ERIC WINSBERG

MODELS OF SUCCESS VERSUS THE SUCCESS OF MODELS: RELIABILITY WITHOUT TRUTH[1]

ABSTRACT. In computer simulations of physical systems, the construction of models is guided, but not determined, by theory. At the same time simulations models are often constructed precisely because data are sparse. They are meant to replace experiments and observations as sources of data about the world; hence they cannot be evaluated simply by being compared to the world. So what can be the source of credibility for simulation models? I argue that the credibility of a simulation model comes not only from the credentials supplied to it by the governing theory, but also from the antecedently established credentials of the model building techniques employed by the simulationists. In other words, there are certain sorts of model building techniques which are taken, in and of themselves, to be reliable. Some of these model building techniques, moreover, incorporate what are sometimes called "falsifications." These are contrary-to-fact principles that are included in a simulation model and whose inclusion is taken to increase the reliability of the results. The example of a falsification that I consider, called artificial viscosity, is in widespread use in computational fluid dynamics. Artificial viscosity, I argue, is a principle that is successfully and reliably used across a wide domain of fluid dynamical applications, but it does not offer even an approximately "realistic" or true account of fluids. Artificial viscosity, therefore, is a counter-example to the principle that success implies truth – a principle at the foundation of scientific realism. It is an example of reliability without truth.

1. INTRODUCTION

The mathematical models that drive many computer simulations of complex physical systems have mixed ancestries. On the one hand, the basic features of these models are motivated by fundamental theory. On the other hand, in order to produce a model that is computationally tractable, simulationists also craft their models using a motley assortment of other components, incorporating many assumptions that are not sanctioned by high theory. Despite their mixed ancestries, many of these very simulations are trusted in

Synthese (2006) 152: 1–19
DOI 10.1007/s11229-004-5404-6

making predictions and building representations of phenomena, and they are often successfully used in engineering applications. Indeed, researchers run simulations of systems about which data from real experiments is difficult or impossible to get – the simulations take the place of experiments and observations – and so they are trusted even in circumstances where they cannot be evaluated by comparing their results with the world.

Given that the construction of these models is guided, but not determined, by theory, what is the source of the credibility of these models? In what follows, I argue that the credibility of a simulation model must come not only from the credentials supplied to it by its theoretical ancestors, but also from the antecedently established credentials of the model building techniques employed in its construction. There are, in other words, model building techniques that are taken, in and of themselves, to be suitable for building credible models. Some of these techniques, moreover, go beyond idealization or approximation, and involve incorporating what can be best described as "falsifications."[2] These are contrary-to-fact (or at least contrary to our best knowledge) modeling principles that are included in a simulation model and whose inclusion is taken to increase the trustworthiness of the simulation's output.

The practice of using falsifications in building credible simulations is worthy of closer scrutiny by philosophers of science. Here, I examine two examples of falsifications from the field of computational fluid dynamics: so called "artificial viscosity" and "vorticity confinement." Both of these techniques are successfully and reliably used across a wide domain of fluid dynamical applications, but both make use of 'physical principles' that do not purport to offer even approximately realistic or true accounts of the nature of fluids. I argue that these kinds of model building techniques, therefore, are counter examples to the doctrine that success implies truth – a doctrine at the foundation of scientific realism. I suggest, furthermore, that the modeling principles I call "falsifications" can provide a useful backdrop for thinking more carefully about what characteristics we take successful model-building principles to have.

2. BACKGROUND

Computer Simulations are techniques for studying mathematically complex systems; they have applications in almost every field of scientific study. “Discretization” techniques transform continuous differential equations into step-by-step algebraic expressions. “Monte Carlo” methods use random sampling algorithms even when there is no underlying indeterminism in the system. “Cellular automata” assign a discrete state to each node of a network of elements, and assign rules of evolution for each node based on its local environment in the network.

In this paper, I will be focusing on simulations techniques from Computational Fluid Dynamics (CFD) that employ methods of discretization. Discretization begins with a set of mathematical equations that depict the time-evolution of the system being studied in terms of rules of evolution for the variables of the model. In the case of computational fluid dynamics, the equations in question are some version of either the Euler equations or the Navier–Stokes equations, depending on the factors to be included, and the coordinate system to be employed.

