Synthese (2009) 169:483–496
DOI 10.1007/s11229-008-9434-3

Does matter really matter? Computer simulations, experiments, and materiality

Wendy S. Parker

Received: 11 December 2006 / Accepted: 17 September 2007 / Published online: 7 November 2008

Abstract A number of recent discussions comparing computer simulation and traditional experimentation have focused on the significance of "materiality." I challenge several claims emerging from this work and suggest that computer simulation studies are material experiments in a straightforward sense. After discussing some of the implications of this material status for the epistemology of computer simulation, I consider the extent to which materiality (in a particular sense) is important when it comes to making justified inferences about target systems on the basis of experimental results.

Keywords Computer simulation · Experiment · Models · Materiality

1 Introduction

Philosophers have had surprisingly little to say about computer simulation modeling, considering its growing importance in science. Much of the work that has been done has looked for insight into computer simulation by comparing it to more established scientific practices, such as theorizing and experimenting (e.g. Dowling 1999; Galison 1997; Humphreys 1994; Rohrlich 1991; Sismondo 1999). Recently, it is the relationship between computer simulation and traditional experimentation that has attracted particular attention (e.g. Guala 2002, 2005; Keller 2003; Morgan 2002, 2003, 2005; Norton and Suppe 2001; Winsberg 2003). This recent work has yielded new insights about both practices and, collectively, constitutes an important thread in the burgeoning philosophical literature on computer simulation.

W. S. Parker (✉)
Department of Philosophy, Ohio University, Ellis Hall 202, Athens, OH 45701, USA
e-mail: parkerw@ohio.edu

Nevertheless, in this paper I argue that several claims emerging from this work—claims that all flag "materiality" as particularly important, in one way or another, for differentiating computer simulation and traditional experimentation—are mistaken. In the course of challenging these claims, I grapple with three important questions about computer simulation and experimentation: Do computer simulations, or perhaps computer simulation studies, qualify as experiments? If so, should such "computer experiments" be regarded as "nonmaterial" experiments, in contrast to the material experiments that traditionally take place in the laboratory? When inferences are made from the results of experiments to conclusions about target systems, are those inferences more justified when experimental and target systems are made of the "same stuff" than when they are made of different materials (as is the case with computer experiments)?

Section 2 considers whether computer simulations, or perhaps computer simulation studies, qualify as experiments. Guala (2002, 2005) suggests a negative answer when he claims that experiments can be distinguished from simulations, including computer simulations, as follows: in genuine experiments, the same "material" causes are at work in the experimental and target systems, while in simulations there is merely formal correspondence between the simulating and target systems. I identify several problems with this proposed distinction and ultimately reject the idea that no scientific study can be simultaneously, and as a whole, both a simulation and an experiment. I argue that, while computer simulations per se do not qualify as experiments, typical scientific studies involving computer simulation models do.

Morgan (2003) seems to agree that computer simulation studies often qualify as experiments. However, she characterizes them as "nonmaterial" experiments, in contrast to the "material" ones that traditionally take place in the laboratory. In Sect. 3, I argue against such a characterization, defending instead the view that any computer simulation study classified as an experiment is first and foremost a material experiment, in a straightforward sense that I explain. I then discuss why recognizing this material status is important when it comes to the epistemology of computer simulation.

"Materiality" has also been a focal point in comparisons of the epistemic power of computer simulations and traditional experiments. According to Morgan (2005), inferences about target systems are more justified when experimental and target systems are made of the "same stuff" than when they are made of different materials (as is the case in computer experiments). In Sect. 4, I argue that the focus on materiality is somewhat misplaced here, because it is relevant similarity, not materiality, that ultimately matters when it comes to justifying particular inferences about target systems. Nevertheless, I suggest an alternative sense in which Morgan may well be right that "ontological equivalence provides epistemological power" (ibid, p. 326).

2 Distinguishing simulations and experiments

Scientists often refer to computer simulations as "numerical experiments" or "computer experiments." But should computer simulations really be classified as experiments? Guala (2002, 2005) seems to suggest a negative answer when, taking inspiration from Simon (1969), he claims that there is a fundamental ontological difference between simulations and experiments:

> The difference lies in the kind of relationship existing between, on the one hand, an experiment and its target system, and on the other, a simulation and its target. In the former case, the correspondence holds at a "deep," "material" level, whereas in the latter, the similarity is admittedly only abstract and formal. …In a genuine experiment, the same material causes as those in the target system are at work; in a simulation, they are not, and the correspondence relation (of similarity or analogy) is purely formal in character. (Guala 2005, pp. 214–215).[1,2]

This way of attempting to distinguish simulations and experiments has some initial appeal, but upon examination turns out to be problematic in multiple respects.

