THE PHILOSOPHY OF SCIENCE AND VALIDATION IN SIMULATION

George B. Kleindorfer
Ram Ganeshan

Department of Management Science and Information Systems
The Pennsylvania State University
University Park, PA 16802. U.S.A.

ABSTRACT

More than twenty-five years ago, Naylor and Finger suggested that the problem of validation in simulation was analogous to the problem of validating scientific theories in general. They went on to prescribe an eclectic approach to validation in simulation that they put together from what they viewed at the time as an exhaustive description of the possible philosophical alternatives. A considerable development has taken place in the philosophy of science since Naylor and Finger wrote their paper. Most notably the justificationist positions in the philosophy of science that Naylor and Finger appealed to have been largely discredited. We attempt here to provide a new examination of the various relevant positions. And we also attempt to show in one way or another how these positions provide additional perspectives on overcoming some of the conceptual difficulties involved in simulation validation.

1 INTRODUCTION

In their classic article on validation, (Naylor and Finger, 1967) first called attention to the relationship between the problem of validating computer simulation models and the problem of validating scientific theories in general. They outlined three different positions in the philosophy of science on the problem of scientific validation and they argued that these positions were equally as relevant to validation in simulation. Empiricism, rationalism and positive economics were the names that they attached to these positions. At the present time, these positions are still being put forward in the simulation literature as exhaustive. They have come to be called the "historical" approaches to validation (Sargent, 1992).

We would like to revisit this subject. In our view, the limitations involved in the positions outlined by Naylor and Finger have been the sources of some of the problems in dealing with simulation validation. In the philosophy of science and in the arguments surrounding methodological questions in economics, weighty and interesting arguments have been advanced for other ways of looking at the problem. In the past two and a half decades, the philosophy of science has developed considerably. Our aim in this piece is to suggest some of the ways in which the main currents in this literature have a bearing on different ways of viewing the general principles of simulation validation.

Why is it that the burden of solving the validation problem seems to be singularly placed on the designers of computer simulation models? Every application paper involving simulation is expected to include a section on validation. The same requirement does not seem to have been as rigorously placed on other kinds of exercises in modeling. For example, one seldom finds in a paper on the application of optimization techniques a section specifically titled "validation." Does the word "simulation" itself evoke the idea of comparison immediately? To simulate means to build a likeness and the question as to the accuracy of the likeness, one version of the validation problem (some might argue the only version), is never far behind. The validation problem 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 logistical system works. As such, Naylor and Finger were right in bringing forward the question of the validity of these models in the same terms that the question has been raised for scientific theories in general.

In many fields, the validation problem has been more actively discussed than it has in management science. In economics, for example, there is a long standing tradition of argument about and a vast literature on the justification of economic methodology itself (See for e.g., Caldwell, 1991, Friedman, 1953,
and Wible, 1981). Naylor and Finger appeal to this tradition in economics and many of the sources they used came directly from that literature. But economics is not the only example: the same problem has been raised in most of the main areas of the physical and social sciences (See for e.g., Bernstein 1983, Feyerabend 1988, Kuhn, 1970, and Popper, 1953). The philosophy of science is the name that has been applied to this historical discussion.

If one cuts out the sections on validation from any sample of simulation papers, one is immediately struck with the wide variation to be found. There will be recitations about model behavior, success in application, reservations and restrictions, personal experiences, description 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 of several problems. For one thing, there is still some confusion about what validation involves. And in addition the restrictions of an empirically oriented validation (for a detailed account on empirical validation, see Law and Kelton, 1991, and Sargent, 1992) cannot be met in many applications. This happens quite often in practice. A designer is asked to simulate a set of scenarios for a prospective system. One cannot empirically or statistically validate a model for a nonexistent system. How can one discuss validation in such a case? This is one of the questions on which we hope to shed some light in our consideration of the arguments from the philosophy of science.

2 THE VALIDATION PROBLEM AND THE PROBLEM OF INDUCTION

The fundamental difficulty underlying the validation of both simulation models and scientific theories has to do in one way or another with the problem of induction. The explicit focus on the problem of induction seems to have arisen primarily out of one particular direction in scientific validation: the desire to justify a theory solely on empirical grounds. Since modelers and theorizers have access to only their own peculiar and limited set of direct experiences, the problem of justifying generalizations beyond that limited personal empirical domain arises. In simulation we wish to infer from our experience of a system that the model that we produce captures the essential logic and parameters of the system and we want to do this based on a limited set of experiences with the system. This is the problem of induction in simulation.