Simulations in CFD, in other words, begin with mathematical models built out of well-established theoretical principles. But to study a particular flow problem, it is necessary to find solutions to those models’ equations, and for many of the systems that are of interest in the computationally intensive sciences, the models that are suggested directly by theory consist of second-order, non-linear differential equations. It is usually impossible, even in principle, to find closed-form solutions to these equations.

What a simulationists must do therefore, is to find a replacement for the model hatched out of the theory – one that that consists of a set of equations that can be iterated numerically, using step by step methods. When a theoretically motivated model is thus “discretized”, and turned into a simulation model, the original differential equations are transformed into difference equations, and crafted into a computable algorithm that can be run on a digital computer. The resulting evolution of the dependent variables on the computer (“data”) is said to “simulate” the evolution of the system in question.

3. AUTONOMY AND SANCTIONING

The above makes the whole process sound very simple, and entirely theory-driven. It is not. One important thing to emphasize about simulation models is what I call, following Mary Morgan and

Margaret Morrison, their semi-autonomy from theory. The claim frequently made by Morrison and Morgan that models are autonomous or independent of theory is meant to emphasize the fact that there is no algorithm for reading models off from theory (Morrison and Morgan 1999). While models generally incorporate a great deal of the theory or theories with which they are connected, they are usually fashioned by appeal to, by inspiration from, and with the use of material from, an astonishingly large range of sources: empirical data, mechanical models, calculational techniques (from the exact to the outrageously inexact), metaphor, and intuition.

Successful simulation models employ all of the above; and even what can be best described as deliberate falsifications[3] are often introduced into the model. In the end, the model that is used to run the simulation is an offspring of the theory, but it is a mongrel offspring. It is for these reasons that I insist that the construction of simulation models is guided, but not determined by theory. But since I don't want to deny the obvious and strong connections these models have to theory, I use the term "semi-autonomous."

A second important feature of simulations is that they are often constructed precisely because data about the systems they are designed to study are sparse. In these circumstances, simulations are meant to replace experiments and observations as sources of data about the world. Simulation methods, for example, are used to study the inner convective structure of stars, or to determine the distribution of pressure and wind speed inside a super-cell storm. Not all of the results of such simulations can be evaluated simply by being compared to the world. If a simulation reveals a particular pattern of convective flow inside a star, we have to be able to assess the trustworthiness of that information without being able to physically probe the inside of the star to check and see whether that result is confirmed by observation. In this sense, we can speak of simulations as being independently sanctioned. By this, I simply mean that if a simulation is to be useful, it must carry with itself some grounds for believing in the results it produces. The process of transformations itself, from theoretically given model to computationally tractable model, must be sanctioned. This is a feature of simulation that I have previously highlighted by talking of its "downward epistemology" (Winsberg 1999).

When these two features of simulations, their semi-autonomy from theory and the fact that they are independently sanctioned, are held up side by side, it raises an interesting question. If the process of constructing simulation models is at best only guided by theory, then how can the simulation be trusted to produce results in situations where data are sparse? What, other than the governing theory, could provide the necessary credentials?

At least part of the answer is that the techniques that simulationists use to construct their models are "self-vindicating" in much the same way that Ian Hacking says of instruments and experiments that they are self-vindicating (Winsberg 2003). That is, when simulationists build a model, the credibility of that model comes not only from the credentials supplied to it by the governing theory, but also from the antecedently established credentials of the model building techniques used to make it.

The principle purpose of simulations is to produce "results". These results come in the form of simulated "data" that are expected to, in specifiable respects, accurately predict or represent the phenomena to be simulated. Whenever these techniques and assumptions are employed successfully, that is, whenever they produce results that fit well into the web of our previously accepted data, our observations, the results of our paper and pencil analyses, and our physical intuitions, whenever they make successful predictions or produce engineering accomplishments, their credibility as reliable techniques or reasonable assumptions grows. That is what I meant when I said that these techniques have their own life; they carry with them their own history of prior successes and accomplishments, and, when properly used, they can bring to the table independent warrant for belief in the models they are used to build. In this respect, simulation techniques are much like experiments and instruments and as Hacking and Peter Galison describe them (Galison 1997; Hacking 1988).

That was the answer I gave to the question "What makes it possible for semi-autonomous simulation models to be credible sources of knowledge about systems for which data are sparse?" That is an epistemological question. Here I want to explore some of the implications of the answer to that question for the metaphysics of science. Roughly, I want to ask if there are any lessons we can learn from the observation that model builders develop trust in the success of some of their techniques, even when their techniques employ contrary-to-fact physical principles.

