that the polarity of these categories begets unproductive, dead-end arguments that structurally feed off each other. The current position of many of these philosophers is well-captured by the quote from Oliver Wendell Holmes that appears at the beginning of this essay.

Richard Rorty (1980, pp. 727–728) put the contrast between objectivism and relativism in these terms:

> "Relativism" is the view that every belief on a certain topic, or perhaps any topic, is as good as every other. No one holds this view. . . . The philosophers who get called "relativists" are those who say that the grounds for choosing between opinions are less algorithmic than had been thought.

In the last decade and a half, several philosophers have attempted to break out of this debate. To our knowledge, the first philosopher to do this explicitly was Bernstein (1983), who writes (1983, p. 23):

> One reason why these controversies seem to generate more heat than light is that the entire discussion is still infected with the legacy of the Cartesian *Either / Or;* many of the participants in these disputes argue as if we must choose between the alternatives of objectivism (e.g., scientific realism) or relativism. But this way of framing the key issues is misleading. We gain . . . when we appreciate that what is really going on is that the whole framework of thinking that poses issues with reference to these and related dichotomies is being called into question.

Further on (p. 166), he writes of this dichotomy:

> Relativism ultimately makes sense (and gains its plausibility) as the dialectical antithesis to objectivism. If we see through objectivism, if we expose what is wrong with this way of thinking, then we are at the same time questioning the very intelligibility of relativism.

Bernstein argues that progress of philosophy has been hampered by the framework which posits validation as an *"either / or"* proposition. See also Martin and Kleindorfer (1991) for a depiction of the logical structure of this debate; specifically, how each pole complements and implicitly assumes the epistemic presuppositions of its opposite.

5. Hermeneutics

Twentieth century Anglo-American philosophy of science has had a largely analytic orientation; that is, its main concerns were with working out the problems connected with the foundationalist project that we outlined in the earlier parts of this paper. The radical questioning of that project and the corresponding Cartesian Anxiety—as Bernstein would call it—energized the debate among Anglo-American philosophers. More recently these discussions have made connections with various continental philosophical positions and with the movement of philosophical hermeneutics in particular.

The origin of the modern hermeneutics is frequently placed in nineteenth century studies of textual interpretation. There is no doubt that Martin Heidigger is the most seminal twentieth century thinker associated with the movement. Hans Georg Gadamer, the principle heir to the Heidiggerian hermeneutic tradition, in his *Truth and Method* (1975), inventively elaborated on Heidigger's hermeneutics to produce among other things a fuller exploration of its epistemological side. The two most influential Anglo-American scholars who have made the connection to Gadamer are Richard Bernstein in *Beyond Objectivism and Relativism: Science, Hermeneutics and Praxis* (1983), and still earlier Richard Rorty in *Philosophy and the Mirror of Nature* (1979). T. S. Kuhn (1977, p. xiii, and p. xv) discussed his own reading of hermeneutics and described it as "decisive" on his view of science.

In *Beyond Objectivism and Relativism*, Bernstein presents Gadamer's hermeneutics as a philosophical fulcrum for, among other things, getting beyond the polarity of the foundationalist versus anti-foundationalist debate. Bernstein delineates a description of rationality that is historically situated and practical, involving choice, deliberation, and judgment. Bernstein argues that in a certain sense both the foundationalists and the anti-foundationalist sides are right and he develops his argument by using an interesting set of metaphors borrowed from Gadamer. One of these metaphors is the activity of play. Returning temporarily to the terms of the polarity, Bernstein and Gadamer describe play—such as exists in a game—as possessing both a subjective or relativistic element that each player brings to the game in the form of his or her viewpoints, abilities, past experiences, and personal idiosyncrasies. It is only through this peculiar contribution by each player that the game comes into existence as an entity separate in some sense from the players in it. This separate existence of the game in effect involves the reach for an objectivity—again casting it in polar terms—over and above the individual viewpoints of the players. Extending this metaphor to scientific activity, he argues that the practice of science involves just such a mix of relativistic and objectivistic perspectives. He attempts to show that the arguments in the philosophy of science, most notably those between the various positions we have discussed so far, may be co-opted into this more general perspective.

Play is not the only metaphor that Bernstein or Gadamer use in their attempts to understand how human knowledge comes about. In *Truth and Method*, Gadamer takes up other metaphors from art, from drama, from the reading of texts, and from language itself, in order to show that the extreme polarities of objectivism and relativism have warped our understanding of human knowledge. In sum, the aim of the move beyond objectivism and relativism is ". . . just enough constraint to guarantee stability and lend meaning to individual purpose yet not so much as to stifle the possibility of autonomous and maybe creative action." (Bernstein 1983, p. 84)

Rationality, in Gadamer's view, is constituted so that understanding, interpretation, and application are simultaneous. This simultaneity in cognition is called the "hermeneutical circle," and is one response to the problem of induction. In simulation, it describes what we have all experienced: that there is a continual play back and forth whereby our understanding of general principles is increased as we interpret the particulars in a given application. In light of that understanding, we simultaneously begin to see the particulars more sharply and are better able to give them meaning. Furthermore, there is the recognition that "playing" with a theory or simulation model is a way of effecting its validation. The interaction between the modeler and the client in mutually understanding the model and the process establishes the model's significance; that is, its warranty.