Historically in the philosophy of science, there seems to have been and are now various forms of three basic approaches in discussing the problem of scientific validation. These have been variously called: (1) justificationism or foundationalism or objectivism, (2) antijustificationism or conventionalism or relativism, and (3) various attempts to cut through the middle between the first two positions. As the terms suggest, justificationism and antijustificationism are usually seen as "either-or" positions. Recently explicit attempts have been made in the philosophy of science to overcome this polarity (Bernstein, 1983): hence, our third category. Almost every important philosopher of science, while professedly taking either a justificationist or an antijustificationist stance, ameliorates or incorporates into his or her position some way or other of overcoming the extremes of the polarity between the first two fundamental positions.

At least up until twenty-five years ago most economists would have classified themselves as justificationists or at least they would have approved of being called objectivists. Of Naylor and Finger’s three philosophical positions on validation, the first two (empiricism, and rationalism), at least as these authors describe them, are justificationist positions. Their third position (positive economics), they present as if it were a justificationist or objectivist position, but a careful reading of Friedman (1953) from which they derive it shows that he was keenly aware of the problem of induction and problems with justifiacinist itself. We would argue that his own position is antijustificationist in its primary thrust although there are those who would disagree with us on this point (Boland, 1979). We will discuss this further below.

3 JUSTIFICATIONISM

A justificationist believes that there is a unique ultimate basis either in experience or in rational thought into which a model or a theory must be resolvable if one is to validate it. Depending on whether one is an empiricist or a rationalist, this basis is either to be found in direct experience, or in clear, certain, self-evident ideas found in one’s own mind. Validation consists of reducing or tying the propositions of the model or theory to the basis and only the basis whether it be empirical or rational. Empiricist foundationalists like to say that the model must be “verified.” This is their code word for the requirement that the model or theory be reducible to the empirical basis. Rationalists have not used the term “verification” but they differ only in claiming another kind of foundation, that of self-evident rational principles, to which a theory must be reduced if it is to be
validated. For rigorous empiricists and rationalists, validation is in principle an absolute: no equivocation is allowed by stopping at a half way mark or at a position that involves any compromise introduced by human judgment. The reduction to the foundation may be viewed as a logical process in which one takes higher level propositions of a theory, and breaks them into 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, and Reichenbach, 1951).

At the present time, empiricist based justificationism is no longer in vogue. And the basic problems with it are very important for simulationists to understand. It faces several difficulties of which we will only mention two here. 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). Basically no one has been able to demonstrate through a noncircular logical argument that there is in principle an empirical or rational basis.

But even if we did have access to a theory-free empirical foundation and if we did restrict ourselves to holding only those propositions that had been built up from it, we would still have to deal with difficulties with the problem of induction. The propositions of a theory or a model are general propositions that are always beyond our particular direct experience: they are analogous to propositions about a population in contrast to a finite sample of observations taken by us from that population.

On the rationalist side of foundationalism, similar difficulties have been encountered. 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. This is not to say that some aspects of rationalism may be at least eligible for use in the validation of simulation models. Take as an example, the ordinary inventory equation:

\[ I(t + 1) = I(t) + P(t) - S(t). \]

The inventory next month is the inventory this month plus what is produced minus what is sold. An old style rationalist would have argued that this is a presupposition and not empirically learned. Without assuming it, he or she would argue that we would not be able to organize our experience of material objects. Many of the relationships that we include in our models may be of this type and rationalist-like elucutions certainly seem sensible for justifying their inclusion.

4 ANTIJUSTIFICATIONISM

The failure of justificationism has consequences that every simulationist learns sooner or later in practicing of our trade. When we frame a simulation model there is always a considerable element of human judgment involved, for example, in knowing where to stop. One of the most common mistakes that beginners make is the inclusion of too much in their models. We always tell them that is important to keep their models as simple as possible and to include only first order effects. But this raises the validation problem. What are the first order effects in the situation being modeled? How do we justify what we include as such an effect? How do we make that choice credible? Once we understand that the absolute validation required by an empirically oriented justificationist is out of the window, we realize that judgement and decision-making with respect to these problems cannot be avoided. In that regard, it is interesting to see the various ways in which this problem has been dealt with in the philosophy of science.