2.1 Objections to Guala's proposed distinction

First, Guala's proposed distinction implies a view of experiments that is too restrictive, even assuming that he means not that *all* material causes at work in the two systems must be the same, but only that the (relatively proximate) ones that give rise to the behaviors of interest in the study must be the same. For surely when scientists give a new drug to rats in the hope of learning about the drug's potential for treating a disease in humans, they are *experimenting* on the rats, even if it turns out that the rats' response to the drug depends upon a physiological pathway that is absent in humans; the scientists are not mistaken in their belief that they are conducting an experiment, even if they are mistaken in their belief about what the typical response of rats who take the drug implies about the typical response of humans who will take the drug.

Guala's proposed distinction also implies a view of simulations that is too restrictive. Even if much of the time the material causes that produce behaviors of interest in simulating systems are not the same as those at work in their associated target systems, the suggestion that they are *never* the same is too strong. Consider a study of a full-scale prototype of a new car in a wind tunnel; it seems correct to say that the movement of air around the prototype in the wind tunnel is a *simulation* of the movement of air around the production version of the car when it is driven on the highway, despite the fact that many of the same material causes (e.g. the shape of the prototype/car, friction between air molecules and the surfaces of the prototype/car, etc.) are at work in producing the flow patterns in the two cases.[3]

[1] A target system is a system about which one ultimately wants to draw conclusions. For instance, when scientists use a computer to simulate the flow of some particular real-world river, the real river is a target system.

[2] Guala speaks of simulations in general, but it seems clear that computer simulations are to be included among them.

[3] It is not clear what Guala requires for there to be the "same material causes" in two systems, but this counterexample is intended to work under several reasonable interpretations.

Furthermore, the proposed distinction implies that no study as a whole can be simultaneously both a simulation of some target system T and an experiment undertaken to learn about that same target system T, since the required relationships with T are mutually exclusive.[4] Yet consider the following possible study using the Bay Model, a scale model of San Francisco Bay developed by the Army Corps of Engineers: a scientist places a carefully-determined type and volume of solid matter at a particular location in the Bay Model, and then observes the subsequent changes in the rate of fluid flow at various points in the Model, with the aim of learning how the placing of fill in a particular portion of the real Bay would impact the rate at which water moves past particular locations in the real Bay. Intuitively, it seems correct to describe this study as both an experiment on the Bay Model (with target system $T =$ the real Bay undergoing and responding to filling) and a simulation of the filling of the real Bay and the subsequent changing of its flow (again, $T =$ the real Bay undergoing and responding to filling).

2.2 Simulations, experiments, and computer simulation studies

As it stands, then, Guala's attempt to differentiate simulations and experiments does not quite succeed. Before investigating further how simulations and experiments might be distinguished, I want to emphasize that, like Guala (see 2002, p. 63), I am not particularly interested in conceptual distinctions per se. Still, I think it worth trying to develop plausible working accounts of simulation and experiment that can be used consistently in the remainder of this discussion and that might spark further analysis by others.

I characterize a *simulation* as a time-ordered sequence of states that serves as a representation of some other time-ordered sequence of states;[5] at each point in the former sequence, the simulating system's having certain properties represents the target system's having certain properties.[6] In virtue of what does one entity represent another? Not in virtue of there being some particular mind-independent relationship between the two entities; rather, it is just that some agent intends that the one entity stand for the other. Thus, in principle, any entity can serve as a representation (or model, or simulation) of any other. This does not imply, of course, that for a given purpose any entity is as *good* a representation (or model, or simulation) as any other.[7] In practice, scientists typically select a simulating system on the basis of its being hoped or believed to be similar to the target system in ways deemed relevant, given

[4] Guala does identify hybrid studies that he considers *part* experiment, *part* simulation, e.g. a study meant to investigate human consumer behavior that involves real human subjects but a simulated market (2002, pp. 71–72).

[5] This characterization of simulation has much in common with that offered by Hartmann (1996). Guala seems to adopt a similar view (see 2002, pp. 61–62).