4. ASYMMETRIES OF SUCCESS

Let me begin to motivate the idea that this observation might have some interesting implications. Recall the claim: We sometimes trust the results of simulation models because

- we place trust in the theories that stand behind these models. We do this, of course, because they have been successful in lots of other applications,
- and equally, because we place trust in the model building techniques that we use to transform theoretical models into simulation models. Here too, we do this because they too have been successful in lots of other applications.

This way of putting things makes it seem as though there is a perfectly simple symmetry at work here. The apparent symmetry obscures a significant difference. We all know that theories and laws are the sorts of the things that are supposed to gain credibility – to be corroborated – every time they are applied successfully. But I take it that it might come as more of a surprise to some that a model building technique would be the right sort of thing in which to develop confidence. Just because a model building technique, a particular way of altering a theoretical model so as to make it more computationally tractable, produces good results in one particular application, one might be tempted to ask why we should expect it to work in another application. Why, in other words, should the success of a model building technique be the sort of thing that is projectable from one application to another?

The reason I think this is an interesting question is that it is often thought that there is only one possible explanation for our confidence in the projectability of scientific success. If we ask why we expect scientific theories that have been successful in past applications to be successful in the future, then the answer that people often give goes something like this: If a proposed theory or law is used successfully in making a variety of predictions and interventions, then it likely that that theory or law is in some way latching on to the real structure of the world – that it is true – and hence it is only natural that we should expect it to be successful in the future. If we did not believe it was latching on to the real structure of the world, so the reasoning goes, then we would have no reason at all to expect it to be successful in the future. Put short, the success of scientific theories is projectable because successful scientific theories are true.

In fact, many think that the truth of scientific theories is the *only possible explanation* of their success in making predictions and interventions. Otherwise, it is claimed, that success would be a miracle. Many also think that our *belief* in the truth of scientific theories is the only possible *explanation* for our common practice of trusting theories in making predictions about novel situations. These, of course, are standard arguments for scientific realism. But what, then, is the explanation of our practice of trusting *model-building techniques* in doing the very same thing?

Arguments for scientific realism, in other words, rest in at least in part on the *conviction* that the projectability of the success of scientific theories calls for an explanation, and that the only possible or viable explanation available is truth. In what follows, I ask whether, in light of the fact that some aspects of scientific practice *also* seem to rely on the projectability of the success of model building techniques, this conviction is warranted.

5. NO MIRACLES

The idea that the success of scientific theories requires an explanation, and that the best explanation for that success is truth, forms the basis for what Phillip Kitcher has called the "Success-to-Truth" rule (Kitcher 2002), which in turn is the engine of the "no miracles" argument for scientific realism. I state the rule crudely below. In the sequel, following Kitcher, I review some considerations that force upon us a more nuance formulation.

If X plays a role in making successful predictions and interventions then X is true.

Of course, the no miracles argument, and the rule of inference on which it depends, has many critics. One kind of criticism questions whether correspondence-truth really has any explanatory value to begin with (e.g., Horwich 1999). That avenue of criticism won't concern us here. Other critics have looked for counter-examples to the principle (e.g., Laudan 1981), usually in the form of historical examples of scientific theories that were successful, but which we no longer hold to be true – such as the humoral theory of disease, or the wave-in-aether theory of light.

Defenders of the "no-miracles" rule have responses to these sorts of counter-examples that invoke qualifications to the rule. Take, for example, the case of the humoral theory of disease. Examples of this type are often described by defenders of the rule as belonging to "immature science." The response, in other words, is that in order for X to be a genuine example of something to which the rule ought to apply, it needs to play a role in making *sufficiently specific and fine grained* predictions and interventions. Since, the argument goes, such theories as the humoral theory of disease and the phlogiston theory of combustion never allowed for sufficiently specific and fine grained predictions and interventions, they are not good counter-examples to a properly formulated rule. The first modification we have to make to the rule, therefore, is to accommodate this concern.