4.1 Instrumentalism

If one does not accept or at least doubts that general propositions can be legitimately and inductively constructed from individual direct empirical observations, then one possible response, that of instrumentalism, 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 theory or a model are demoted to be merely organizers or convenient arrangements or instruments that we use to order our observations no matter whether the modeled system be scientific or technical like a simulation model. An instrumentalist would say that the general structures of a model or a theory have an unavoidable conventualistic element. He or she would go on to say that we choose them for aesthetic reasons or reasons of convenience more than anything else. But in looking at them in this instrumentalist fashion, we acknowledge that it is not possible to empirically justify why we put them in any certain form.

There are two possible versions of this instrumentalist approach. One has a justificationist flavor. For this position, direct theory-free empirical observations are still assumed to be available. Another
version holds that theory and fact are not separable; that is, facts are said to be "theory laden." This latter form of instrumentalism is antijustificationist not only with respect to theory but also with respect to the facts themselves.

The "positive economics" of Milton Friedman is an instrumentalism of the latter type although Naylor and Finger, oriented as they are toward empirical verification, lean toward portraying it as being justificationist in spirit. That Friedman is an antijustificationist seems clear to us in statements like the following: "A theory is the way that we perceive 'facts,' and we cannot perceive 'facts' without a theory" (Friedman, 1953, p. 34). As to the general assumptions underlying the theory of the firm in economics, Friedman holds that it is pointless to attempt to verify these general assumptions empirically. This is the direction in Friedman's economic methodology that has been characterized as instrumentalist (Boland, 1979, 1980, Del Alessi, 1968, Friedman, 1953, Koopmans, 1957, and Rotwein, 1980).

Friedman emphasizes that success in economic prediction is the criterion to be used in judging a theory and not the realism of the assumptions themselves (Friedman, 1953). This is the instrumentalist part of Friedman's dictum on economic modeling that is most often cited. But there is a definite problem here as well as we are sure that Friedman as an antijustificationist would acknowledge. Predictions themselves are not neutral theory-free facts. Their status and meaning on the output side of the model are even more theory determined than the "facts" on the input side of a model. The appeal to success in prediction as a basis for validating a model is just as circular as any appeal to a supposed neutral empirical basis. To settle on this appeal and to decide to accept it is a conventionalistic agreement, not a directly observable value-free fact.

No matter how we decide to characterize Friedman's position in particular, the instrumentalist position in general always leaves a trail of dissatisfaction in its wake. The abandonment of the realism of assumptions is the source of the discomfort. And for many the validation problem is centered on this question. To be fair to Friedman, there are places in his essay where he indicates that what might be construed as his antirealism does not capture his position. For example, in the conclusion of his essay he speaks of progress in economics as expanding the "body of generalizations, strengthening our confidence in their validity, ..." (Friedman, 1953, p. 38). There are other places where he makes an appeal to a criterion of validity that seems to go beyond a narrowly conceived instrumentalist position.

Despite whatever dissatisfaction we might feel about instrumentalism as a general philosophical position, it seems to us that the approach has a definite place in connection with validation in simulation. Most of the systems that we simulate have an unavoidable conventionalistic element and for parts of these models our arguments should go along instrumentalist lines. For example, when we are 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 or simplicity or some other such instrumentalist-like criterion.

### 4.2 Falsificationism

Falsificationism is another response to the questionable empirical status of general propositions in scientific theorizing. Sir Karl Popper with whom it is usually associated classifies it as a form of conventionalism (Popper, 1959, pp. 53-56). Again there are justificationist and antijustificationist versions of this position although Popper himself in his Logic of Scientific Discovery has given some of the most powerful arguments against justificationism and has himself explicitly adopted an antijustificationist stance.

First of all, for a falsificationist there is no answer in methodology as to where general propositions themselves come from. Popper himself believes 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. For Popper, scientific theories may come from meta-scientific or personal psychological sources. Whatever their origins, the point of scientific activity is to test theories empirically. Conventionalism comes into this position in determining what is meant by an empirical test. For Popper the tests are carried out by reference to what he calls "basic" statements. These are empirically oriented statements that are agreed upon by the discussants of the scientific theory. They are not basic in a justificationist sense. Popper, these propositions are always open for discussion and they are acknowledged not to be neutral, or theory-free.

For Popper conventionalism comes into falsificationism not at the top of the theory where the general propositions making it up are chosen but rather at the bottom of the project where the basic 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 is open to inter-
pretation, the agreed-upon empirical "observations" are the deciding factor in determining whether the theory is rejected or not.

Modern classical statistics is directly related to the falsificationism that we have been describing. Classical statistical inference is in effect a set of conventions that have been adopted for making statements about a population from finite samples from it. As such the statistical methods in the simulation literature for validating models should be classified as falling under falsificationism in the philosophy of science.