[6] Note that the sequence of states that constitutes a simulation is something defined and chosen by the scientist—as a simulating system evolves, neither all of the states that it occupies nor all of its properties at any given time must be assumed to represent states or properties of the target system.

[7] I agree with Callender and Cohen (2006) that the question of what makes X a representation of Y should be separated from the question of what makes X a *good* representation of Y for some purpose.

the goals of the simulation study. For the kinds of goals typical in scientific studies (e.g. prediction or explanation), the aim is usually for there to be particular formal similarities between the simulating and target systems.[8]

An *experiment* can be characterized as an investigative activity that involves intervening on a system in order to see how properties of interest of the system change, if at all, in light of that intervention.[9] An *intervention* is, roughly, an action intended to put a system into a particular state, and that does put the system into a particular state, though perhaps not the one intended. Paradigmatic examples of interventions in traditional experimental contexts include giving a medical treatment to a patient, heating a volume of liquid over a burner, or removing a component of a system. Typically, the investigative activity that constitutes an experiment occurs under at least partially controlled conditions. Absent from this characterization of experiments, in contrast to that given by Guala, is any mention of a target system. It is true that when scientists conduct traditional experiments in the laboratory or in the field, they typically hope that what they learn about the particular system on which they intervene (i.e. the experimental system) can be generalized to provide information about other systems. It is also true that, in order to justify these generalizations, scientists typically aim for various material similarities to obtain between their experimental systems and those target systems.[10] Nevertheless, on the view advocated here, scientists' so aiming (much less whether any hoped-for similarities actually obtain) is irrelevant to whether their activities count as experiments.

These characterizations imply at least the following fundamental difference between simulations and experiments: while a simulation is a type of *representation*—one consisting of *a time-ordered sequence of states*—an experiment is an *investigative activity* involving *intervention*. This difference, however, does not entail that "simulation" and "experiment" must have mutually exclusive extensions, contrary to what follows from Guala's proposed distinction. Most straightforwardly, on the characterizations given here, one experiment might serve as a simulation of some other experiment, perhaps one that involves intervening on the target system directly, as with the Bay Model study described above.

Up to now the discussion has focused on simulations in general, not specifically on computer simulations. How are computer simulations classified according

---

[8] A *formal similarity* consists in the sharing of some formal property, where a formal property is a property that can only be instantiated by a *relation* between/among other properties. For example, if the evolution of selected properties of two physical systems can be described accurately using a single mathematical equation, then the systems share the formal property of having properties that are related as described by that equation, and the systems are formally similar in that respect. (Both here and in the discussion of material similarity below, I assume that properties can be specified as broadly as one likes, e.g. they can be specified as disjunctions of other properties.)

[9] Many other authors emphasize the close connection between intervention and experiment, including Hacking (1983), Tiles (1993), Radder (1996), Morgan (2003) and Woodward (2003).

[10] A *material similarity* consists in the sharing of some material property, where a material property is a property that can only be instantiated by a material/physical entity. Thus, having in common the property of being brittle would constitute a material similarity, as would having in common the property of having a kidney, while having in common the property of being a monotonic function or of generating trajectories within some region of a state space would not. Note that material similarity is neither the negation nor the complement of formal similarity.

to the characterizations of simulation and experiment presented here? A computer simulation is a sequence of states undergone by a digital computer, with that sequence representing the sequence of states that some real or imagined system did, will or might undergo.[11] So computer simulations do qualify as simulations, as expected. But precisely because computer simulations consist of sequences of computer states—not activities undertaken by inquiring agents—they clearly do not qualify as experiments.

However, we might draw a further distinction between a computer simulation and a *computer simulation study,* which consists of the broader activity that includes setting the state of the digital computer from which a simulation will evolve, triggering that evolution by starting the computer program that generates the simulation, and then collecting information regarding how various properties of the computing system, such as the values stored in various locations in its memory or the colors displayed on its monitor, evolve in light of the earlier intervention (i.e. the intervention that involves setting the initial state of the computing system and triggering its subsequent evolution). So defined, a computer simulation study *does* qualify as an experiment—an experiment in which the system intervened on is a programmed digital computer.[12] The importance of looking at computer simulation studies in this way—as experiments on real, material/physical systems—will be the focus of the next section.[13]

3 The material status of computer simulations

Mary Morgan seems to hold a view of experiments very similar to the one just offered and to agree that many computer simulation studies in science qualify as experiments (see Morgan 2002, 2003). Nevertheless, she aims to call attention to important differences among studies that qualify as experiments, and she is particularly concerned with differences in their "degree of materiality" (2003, p. 231). Her analysis of the degree of materiality of different types of experiment is detailed and complex. To some extent, however, her characterization of computer simulation studies as *nonmaterial* experiments (e.g. ibid, p. 217) seems to depend on their being understood as experiments on mathematical models or mathematical objects (ibid, p. 225 and 232) rather than on real, physical stuff. It is this rather common way of thinking of computer simulation studies that I want to take issue with in this section.