Another canonical counter-example to the no-miracles rule is the wave-in-aether theory of light. So called "structural realists" have taken this example to be paradigmatic of a certain kind of counter-example, and have crafted their version of realism in response. What they have urged is that these historical examples of successful but untrue theories can be divided into two parts: (a) a part that is no longer taken to be true, but that did not play a genuinely central role in the relevantly successful predictions and interventions, and (b) a part that did play a genuinely central role, but that is still taken to be true. In the case of the aether theory (a) is the mathematical form of the theory, which they argue can still be regarded as true, and (b) is the ontology of the theory (namely the existence of the aether) which it is argued did not play a genuinely central role in making predictions or interventions. So while structural realists admit that many scientific theories have been successful without being true in their entirety, they argue that the successes of these theories have come entirely in virtue of the fact that one component of the theory, its structure, has accurately reflected reality. Structural realists, in other words, argue that the rule of inference should be modified so that it only applies if X, *in its entirety*, plays *a genuinely central role* in making the successful predictions and interventions.

It has also been widely recognized that many successful theories and models, both from the past and present, cannot be held to be literally true or to represent exactly. The model of a simple harmonic oscillator can quite successfully predict the behavior of many real physical systems, but it provides at best only an approximately accurate representation of those systems. Newtonian mechanics is a very successful theory, but it is at best a limiting case of a true theory. Even maps can be used to successfully navigate their intended territories, but they omit many details, and distort certain features. Because of these considerations, many defenders of the no-miracles rule accept that it needs to be modified so that it guarantees, not the truth of X tout-court, but truth in some qualified sense – for example the approximate truth of X, or the fact that X is a limiting case of something true. For now, we can rewrite the consequent of the rule to read: "X is *(in some qualified sense)* true."

Ad hoc hypotheses are another class of obvious exceptions to the no-miracles rule. In order to qualify for application in the rule, it is widely recognized that X must achieve success across a wide range of applications. Better still is if the success of X is projectible – that is, that we not only thing that X can be used for some domain for which it was designed, but that we have the expectation that it will be useful in future domains. In short, the success of X has to be *systematic* (as opposed to *ad hoc*.)

One final modification to the rule as stated above is required. It is fairly obvious that the scope of the variable X in the rule cannot range over all entities. No one would deny that a calculator, a triple-beam balance, and even a high-energy particle accelerator can all play genuinely central roles in making specific and fine grained predictions and interventions. But no one would want to have to defend the view that any of these entities is "true", even in any qualified sense. So, to be perfectly pedantic about it, we need to be perfectly clear that the no-miracles rule applies only if the X in question is the right sort of entity to be a candidate for truth and falsity, or at the very least (in order to include such things as maps and models), similarity of structure with the world.

A properly formulated "no miracles" rule of inference thus reads as follows:

If . . .

> (the right sort of) X (in its entirety) plays a (genuinely central) role in making (systematic) successful (specific and fine grained) predictions and interventions.

Then . . .

> X is (with some qualification) true.

So, if some X is going to be offered as a counter-example to the no-miracles rule in a way that advances the debate about that rule, then it had better be the case that:

- X plays a genuinely central role in making predictions and interventions.
- Those predictions and interventions are specific and fine-grained.
- X cannot be separated into a part that is false and a part that does the relevant work.
- The predictions and interventions we use X to make occur across a wide range of domains, and we sometimes confidently apply X in new domains.
- X is the relevant sort of entity for consideration as true or false.
- X cannot plausibly be described as true, even in some suitably qualified sense.

I want to argue here that the field of Computational Fluid Dynamics offers some plausible candidates for counter-examples to the no-miracles rule that meet all of the above criteria. In what follows, I discuss two of them: "artificial viscosity" and "vorticity confinement."

6. THE SUCCESS OF FALSIFICATIONS

One of the earliest uses of finite difference simulations arose in connection with the Manhattan Project during World War II. John von Neumann and his group of researchers used finite difference computations to study the propagation and interaction of shock waves in a fluid, a subject crucial to the success of the Atomic Bomb.

We generally think of shock waves as abrupt discontinuities in one of the variables describing the fluid, but it was quickly recognized that treating them in this way would cause problems for any numerical solution. The reason is that a shock wave is not a true physical discontinuity, but a very narrow transition zone whose thickness is on the order of a few molecular mean-free paths. Even with today's high speed and high memory computers, calculating fluid flow with a differencing scheme that is fine enough to resolve this narrow transition zone is wildly impractical. On the other hand, it is well known that a simulation of supersonic fluid flow that does not deal with this problem will develop unphysical and unstable oscillations in the flow around the shocks. These oscillations occur because of the inability of the basic computational method to deal with the discontinuities associated with a shock wave – the higher the shock speed, the greater the amplitude of these oscillations becomes. At very high speeds, such a simulation quickly becomes useless. To make it more useful and accurate, what simulationists need to do is to somehow dampen out these oscillations.