It is relevant to ask here how a falsificationist would handle the problem of validation if there were no empirical data available. In a more general sense, what Popper values in methodology is the ability to criticize our theories. This is why we subject scientific theorizers to the requirement that their theories be operationalizable in so-called basic statements. In a general sense, then, Popper would ask the designers of theories and models to put them in a shape that makes them criticizable. For Popper validation is the survival by the model or theory of such an open and thorough going criticism even if that criticism cannot be grounded in a fully empirical manner. Popper would ask that models be constructed in such a way that they are fully accessible to critics, and validation for him would be the result of explicit intersubjective criticism.

4.3 Kuhnianism

T.S. Kuhn's book, The Structure of Scientific Revolutions, has become so well known that to bring it up has almost become trite. But most readers of it do not see it placed in the controversy between justificationism and conventionalism that we have been describing in this essay. Instead the book is usually read in isolation as a history or sociology of science.

The normal empiricist justificationist 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 can not be squared with what happened historically. Instead of the accumulation of neutral theory-free observations, Kuhn argued that scientific theories derive from a gestalt, a set of exemplars, or what he calls a paradigm. As a pattern, the paradigm serves as an implicit guide in defining the relevant data and making it coherent with respect to the theories developed around the paradigm. This is also true with respect to the methodology in any given area of science; that is, the paradigm of the field serves as an implicit guide as to what is in fact correct methodology. As the given field develops what was implicit in the paradigm becomes explicit in the methodological rules and models that result. Historically, Kuhn describes paradigms as going through a life cycle of genesis, normal activity and finally rejection as difficulties are not met and new paradigms are invented.

Discrete event simulation, for example, may be viewed as a particular paradigm for modeling industrial and logistical systems. According to this paradigm, such systems are visualized as a set of events that take place at instants of time that can be scheduled in a calendar. Computerized entities representing material flows can be switched back and forth from a calendar list and waiting line lists in a logic that captures the essential behavior of such systems. This paradigm was crystallized thirty or so years ago, and since that time the primary activity of the field of discrete event simulation has fallen into the normal scientific pattern of working itself out for all kinds of systems. The problem of validation with respect to such models is the asking of the question as to whether a situation has been well captured by this particular paradigm. Kuhn would say that the paradigm is not up for questioning; rather, as the paradigm is still in a progressive phase (at least most of us think so) the questions we ask have to do with the adequacy of the model in this application. If there is a problem with the validity of the model, that problem is not with the paradigm but rather with our operationalization of it. What we have just said about discrete event simulation could be equally said about continuous simulation or any other kind of simulation.

In a Kuhnian sense, validation consists of persuading someone that one's model falls into a well-accepted way of seeing a problem. Again as in Friedman's case, it is an appeal to past success. In the field of simulation, Kuhn's approach might be the kind of defense that we are invoking when we call in an expert in a given area of practice to check over our simulation model. For Kuhn an expert is someone who is a recognized adept in a reigning paradigm. We are probably doing the same thing when we invoke "face validity" as a criterion. Turing's test also does something like this (Schuben, 1980). In a Kuhnian sense it seems to us that we could validate a model by exhibiting examples of like models used successfully elsewhere on similar systems. For Kuhn, the criteria that are involved in the decisions as to the adequacy of the model involve values like simplicity, consistency, fruitfulness of the theory or model, and the like (Kuhn, 1970, p. 199).

Kuhn's views on the nature of scientific development have come under heavy criticism for reasons
that have to do with the problem of validation that we are discussing in this essay. Some have seen his account of scientific development as including an arbitrary and relativistic element that leaves out the question of scientific validity altogether. Kuhn refuses to accept the relativistic label: instead by being true to past historical practice, he wants us to avoid justificationist prescriptions that he believes will stultify science. Perhaps the same admonition might apply to analogous requirements and prescriptions when it comes to validating simulation models.

4.4 The Methodology of Scientific Research Programs (MSRP)

Imre Lakatos has sought to put together a position that lies somewhere between Kuhn's alleged relativism and Popper's empirically oriented falsificationism (Lakatos, 1970). Like Kuhn, Lakatos thinks of a scientific movement as possessing a central core: he calls it "the hard core." Surrounding this core is a belt of problems that are attacked from the perspective of the core and developed in a series of theoretical steps over a period of scientific growth. He calls this series a "research program." Lakatos tries to lay out a prescriptive model for this development: this is his MSRP. He argues that progressive development avoids ad hoc 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. He analyzes historically the various forms of ad hoc dodges that have taken in bolstering theories that are failing empirically. This appeal to an empirically oriented adequacy is one of the elements that allies him with Popper's falsificationism.