In fact, I want to stress the importance of instead understanding computer experiments as, first and foremost, experiments on real material systems. The experimental system in a computer experiment is the programmed digital computer—a phys-

[11] See also footnote 6.

[12] Since a simulation is one element in a computer simulation study, such studies have a representational element in that respect at least; characterizing computer simulation studies as experiments on material/physical systems in no way denies that representational element. The example of the Bay Model experiment above shows that, in addition, a computer simulation study as a whole might in some cases serve as a representation (e.g. of some other experiment).

[13] Winsberg (2003) suggests that other authors who characterize computer simulation studies as experiments in which the programmed computer is the experimental system beg the core question of the epistemology of simulation, because they *assume* that the programmed computer "reliably mimics" the natural target system (ibid, p. 115). The view advocated here makes no such assumption.

ical system made of wire, plastic, etc. As described in the last section, a computer simulation study involves putting a computing system into an initial state, triggering its subsequent evolution (the simulation), and collecting information regarding various features of that evolution, as indicated by print-outs, screen displays, etc. It is those data regarding the behavior of the computing system that constitute the immediate results of the study. In a computer simulation study, then, scientists learn first and foremost about the behavior of the programmed computer; from that behavior, taking various features of it to represent features of some target system, they hope to infer something of interest about the target system.

Suppose one does not deny that computer simulation studies are material experiments in this sense, but claims that such studies are also experiments on abstract mathematical systems. Can such a view be defended? I will not attempt to answer this question here, but I will make several observations. First, the advocate of such a view should explain how intervention on an abstract system takes place during a computer simulation study and whether (and, if so, the sense in which) the abstract system moves through a series of states during the computer simulation. Second, if one or more abstract mathematical systems can be intervened on during a computer simulation study, presumably they will be ones that are realized by the computing system;[14] but if it turns out that any abstract mathematical system realized by the computing system can be intervened on during a computer simulation study, then it seems that many traditional experiments can also be experiments on abstract mathematical systems, since their experimental systems also realize various mathematical systems. This would make it nothing special that computer simulation studies can be considered experiments on abstract mathematical systems. Third, and perhaps most importantly, if a computer simulation study is an experiment on one or more abstract mathematical systems, we need to take care when identifying those mathematical systems; typically a computer simulation indicates *at best* only approximate solutions to the equations of the mathematical model that scientists, on the basis of background theory and other available knowledge of the target system, would most like to use to represent that system. Why?

For one thing, these preferred mathematical model equations typically do not even appear in the program that generates the computer simulation. Scientists often undertake computer simulation studies when they cannot solve the preferred equations analytically. Typically, the preferred equations are discretized, simplified and otherwise manipulated until they are in a form such that solutions can be estimated using methods that require only brute-force calculation, which the computer can provide; the resulting equations are what appear in the program that generates the computer simulation (i.e. are the *programmed equations*). In addition, as I just hinted, the numerical methods used to solve the programmed equations often are *designed* to deliver only approximate solutions—solutions to within some epsilon—rather than exact solutions. Furthermore, even if the solution method were not so designed, and even if the

[14] Following (Norton and Suppe 2001, p. 105), system B is a *realization* of system A when there is a many-one behavior-preserving mapping from the states of B onto the states of A. Note, however, that by "states" here I mean states of the system as defined by the scientists conducting the study, which will pick out only some system variables and only at particular times or intervals of time (see footnote 6).

preferred model equations were identical to the programmed ones, a simulation still almost surely would not indicate exact solutions to the preferred equations. This is because a digital computer can encode or represent numbers only to finite precision. As a consequence, the results of calculations made during a simulation will be rounded off, and the solutions generated during the simulation will differ at least a little bit from the exact solutions to whatever equations appear in the computer model. Thus even if it makes sense to say that a computer simulation study is an experiment on *some* mathematical system, perhaps one realized by the physical computing system, the solutions indicated by the computer simulation typically are not identical to the corresponding solutions for either the preferred model equations or the programmed equations.