6. Implications for Validation in Simulation

Whereas the philosophy of science has largely begun to move away from Cartesian anxiety and the presupposition of the "*either / or*," this legacy still pervades the discussion on validation in simulation. Consider the following passage from Oreskes et al. (1994, p. 643) regarding the validation problem in the earth sciences:

> The central problem with the language of validation and verification is that it implies an *either / or* situation. In practice, few (if any) models are entirely confirmed by observational data, and *few* are entirely refuted.

An extreme objectivist believes that model validation can be divorced from the model builder and its context. He or she maintains that models are either valid or invalid, and that validation is an algorithmic process which is not open to interpretation or debate. By contrast, an extreme relativist believes that the model and model builder are inseparable. As such, all models are equally valid or invalid and model validity is a matter of opinion. While these two poles appear antithetical, there is a sense in which they are the same. In the former case, meaningful dialogue is stifled by an appeal to the foundation; in the latter, it is suppressed by asserting whatever the current opinion happens to be. At bottom, neither pole is satisfying or tenable, since, at both ends, the model is spared a thorough, external critical review, and the model builder is let off the hook. Our purpose here is to place the model builder back on the hook.

At the same time, both poles of the debate are right in a sense. Objectivism seeks a common framework with which to evaluate and compare models and a sense in which the validation process transcends the model builders and users. By contrast, the relativist position highlights the need for a dialogue between the model builder and other model stakeholders. According to Barlas and Carpenter (1990), validation is "a matter of social conversation rather than objective confrontation."

We would argue that most practitioners have instinctively adopted a middle ground in this debate. For example, while acknowledging the importance of empirical (which he termed "scientific") approaches to validation, Carson noted that, "Validation cannot be carried out by the modeler alone . . . Communication with the client plays a large role in building a valid model and establishing its credibility" (1989, p. 552).

In relation to scientific theorizing in general and simulation in particular, the hermeneutic position would assert that the validation of a model can be achieved in any reasonable manner. What precisely is meant by reasonable is that, through historically situated dialogue, judgment, and practical discourse, we are able to discern the difference between the good and the bad, the worthwhile and the frivolous, the "true" and the "false." Also this practical judgmental and interactive orientation brings in an ethical dimension to scientific validation. Both Bernstein and Gadamer throughout their work relate this view back to *phronesis*, the term that Aristotle used for "practical wisdom," that is, the wisdom that is called into play in council on practical affairs.

As an example of a framework for validation in simulation that would be consistent with Bernstein's hermeneutics, we look to the court system. To obtain a conviction, a prosecutor does not have to prove the guilt of the defendant in any foundationalist sense, but rather "beyond a reasonable doubt." This opens the legal system to the charge of relativism, since each juror's definition of "reasonable" is presumably a function of his or her own biases and prejudices. Yet this charge of relativism is unfounded. By and large, it is the merits of the case as defined within the parameters of the law that determine a trial's outcome.

If we extend the court metaphor to simulation, then we must assume an *openness* denied by both objectivists and relativists in which we can conduct meaningful dialogue on a model's warrantability. This openness would imply an ability to meaningfully compare different models. The model builder or builders would be free to establish and increase the credibility of the model through any reasonable means. This process would also involve other model stakeholders, such as model users and referees of journal articles, who would share part of the responsibility of effecting model validation. This framework would not preclude any of the positions we have outlined earlier. Rather, there are situations where appeals to each of these positions would be called for. Thus, the purpose should be to lend just enough structure to provide stability and lend meaning to questions of validation, yet not so much as to diminish the importance of individual freedom and ethical behavior in model validation.[1]

[1] The authors would like to thank the associate editor and two anonymous referees, whose thoughtful criticism led to major improvements in the paper.

References

Barlas, Y., S. Carpenter 1990. Philosophical Roots of Model Validation: Two Paradigms. *Systems Dynamics Review* **6** 48–166.

Bernstein, R. J. 1983. *Beyond Objectivism and Relativism: Science, Hermeneutics, and Praxis*. University of Pennsylvania Press, Philadelphia, PA .

Blaug, M. 1976. Kuhn versus Lakatos or Paradigms versus Research Programmes in the History of Economics. S. Latsis, ed. *Method and Appraisal in Economics*. Cambridge University Press, Cambridge, UK91–108 .

Boland, L. A. 1979. A Critique of Friedman's Critics. *J. of Economic Literature* **17** 503–522.

Caldwell, B. 1991. Clarifying Popper. *J. of Economic Literature* **29** 1–33.

Carson, J. S. 1989. Verification and Validation: A Consultant's Perspective. E. A. MacNair, K. J. Musselman, and P. Heidelberger eds. *Proc.* 1989 *Winter Simulation Conf.* 552–557.