The generally accepted way to do this, which was originally devised by von Neumann and Richtmyer while working at Los Alamos (von Neumann and Richtmyer 1950), is to introduce a new term, an "unphysically large value of viscosity" into the simulation which is called "artificial viscosity."[4] The inclusion of this term in the simulation is designed to widen the shock front, and blur the discontinuity over a thickness of two or three grid zones. The trick is to apply this viscosity only to those portions of the fluid that are close to the shock front. This is achieved by assigning a magnitude to the fake viscosity that is a function of the square of the divergence of the velocity field – which happens to be a vanishing quantity everywhere but close to the shocks. The end result is that the method enables the computational model to calculate certain crucial effects that would otherwise be lost inside one grid cell; in particular, the dissipation of kinetic energy into heat.

Artificial viscosity is not the only non-physical "effect" used in simulations of physical systems. Another example from CFD is what is known as "Vorticity Confinement" (Steinhoff and Underhill 1994). The problem to be overcome in this case arises because fluid flows often contain a significant amount of rotational and turbulent structure at a variety of scales. When that structure manifests itself at scales too small to be resolved on a grid of the size used in the simulation, significant flow features can become "damped out." This undesirable effect of the differencing scheme is called "numerical dissipation." The solution is to use a technique called vorticity confinement.[5] The method consists in finding the locations where significant vorticity has been numerically damped out, and to add it back in using an artificial "paddle wheel" force. Much as in the case of artificial viscosity, this is all done by a function that maps values from the flow field onto values for the artificial force.

For the rest of the discussion, I will confine my remarks to the example of artificial viscosity. Most of what I say, however, can be repeated mutatis mutandis, about the paddle wheel force, as well, presumably, as a variety of other falsifications used in simulation –

artificial viscosity is simply the oldest and most established of these techniques. Some remarks about artificial viscosity are now in order:

- Artificial viscosity is clearly a successful tool of scientific investigation, prediction and intervention.
- The success if artificial viscosity, furthermore, is systematic and projectable. The success is "projectable" in the sense that physicist and engineers use this technique of CFD to make *novel predictions* about flows containing shocks – its inclusion in a simulation of flows with strong shocks adds to the confidence that researchers have in their results. The success is "systematic" in the sense that its use has been studied extensively, and there is a wide body of knowledge about how to use this modeling principle effectively in an off-the-shelf manner. It is used, furthermore, in an enourmous variety of applications ranging from engineering applications to astrophysics.
- Artificial viscosity is not like the humoral theory of disease. It is used to make very fine-grained and detailed predictions, descriptions and interventions. Without it, the success of the Manhattan Project might not have come to pass.
- Artificial viscosity is not like the claims about the aether. It can perhaps be said of the 19th century wave theory of light that the claims it contained about the aether were superfluous – that these claims "could be dropped from the theory without affecting the success of the practice" This is clearly not the case of artificial viscosity. Artificial viscosity plays a crucial role in damping oscillation instabilities which would otherwise render simulation results useless.
- I said earlier of simulation *techniques* that they can bring to the table independent warrant for belief in the models they are used to build. There are, in short, such things as widely successful model building techniques. This fact alone, however, might not worry the proponent of the no-miracles rule, because *techniques* are not the sort of things that are candidates for being true or false, or for providing accurate representations. But artificial viscosity does take the form of a claim. We can easily think of artificial viscosity as the claim, made of fluids under its domain, that they display a viscosity that is proportional to the square of the divergence of their velocity field. In this sense, artificial viscosity is indeed a candidate for truth or falsity. The same is true of the technique of vorticity confinement. The technique calls for

> the application of a physical principle: the principle that certain kinds of fluid flows give rise to a paddle wheel force that arises in proportion to certain characteristics of the flow. This principle is, as well, a candidate for truth or falsity.

Finally, we need to address the question of whether we can understand "X" to be (perhaps in a qualified sense) true. To begin to do this, we need first to distinguish the model that drives the simulation, the principles that go into it, and the results that the simulation produces. Recall that we consider a simulation model to be successful if we have reason to think that predictions and representations of the phenomena that it produces – its results – are accurate in the respects that we expect them to be. Nothing whatsoever about my argument should prevent us from understanding these simulation results in realistic terms if we so choose.