The upshot is that, like the material/physical systems that scientists set out to simulate in the first place, the mathematical systems specified by the preferred model equations and by the programmed equations must be thought of as *target systems*, and conclusions about them on the basis of computer simulation results cannot be automatic, but rather require justification.[15] Of course, scientists who actually perform simulation studies often recognize this. In many fields, there is even a special (and somewhat misleading) name—*verification*—given to the process of collecting evidence that a computer simulation can indicate solutions that are close enough to those of the programmed equations and/or the preferred model equations.[16] Verification activities are concerned not only with such things as the adequacy of the numerical methods that are employed to estimate solutions to the programmed equations, but also with the design and actual implementation of the computer program that is run to generate the simulation (see e.g. Oberkampf et al. 2004). Possible errors in the coding of the program are considered important and are investigated as part of the verification process. And if the computer hardware itself is not working properly or as expected, that too can sometimes be revealed through verification activities.

The material status of computer simulations is no less important when the aim is to learn about material/physical target systems, rather than mathematical ones. In fact, the verification activities described above are often undertaken as one step in the process of establishing that some computer simulation results can be trusted to provide information of interest about some real-world material target system. With that kind of approach, the aim is often to justify a chain of inferences, from simulation results to conclusions about solutions to the programmed equations, from conclusions about solutions to the programmed equations to conclusions about solutions to the preferred model equations, and from conclusions about solutions to the preferred model equations to conclusions about the real-world target system. But such an approach need not be taken; it is possible to collect evidence regarding the trustworthiness of a computer simulation model (for a given purpose) through direct comparison of its results with

[15] Another way to put this point is to say that, for all three systems—the real-world material/physical system, the mathematical system specified by the preferred model equations, and the mathematical system defined by the programmed counterpart equations—it is the "external validity" of the computer experiment that is at issue. Experimental results are externally valid with respect to some target system if the results apply to (or, more loosely, provide desired information about) the target system.

[16] What counts as close enough, and which solutions need to be that close, depends on the goals of the study.

observations of the real-world target system. Computer models of the atmosphere that are used to simulate and forecast the weather, for instance, are evaluated and improved in light of repeated testing against meteorological observations, with relatively little concern for what their results might indicate about the properties of various mathematical systems.[17] Such evaluations involve comparison of the observed behavior of one material/physical system (the programmed computer(s) that produces the forecasts) with the observed behavior of another material system (the atmosphere), and there is a keen awareness of this materiality among the scientists who manage and develop the forecasting system; they recognize the potential for such things as hardware changes and addition of new code to impact the accuracy of their forecasts, and they watch for and try to avoid problems that might arise from such changes.[18]

Nevertheless, discussions of computer simulation often give the impression that computer simulation studies are just mathematical modeling exercises, and that the epistemology of computer simulation is all about whether the preferred model equations adequately represent the real-world target system. But the foregoing shows that such a view is doubly mistaken. First, in a very straightforward sense, computer simulation studies are *material* modeling exercises and *material* experiments. Second, the epistemology of computer simulation must attend to this materiality, or it will be impoverished, if not fundamentally confused. Ordinarily, the mathematical system specified by the preferred model equations is no less an external target system than the real-world material system that motivated the computer simulation study in the first place. Inferences about these and any other target systems will take as their starting point the observed behavior of a material system—the programmed digital computer—and accounts of the epistemology of computer simulation should reflect this.

4 Materiality, relevant similarity and justified inference

Materiality has also been a focal point in comparisons of the epistemic power of computer simulations and traditional experiments. According to Morgan, traditional experiments "have greater potential to make strong inferences back to the world" (2005, p. 317), because "ontological equivalence provides epistemological power" (ibid, p. 326): in traditional laboratory experiments, the experimental and target systems can be made of the "same stuff," and when they are, scientists are more justified in crossing the inferential gap from experimental results to conclusions about the target system (ibid, p. 323; see also Morgan 2003, p. 230 and Morgan 2002, p. 54). But is it true that inferences about target systems are more justified when experimental

[17] Winsberg (1999) discusses the epistemology of computer simulation more generally and notes that it is ultimately aimed toward warranting conclusions about real-world target systems rather than abstract theoretical structures.