Churchman, C. W. 1973. Reliability of Models in the Social Sciences, *Interfaces* **4** 1 1–12.

Cyert, R. M., E. Grunberg 1963. Assumption, Prediction, and Explanation in Economics. R. M. Cyert and J. G. March, eds. *A Behavioral Theory of the Firm*, Prentice Hall, Englewood Cliffs, NJ, 298–311.

Feyerabend, P. 1988. *Against Method* (revised edition). Verso, New York.

Forrester, J. W. 1961. *Industrial Dynamics*. MIT Press, Cambridge, MA.

Friedman, M. 1953. The Methodology of Positive Economics. *Essays in Positive Economics*. University of Chicago Press, Chicago, IL, 3–43.

Gadamer, H. G. 1975. Truth and Method. 4th ed. Seabury Press, New York.

Giere, R. N. 1984. *Understanding Scientific Reasoning*. 2nd ed. Holt, Rinehart, and Winston, New York.

Goldberg, J., R. Dietrich, J. Chen, T. Valenzuela, E. Criss 1990. A Simulation Model for Evaluating a Set of Emergency Vehicle Base Locations: Development, Validation, and Usage. *Socio-Economic Planning Science* **24** 125–141.

Herskovitz, P. J. 1991. A Theoretical Framework for Simulation Validation: Popper's Falsificationism. *International J. of Modeling and Simulation* **11** 56–58.

Holmes, O. W., Jr. 1996. Holmes and Frankfurter. University Press of New England, Hanover, New Hampshire. *The New York Times Book Review*, 15 Dec.

Homer, J. 1996. Why We Iterate: Scientific Modeling in Theory and Practice. *System Dynamics Rev.* **12** 55–80.

Howson, C. 1976. *Method and Appraisal in the Physical Sciences: The Critical Background to Modern Science*, 1800–1905. Cambridge University Press, Cambridge, U.K.

——, P. Urbach 1989. Scientific Reasoning: The Bayesian Approach. *Open Court*. La Salle, IL.

Kuhn, T. S. 1970. *The Structure of Scientific Revolutions*. 2nd ed., University of Chicago Press, Chicago, IL.

——, 1977. The Essential Tension: Selected Studies in Scientific Tradition and Change. University of Chicago Press, Chicago, IL.

Lakatos, I. 1970. Falsification and the Methodology of Scientific Programmes. Lakatos, I., Musgrave, A., eds. *Criticism and the Growth of Knowledge*, Cambridge University Press, Cambridge, U.K., 91–196.

Law, A. M., W. D. Kelton 1991. *Simulation Modelling*, 2nd ed. McGraw-Hill, New York.

Latsis, S. 1976. *Method and Appraisal in Economics*, Cambridge University Press, Cambridge, U.K.

Martin, J. E., G. B. Kleindorfer 1991. The Argumentum Ad Hominem and Two Theses About Evolutionary Epistemology: 'Godelian' Reflections. *Metaphilosophy* **22** 63–88.

Naylor, T. H., J. M. Finger 1967. Verification of Computer Simulation Models. *Management Sci.* **14** B92–B101.

Oreskes, N., K. Shrader-Frechette, K. Belitz 1994. Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science* **263** 641–644.

Popper, K. 1959. *The Logic of Scientific Discovery*. Harper and Row, New York.

Reichenbach, H. 1951. *The Rise of Scientific Philosophy*. University of California Press, Berkeley, CA.

Richardson, G. P., A. L. Pugh III 1981. *Introduction to System Dynamics Modeling with DYNAMO*. MIT Press, Cambridge, MA.

Rorty, R. 1979. *Philosophy and the Mirror of Nature*, Princeton University Press, Princeton, NJ.

Sargent, R. G. 1992. Validation and Verification of Simulation Models. J. J. Swain, D. Goldsman, R. C. Crain, J. R. Wilson, eds. *Proc.* 1992 *Winter Simulation Conf.* Arlington, Virginia, 104–114.

Schruben, L. W. 1980. Establishing the Credibility of Simulations. *Simulation* **34** 101–105.

Sterman, J. D. 1985. The Growth of Knowledge: Testing a Theory of Scientific Revolutions with a Formal Model. *Technological Forecasting and Social Change* **28** 93–122.

—— 1996. Personal communication.

Velikovsky, I. 1950. *Worlds in Collision*, Victor Gollancz, London.

Weinberg, R. J. 1960. *An Examination of Logical Positivism*, Harcourt, Brace and Company, New York.

Wible, J. 1982. Friedman's Positive Economics and Philosophy of Science. *Southern Economic J.* **49** 350–360.

Wittenberg, J. 1990. On the Very Idea of a System Dynamics Model of Kuhnian Science. *Systems Dynamics Review* **8** 21–33.

Accepted by Pierre L'Ecuyer; received June 2, 1994. This paper has been with the authors 13 months for 3 revisions.