Things are slightly more complicated when we look at the models that drive the simulations. Since these models incorporate false assumptions, we cannot view in them in literally realistic terms; they surely do not offer literally true accounts of the actual functional relationships that exist between the various properties of the fluid. But we are concerned here not only with literal truth, but with the possibility of approximate truth.

I take it, after all, that this is for many philosophers the standard way to think about successful contemporary theoretical structures that, for whatever reasons, don't seem like plausible candidates for being true descriptions of the world: the fact that they are successful is taken to be evidence that they must approximate some (perhaps as-yet undiscovered) ideal theory.

Approximate truth is a slippery subject. There is nothing like widespread agreement among philosophers about what a theory of approximate truth should look like, or even, for that matter, if one is even desirable or possible. Luckily, we can set these worries aside for our purposes, and ask only whether it is even plausible to think that the principle of artificial viscosity would come out as approximately true on any account of approximate truth. A separate difficulty, on the other hand, arises from the fact that there are two different ways to think about the question in this case: one that focuses on the models that drive the simulations, and the other on the principles that inform the model construction.

One way to ask the question is this: Should we be willing to say that there might be *some qualified sense* in which students of fluid dynamics accept as true (or should accept as true) the claim that certain fluids display a viscosity that is proportional to the square of the divergence of their velocity field? Here, I think the answer is clearly no.

Viscosity, in fluid dynamics, is the measure of how resistive the fluid is to flow. Most fluids can be well modeled by assuming that the sheer stress between parcels of fluid is proportional to their relative velocities and some constant, called the viscosity, which is taken to be a physical property of the fluid itself. These are called "Newtonian" fluids. In very high Reynolds number flows, like the kinds of flows we have been talking about, viscosity often becomes an insignificant parameter – it is often left out of the model. Of course, no one really thinks that any fluid is truly Newtonian, and certainly no fluid is inviscid. But treating a fluid as Newtonian, or inviscid, might arguably be thought of as providing a good approximation to the actual forces that small parcels of fluid experience as they slide against each other. On the other hand, when simulationists refer to "artificial viscosity" and label it an "unphysically large" "viscosity-like term", they are signaling that the term in the equation is not meant to capture, not even approximately, relationships between properties of the fluid and forces that occur within them.

Defenders of the idea that success implies truth, however, might want to argue that this is wrong question to ask. The relevant sort of question, one might argue, is not whether the claim that certain fluids display a viscosity that is proportional to the square of the divergence of their velocity field is approximately true. The relevant sort of question is whether or not the models that drive the simulations, which we build *using* artificial viscosity, accurately represent the real functional relationships that exist between the various properties of the fluid, at least approximately so. After all, it is not as if we use "the theory of artificial viscosity" by itself to make successful predictions and interventions. Its only when we couple this little bit of mathematical structure together with some form of the Navier–Stokes equations or the Euler equations – equations describing relationships between other variables of the fluids state – that we hope to make any useful predictions.

If we put the question this way, asking not whether the claim about artificial viscosity is approximately true, but asking rather whether the models we build – using artificial viscosity as one piece of the puzzle – are approximately realistic, the answer becomes less clear. Arguably, when each such model is considered individually, and in its entirety, there is no compelling reason to deny that, while all of these models contain false features, some of the models might very well count as reasonably accurately representing the relevant features of the fluid. There is arguably no compelling reason to resist viewing these models in quasi-realistic terms.

What I want to argue by way of reply to this objection is that defenders of the idea that success implies truth cannot simply avoid the first version of the question and hide behind the second. It is correct to say that it is the local models – models put together using falsifications, but also using lots of other bits of theory – that are the engines of local successes. And it may indeed be arguable that these local models accurately represent, at least reasonably so. But it is also the case that the little bit of mathematical structure known as "artificial viscosity" is entitled to its own, in this case much less local, record of success. Indeed, it is only the artificial viscosity itself, and not the local models contructed with it as an ingredient, that enjoys genuinely systematic and projectable success.