[18] There can be reason for concern even when no hardware/software changes have been recently made. For instance, in the early days of computerized weather forecasting, hardware was rather unreliable; breakdowns occurred regularly and were a major source of anxiety and frustration (see e.g. Cressman 1996 and Nebeker 1995).

and target systems are made of the “same stuff” than when they are made of different materials (as is the case with computer experiments)?

It seems clear that it is not always true. Consider the case of weather forecasting again. Scientists want to reliably predict, to within a few degrees, noontime temperatures in various cities. But it is extremely difficult (if not impossible) to construct a same-stuff laboratory model that can be used to make such predictions, in part because it is extremely difficult (if not impossible) to put any such model into an initial state that reflects enough of the real atmosphere’s complicated temperature structure. This, along with a host of other considerations, render scientists not at all justified in inferring from the behavior of such laboratory models the noontime temperatures that interest them, despite the fact that the laboratory models are made of the “same stuff” as the real atmosphere.[19] In practice, computer simulation models are used instead. And since scientists know from past usage that some of these models often do succeed in predicting, to within a few degrees, noontime temperatures for at least some cities of interest, they can be much more justified in drawing conclusions about those temperatures on the basis of the predictions of their computer simulation models (at least for those cities for which the model at hand has a track-record of many past successful predictions) than on the basis of experiments on same-stuff laboratory models of the atmosphere.[20] As this example illustrates, it is not always the case that inferences about target systems are more justified when experimental and target systems are made of the “same stuff” than when they are made of different materials.

It might be suggested, however, that Morgan’s claim should be read as one concerning what holds in ceteris paribus conditions: Ceteris paribus, inferences about target systems are more justified when experimental and target systems are made of the “same stuff” than when they are not.[21] But Morgan includes no such qualification herself, and it is far from obvious how such a ceteris paribus clause would even be interpreted; what exactly is the “all else” that must be “equal” in the different experiments? One interpretation is easy to dismiss: it cannot mean that everything else must be the same in the two experiments, since there is a straightforward sense in which one cannot do the “same” intervention or make the “same” observations in two experiments in which the systems being intervened on and observed are quite different in material and structure. It is unclear whether there exists some other reasonable interpretation of the ceteris paribus clause on which Morgan’s claim, so qualified, is both true and widely-enough applicable to be interesting from the point of view of actual scientific practice. But instead of trying to find such an interpretation, I would like to move in what seems to me a more fruitful direction. For regardless of whether such an interpretation can be found, I would argue that the focus on materiality (in the sense of having experimental and target systems made of the “same stuff”) is somewhat misplaced here. In the remainder of this section, I will explain why and then will draw

[19] Laboratory models of the atmosphere usually are constructed of liquids rather than air, but these models also fail to be useful for many of the predictive tasks that interest scientists, and for many of the same reasons.

[20] For a discussion of complications involved in arguing that computer simulation models of the atmosphere provide good evidence for conclusions about tomorrow’s weather, see Parker (2008).

[21] An anonymous referee suggested that this is in fact how Morgan’s claim should be understood.

on that analysis in suggesting that Morgan is nevertheless right that materiality has epistemic significance, even if not quite the significance that she seems to claim.

As a first step, let us reflect further on what sorts of things scientists need to be concerned with when attempting to justify inferences from experimental results to conclusions about target systems. I would argue that some of the most important things include: (a) whether the desired intervention was actually performed, (b) whether adequate conditions of control were actually achieved, (c) whether the desired observations of the experimental system (i.e. those that, according to the experimental design, will provide the basis for inference) were actually made, and (d) whether the experimental and target systems were actually similar in the ways that are relevant, given the particular question to be answered about the target system. Since a failure with respect to any of (a) through (d) can undermine an inference from experimental results to some conclusion about a target system, evidence regarding any of (a) through (d) is relevant as scientists consider the extent to which such an inference is justified. But it is (d), I claim, that is the key to understanding both why the focus on materiality is somewhat misplaced and why there is nevertheless something to be said for materiality when it comes to learning about target systems.