Lakatos's MSRP has evoked much discussion and support among economists (See for e.g., Blaug, 1976, Caldwell, 1991, Howson, 1976, and Latsis, 1976). In simulation, MSRP may be germane to the validation problem as it has been in economics. We may argue that the research program of discrete event simulation, to take up that example again, has been progressive. We have been able to adapt the hard core of the methodology to attack an ever broadening set of problems. In the Lakatosian sense this area of simulation constitutes a research program whose credibility is increasing. The appeal to other successful applications as part of this widening program would be one side of a Lakatosian validation of a particular model. The other appeal would be to openness to criticism especially in regard to the empirical side of the model.

We might ask what an ad hoc dodge would look like in a field like discrete event simulation. We do not think that such dodges are hard to find. One way to save the validity of a model is to force the system that it describes to act like the model. This happens everyday. Instead of making the system the independent variable that determines the model, we make it work the other way around and then claim that the model is valid as a result. This approach works until the system breaks down. Then we claim that we didn't adequately validate the model. Again what is underscored here is an antijustificationist emphasis on human judgment that cannot be separated from the modeling activity itself.

4.5 The Bayesian Approach

The Bayesian approach is concerned principally with dealing with the problem of induction in a consistent way. Some of the opponents of Bayesianism see it as the last holdover from justificationism (Giere, 1984, p. 336). But probably most Bayesians are not justificationists. Howson and Urbach, for example, plainly accept the thesis that empirical observations are not indubitable givens but rather fallible theoretically oriented posits (Howson and Urbach, 1989, p. 94). Subjective probabilities are measures of credibility of a proposition or a theory. 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. We do not know of any work that has been specifically done in applying the Bayesian approach to the validation problem in simulation.

5 CUTTING THROUGH THE MIDDLE: BERSTEIN AND GADAMER

In his book, Beyond Objectivism and Relativism, Richard Bernstein has attempted to get beyond the polarity of the justificationist-antijustificationist debate. He argues that in a certain sense both sides are right and he develops his argument by using an interesting set of metaphors that he borrows from the German philosopher, H.G Gadamer. One of these metaphors is the activity of play. Bernstein says that playing a game involves both a subjective and relativistic element that each player brings to the game in the form of his or her own abilities, past experiences, and personal idiosyncrasies. It is only through this peculiar contribution by each player to the game 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 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 coopted into this more general perspective in one way or another. Play is not the only metaphor that Bernstein or Gadamer use in their attempts to understand how human knowledge comes about. In Gadamer's book, *Truth and Method*, there are a series of metaphors of which play is only the first one. He 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 the nature of human knowledge and how it is acquired.

In simulation there has been a recognition that playing with the simulation model is a way of bringing about a validation of it. Animation, especially interactive versions of animation, would resonate well with this play-like form of validation that Bernstein describes.

6 CONCLUSION

In summary our view of the validation question might be characterized as conventionalistic. We agree with Naylor and Finger that an eclectic approach to validation that blends together the various positions may be the most fruitful. We admit that this involves a factor of human judgment in the validation problem: in this sense we are unremitting antijustificationists. But we have tried to show that even though they may not explicitly know it, most simulationists have the common sense to be antijustificationists in practice. The philosophy of science now shows that the arguments that may be brought to bear in the validation of scientific theories are much wider than previously thought. We hope that by reviewing the various positions at the present time that simulationists will learn to live more comfortably with the validation problem by viewing it from a more general antijustificationist perspective.

REFERENCES


AUTHOR BIOGRAPHIES

GEORGE B. KLEINDORFER teaches computer simulation techniques in the Smeal College of Business Administration at the Pennsylvania State University. He holds advanced degrees in Mathematics, Law, and Industrial Administration. In Operations Research, he has worked on logistics problems, network theory, and combined discrete-continuous simulation. His current research interests focus on understanding costing in the simulation of production systems.

RAM GANESHAN is a doctoral candidate in Operations Management and Logistics at the Pennsylvania State University. He has a Master’s degree in Operations Research from the University of North Carolina at Chapel Hill and a Bachelor’s degree in Liberal Arts and Business from the Birla Institute of Technology and Science, Pilani, India. His research interests include global manufacturing systems, inventory control, philosophy of science, and The Beatles.