Recall a claim that I made above: Simulations are often used to learn about systems for which data are sparse. If such simulations are going to be at all useful, the models they use have to be trusted to produce good results despite all of the deviations from pure theory that go into making them. The credibility of those models comes not only from the credentials supplied to it by the governing theory, but also from the antecedently established credentials of the model building techniques employed by the simulationist. Now take, for example, simulations of fluid flows with strong shocks. There are dozens of different modeling schemes used to model these sorts of systems. They use varying difference schemes, they exploit different symmetries, they deal with truncation error in different ways, etc. One thing that the members of a large cross-section of these local models have in common is that they employ artificial viscosity to prevent unstable oscillations around the shocks. On a global basis, part of the reason we have for thinking that these local models are sanctionable is our conviction that artificial viscosity itself is a useful off-the-shelf piece of mathematical structure that can be successfully used to build such models. In other words, it is not just the case that each of the models that employs artificial viscosity is itself locally capable of being used to successfully make predictions and be used for interventions. It is also the case that artificial viscosity itself, when artfully and skillfully applied, is capable of being successfully used for building lots of different sorts of local models in lots of different contexts. The success of artificial viscosity is far broader than any one of the local models that includes it. It is a piece of mathematical structure that has its own degree of success and trustworthiness in its own domain – a domain that is much larger than that of any of the local models that are build with its help.

7. RELIABILITY WITHOUT TRUTH

If we cannot say of artificial viscosity that it is true, or even approximately true, what property does this model building principle have that allows us to expect it to yield sanctionable models – models that we trust in part *because* they include artificial viscosity in their set of equations? For those that believe that scientific success requires an explanation, what property of these model-building principles can we say accounts for our expectation of their future success? The right answer to this question, I think, is to say that model building principles like artificial viscosity are *reliable*. Borrowing from an idea introduced by Arthur Fine, I mean to employ this term to designate a semantic concept – reliability.

In his attempt to deconstruct the realism/anti-realism debate, Fine has argued that all arguments for realism based on the success of science fail because wherever the realist argues for truth, the instrumentalist can always settle for reliability, and *vice-versa*. (Fine 1991)

> instrumentalism takes reliability as its fundamental concept and differs from realism only in this: Where the realist goes for truth in the sense of a correspondence with reality, the instrumentalist goes for general reliability...Where the realist says that science does (or should) aim at the truth, the instrumentalist says that science does (or should) aim at reliability...The realist cannot win this game since whatever points to the truth, realist style, will also point to reliability. (Fine 1996, p. 183)

I would note, of course, that the pragmatic notion of reliability that Fine is suggesting is quite distinct from the view often discussed in the epistemological literature know as "reliabilism". Reliabilism is a view about what further characteristics *true* beliefs need to have in order for us to count them as knowledge. On this view, the beliefs need to have been generated by a reliable process or method. In contrast, the notion of reliability being treated here is one that is meant to take the place of truth as a basic semantic notion. This contrast between the notion of reliabilism in epistemology and the notion of reliability being used here is clearly related to the discussion of technique versus claim discussed above. To take a *claim* to be reliable is to trust that it can be relied on in much the way that reliabilists demand that our justification-producing methods can.

One other difference: reliabilists often define a reliable process or method in terms of something like the relative frequency with which it produces true results. Hence their notion of reliability is actually parasitic on truth. Clearly, our notion of reliability needs to avoid that.[6] Hence, I characterize reliability (for modeling principles) in terms of being about to produce results that fit well into the web of our previously accepted data, our observations, the results of our paper and pencil analyses, and our physical intuitions, and to make successful predictions or produce engineering accomplishments.

Of course, the semantic concept that Fine is talking about is *general reliability*. To take a claim as *generally* reliable "amount(s) to trusting it in all our practical and intellectual endeavors ... to be(ing) committed to understanding and dealing with the world from the perspective of that theory."(Fine 2001, 112)

Clearly, none of the contrary-to-fact model building principles used in computer simulation can be taken to be *generally* reliable in the robust sense that Fine intends. Artificial viscosity has restricted scope – it takes art and skill to know how and when to apply it. It contradicts the content of our theory of fluids. For these and other reasons, we simply cannot be committed to understanding and dealing with the world from its perspective. But reliability, unlike truth, comes in degrees. Instrumentalists like Dewey, according to Fine, believed that "in inquiry we strive for concepts and theories that are generally reliable – although we often make do with less." If we take "*generally reliable*" to be a regulative ideal, then "*broadly reliable*" is a real-world instantiation.

This weaker notion of reliability dovetails well with much of the recent work in philosophy of science inspired by Nancy Cartwright's anti-fundamentalism; work that rejects the notion of universally true (or even generally reliable) theories and laws. In this tradition, theories and laws are seen as providing a *framework* for building models, schematizing experiments, and representing phenomena. They have very broad, but not universal, domains of application. Rather than taking theories and laws to be *universally true* and delimiting the character of all possible worlds, the anti-fundamentalist sympathizer takes them to be *broadly reliable* for a wide array of practical and epistemic tasks.