The focus on materiality is somewhat misplaced because what is ultimately of interest when it comes to justifying inferences about target systems is not materiality, but relevant similarity. To justify such an inference, scientists need evidence that the experimental system is similar to the target system in whatever respects are relevant, given the particular question they want to answer about the target system. The relevant similarities might be formal similarities, material similarities or some combination of the two, depending on the type of experiment and the target question at hand. But, crucially, having experimental and target systems made of the same materials does not guarantee that all of the relevant similarities obtain. The weather forecasting example above illustrates this, though it might not seem so at first, because in presenting the example I emphasized scientists' inability to carry out the desired intervention. But their inability to carry out that intervention amounts to an inability to ensure that the experimental and target systems are similar enough in one of the ways that is recognized as particularly important (relevant) given the goal of forecasting temperatures, namely, in the spatial distribution of temperature in the systems at a particular earlier time.

But even setting aside such problems with intervention—and potential problems with control and observation as well, for that matter (see (a–c) above)—traditional laboratory experiments undertaken to learn about fluid phenomena, like the weather, provide prime examples of how experiments on same-stuff models can fail to be relevantly similar to target systems. In such experiments, it is often critical to ensure that various dimensionless parameters of the fluid flow are quite similar in the experimental and target systems. For instance, it may be essential to ensure that the ratio of quantities that defines the Reynolds number of the flow is not too different in corresponding regions of the experimental and target systems. But because some of these parameter values depend on things other than just which fluid is being used—e.g. on such things as the depth of the fluid and the size, shape, roughness and movement of any container holding it—the relevant similarities can fail to obtain (and scientists can have good reason to believe that they in fact do fail to obtain) even when the experimental and target systems are made of the "same stuff." So, even when experimental and target systems are made of the same materials, there may be every reason to think that the systems are not similar in all of the relevant respects and thus that particular inferences about target system are not justified. This is why Morgan's emphasis on materiality is somewhat misplaced here.

Nevertheless, I do not mean to suggest that materiality is entirely irrelevant from the point of view of learning about target systems; there does seem to be a sense in which Morgan is right that "ontological equivalence provides epistemological power" (Morgan 2005, p. 326). For it does seem true that, oftentimes, experimental and target systems that are made of the "same stuff" will be similar in many scientifically interesting respects. If so, that in turn suggests that often there will be much that could be learned about a target system, at least in principle, through intervening on and observing an experimental system made of the "same stuff." That is, in many cases, there may well be many scientific questions of interest about a given target system for which an experimental system made of the same materials actually *will* be similar in the relevant respects. And for that reason, especially when scientists as yet know very little about a target system, their best strategy may well be to experiment on a system made of the "same stuff" (or perhaps even, I would argue, of the "same kind") as the target system, in the hopes that the various similarities between them will include the ones that are relevant, given the sorts of things they would like to find out about the target system.[22]

It would be a mistake, however, to infer from the reasonable assumption that there exists some nontrivial set of questions that in principle are answerable via experiments in which experimental and targets systems are made of the same materials, the conclusion that any particular question of interest is a member of that set. Likewise, even if it is true that experimenting on a system made of the "same stuff" as a target system is often among the most promising strategies for adding to one's store of knowledge about the target system, this does not imply that one will be more justified in making any particular inference about a target system when experimental and target systems are made of the "same stuff" than when they are made of different materials. In some cases, as the weather forecasting case illustrates, scientists can have good reason to think that the results of a particular computer experiment are more likely to provide desired information about a target system than the results from even the best-available traditional experiment in which experimental and target systems are made of the same materials.

5 Conclusions

Contrary to the suggestion of Guala (2002, 2005), simulations and experiments cannot be distinguished by considering whether the "same material causes" are at work in various systems. A simulation is a type of representation—one consisting of a

[22] The recognition that such experiments may well be the best strategy in such circumstances is perhaps what lies behind Morgan's support for traditional experimentation as "a preferable mode of enquiry," especially in economics (see 2005, p. 326).

time-ordered sequence of states—while an experiment is a type of investigative activity involving intervention. Although computer simulations per se do not qualify as experiments, computer simulation studies do, because they involve intervening on a system in order to study how properties of interest of that system change, if at all, in light of the intervention.

The system directly intervened on during a computer simulation study is a material/physical system, namely, a programmed digital computer. Computer simulation studies (i.e. computer experiments) are thus material experiments in a straightforward sense. Characterizing them instead as nonmaterial experiments on mathematical models or mathematical objects, as Morgan (2003) and others sometimes do, can be misleading and can encourage impoverished views of the epistemology of computer simulation modeling. It is the observed behavior of a material/physical system—the programmed digital computer—that constitutes the immediate results of a computer experiment, and accounts of the epistemology of computer simulation should reflect this.

It is not the case that traditional experiments have greater potential to make strong inferences back to the world than computer experiments do; for both types of experiment, the justification of inferences about target systems requires that scientists have good reason to think that the relevant similarities—whether material, formal or some combination of the two—obtain between the experimental and target systems, where relevance is a function of the particular questions being asked about the target system. Since relevant similarity is what ultimately matters when it comes to justifying particular inferences about target systems, the intense focus on materiality here is somewhat misplaced. Nevertheless, there does seem to be an important sense in which Morgan (2005) is right that ontological equivalence provides epistemological power, even if it is not quite the power that she seems to claim.

Acknowledgements Thanks to Dan Steel, Eric Winsberg, Francis Longworth, Phil Ehrlich, John Norton, Michael Hoffman and two anonymous referees for helpful comments on earlier drafts of this paper.

References

Callender, C., & Cohen, J. (2006). There is no special problem of scientific representation. *Theoria, 55*, 67–85.

Cressman, G. (1996). The origin and rise of numerical weather prediction. In J. R. Fleming (Ed.), *Historical essays on meteorology 1919–1995* (pp. 21–39). Boston: American Meteorological Society.

Dowling, D. (1999). Experimenting on theories. *Science in Context, 12*(2), 261–273.

Galison, P. (1997). *Image and logic*. Chicago: University of Chicago Press.

Guala, F. (2002). Models, simulations, and experiments. In L. Magnani & N. Nersessian. (Eds.), *Model-based reasoning: Science, technology, values* (pp. 59–74). New York: Kluwer.

Guala, F. (2005). *The methodology of experimental economics*. Cambridge: Cambridge University Press.

Hacking, I. (1983). *Representing and intervening.* Cambridge: Cambridge University Press.

Hartmann, S. (1996). The world as a process. In R. Hegselmann, K. Troitzsch, & U. Mueller (Eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view* (pp. 77–100). Dordrecht: Kluwer.

Humphreys, P. (1994). Numerical experimentation. In P. Humphreys (Ed.), *Patrick Suppes: Scientific philosopher* (Vol. 2, pp. 103–121). Boston: Kluwer.

Keller, E. F. (2003). Models, simulation, and 'computer experiments'. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 198–215). Pittsburgh: University of Pittsburgh Press.

Morgan, M. (2002). Model experiments and models in experiments. In L. Magnani & N. Nersessian (Eds.), *Model-based reasoning: Science, technology, values* (pp. 41–58). New York: Kluwer.

Morgan, M. (2003). Experiments without material intervention: Model experiments, virtual experiments and virtually experiments. In H. Radder (Ed.), *The philosophy of scientific experimentation* (pp. 216–235). Pittsburgh: University of Pittsburgh Press.

Morgan, M. (2005). Experiments versus models: New phenomena, inference, and surprise. *Journal of Economic Methodology, 12*(2), 317–329.

Nebeker, F. (1995). *Calculating the weather: Meteorology in the 20th century*. New York: Academic Press.

Norton, S., & Suppe, F. (2001). Why atmospheric modeling is good science. In C. Miller & P. N. Edwards (Eds.), *Changing the atmosphere: Expert knowledge and environmental governance* (pp. 67–105). Cambridge: MIT Press.

Oberkampf, W., Trucano, T., & Hirsch C. (2004). Verification, validation, and predictive capability in computational engineering and physics. *Applied Mechanics Reviews, 57*(5), 345–384.

Parker, W. S. (2008). Computer simulation through an error-statistical lens. *Synthese, 163*(3), 371–384.

Radder, H. (1996). *In and about the world: Philosophical studies of science and technology.* Albany: SUNY Press.

Rohrlich, F. (1991). Computer simulation in the physical sciences. In A. Fine, M. Forbes, & L. Wessels (Eds.), *PSA 1990* (Vol. 2, pp. 507–518). East Lansing: Philosophy of Science Association.

Simon, H. (1969). *The sciences of the artificial*. Boston: MIT Press.

Sismondo, S. (1999). Models, simulations, and their objects. *Science in Context, 12*(2), 247–260.

Tiles, J. E. (1993). Experiment as intervention. *British Journal for the Philosophy of Science, 44*(3), 463–475.

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.

Woodward, J. (2003). *Making things happen: A theory of causal explanation.* Oxford: Oxford University Press.