Fine remarked that many philosophers might find a substitution of reliability for truth to be nothing but a "semantic slight of hand." Indeed, many fundamentalists feel the pull of the metaphysical intuition that behind any broadly reliable model building principle must lie a universally true law. Successful model building principles like artificial viscosity, however, provide a nice example to show that, at least in the case of *broad* reliability, no slight of hand is involved. No other semantic concept, not truth and not even approximate truth, adequately describes the proper attitude to have toward artificial viscosity and other model building principles like it. The success of these models can thus provide a model of success in general: reliability without truth.

NOTES

[1] Thanks to Robert Batterman, Ludwig Fahrbach, Arthur Fine, Stephan Hartmann, David Hyder, Johannes Leonard, Margaret Morrison, and Daniel Weiskopf, as well as other attendees of the ZIF conference on models and simulations in Bielefeld, 4S in Atlanta, and my talk at the University of Konstanz, for helpful comments and criticisms.

[2] I note that, of course, the word "falsification" has a long history in the philosophy of science associated with the ideas of Sir Karl Popper. For Popper, falsification involves rejecting a hypothesis because its predictions fail to match the evidence. Here, I intend the term to refer to something quite different: the practice of incorporating an assumption that you know to be false into a model precisely because you expect that the addition of this assumption will improve the predictive or representational accuracy of the model's output – i.e., the "data" produced by the simulation will better represent the phenomenon being simulated. The contrast between these two senses of the term borders on the ironic, but I can think of no better term to replace it with.

[3] In the sense defined above.

[4] For a modern discussion of artificial viscosity and its applications, (see Caramana et al. 1998; Campbell 2000). The original presentation can be found in (von Neumann and Richtmyer 1950).

[5] See (Steinhoff and Underhill 1994).

[6] I thank Daniel Weiskopf for pointing this out.

REFERENCES

Campbell, J.: 2000 'Artificial Viscosity for Multi-dimensional Hydrodynamics Codes', http://cnls.lanl.gov/Highlights/2000-09/article.htm

Caramana, E. J., M. J. Shashkov and P.P. Whalen : 1988, 'Formulations of Artificial Viscosity for Multi-Dimensional Shock Wave Computations', *Journal of Computational Physics* **144**, 70–97.

Fine, A.: 1991, 'Piecemeal Realism', *Philosophical Studies* **61**, 79–96.

Fine, A.: 1996, *The Shaky Game*, University of Chicago Press, Chicago.

Fine, A.: 2001, 'The Scientific Image' Twenty Years Later', *Philosophical Studies* **106**, 107–122.

Galison, P.: 1997, *Image and Logic: A Material Culture of Microphysics*, University of Chicago Press, Chicago.

Hacking, I.: 1988, 'On the Stability of the Laboratory Sciences', *The Journal of Philosophy*, **85**(10), 507–515.

Horwich, P.: 1999, *Truth,* Oxford University Press, Oxford.

Kitcher, P.: 2002, 'On the Explanatory Role of Correspondence Truth', *Philosophy and Phenomenological Research* **64**, 346–364.

Laudan, L.: 1981, 'A Confutation of Convergent Realism' *Philosophy of Science* **48**, 218–249.

Morgan, M. and Morrison, M.: (eds.): 1999, *Models as Mediators*, Cambridge University Press, Cambridge.

Steinhoff, J. and Underhill, D.: 1994, 'Modification of the Euler Equations for 'Vorticity Confinement': Application to the Computation of Interacting Vortex Rings', *Physics of Fluids* **6**, 2738–2744.

von Neumann J, and R. D. Richtmyer: 1950, 'A Method for the Numerical Calculation of Hydrodynamical Shocks', *Journal of Applied Physics* **21**, 232–247.

Winsberg, E.: 1999, 'Sanctioning Models: The Epistemology of Simulation', *Science in Context* **12**(2), 275–292.

Winsberg, E.: 2003, 'Simulated Experiments: Methodology for a Virtual World' *Philosophy of Science* **70**, 105–125.

Department of Philosophy
University of South Florida
FAO 226, 4202 Fowler Ave.
Tampa, FL 33620, U.S.A.
E-mail: winsberg@cas.usf.edu