



Eckhart Arnold^b

Tools for Evaluating the Consequences of Prior Knowledge, but no Experiments. On the Role of Computer Simulations in Science

— — — Pre-print — — —

Stuttgart, July 2011

^b Institute for Philosophy, University of Stuttgart,
Seidenstraße 36,
70174 Stuttgart, Germany

eckhart.arnold@philo.uni-stuttgart.de

Abstract There is an ongoing debate on whether or to what degree computer simulations can be likened to experiments. Many philosophers are sceptical whether a strict separation between the two categories is possible and deny that the materiality of experiments makes a difference [Morrison, 2009, Parker, 2009, Winsberg, 2010]. Some also like to describe computer simulations as a “third way” between experimental and theoretical research [Axelrod, 2006, Küppers and Lenhard, 2005, Rohrlich, 1990].

In this article I defend the view that computer simulations are not experiments but that they are tools for evaluating the consequences of theories and theoretical assumptions. In order to do so the (alleged) similarities and differences between simulations and experiments are examined. It is found that three fundamental differences between simulations and experiments remain: 1) Only experiments can generate new empirical data. 2) Only Experiments can operate directly on the target system. 3) Experiments alone can be employed for testing fundamental hypotheses. As a consequence, experiments enjoy a distinct epistemic role in science that cannot completely be superseded by computer simulations.

This finding in connection with a discussion of border cases such as hybrid methods that combine measurement with simulation shows that computer simulations can clearly be distinguished from empirical methods. It is important to understand that computer simulations are not experiments, because otherwise there is a danger of systematically underestimating the need for empirical validation of simulations.

Keywords Computer Simulations, Virtual Experiments, Epistemology of Simulations

Preprint Series
Stuttgart Research Centre for Simulation Technology (SRC SimTech)

SimTech – Cluster of Excellence
Pfaffenwaldring 7a
70569 Stuttgart

publications@simtech.uni-stuttgart.de
www.simtech.uni-stuttgart.de

Issue No. 2011-X

Contents

1	Introduction	2
2	Common features of simulations and experiments	3
3	Distinguishing features of experiments and simulations	5
4	Borderline cases	9
5	Summary and conclusion: Computer simulations as a tools for drawing conclusions from prior knowledge	12

1 Introduction

In the literature on computer simulations there is an ongoing debate about whether simulations are experiments [Morrison, 2009, Winsberg, 2009]. In this article I defend the view that computer simulations and experiments are two separate and clearly distinguishable categories and that notwithstanding the fact that the usage of the words “simulation” and “experiment” may be somewhat vague, a sharp line can be drawn between simulations and experiments from an epistemological point of view. In order to do so, I discuss the various similarities and differences between simulations and experiments that have been pointed out in the literature. Only those differences are considered as crucial that are epistemically relevant in the sense that they make a difference regarding the justification of the simulation or experiment in question.¹ The discussion of similarities and differences between simulations and experiments shows that there are at least three fundamental and epistemically relevant differences between simulations and experiments which justify a sharp distinction between the category of simulations and the category of experiments: 1) New empirical data can only be gathered by doing experiments or real measurements, but not by doing simulations. 2) Only experiments can operate directly on the target system, simulations never do. 3) Experiments, empirical measurements and observations alone can be used for testing fundamental hypotheses. Simulations can only be used for testing the consistency of a hypothesis with a background theory, but not for testing the fundamental theories themselves.

As the formulation of these differences suggests, they do not become acute in every single instance of a simulation or experiment. Rather, the categories of simulations and experiments are both large and inhomogeneous and the concepts of simulations and experiments must therefore remain somewhat vague. Also, there is an overlap region between simulations and experiments, because some experiments essentially fulfil the function of *analog simulations*. In order to clarify the situation the border cases of the *experimentum crucis* that can for principle reasons never be substituted by a simulation as well as the cases of *simulation-like experiments* and *experiment-like simulations* will be discussed.

The most critical case, however, is that of *hybrid methods* which combine empirical measurements with sophisticated simulation-like computations. I argue that a hybrid method remains essentially an empirical method (i.e. experiment or measurement) – in contradistinction to a computer simulation based on empirical input data – as long as the result the hybrid method yields is a result about exactly the same system from which the input data was taken.

In the conclusion I argue that the differences between simulations and experiments follow from the fact that a computer simulation cannot yield any other results than those that are logically implied by its premises. This fact places computer simulations firmly on the theoretical side of science in contradistinction to its empirical side (experiments, observations, experiences), which contradicts the view that computer simulations are a “third way of doing science” [Axelrod, 2006, p. 90] between induction (empirical research) and deduction (theory) and relativises the view that “materiality” is not a proper distinguishing criteria [Parker, 2009].

¹ By justification I mean answers to the questions if and what they allow us to learn about nature and why we should believe it to be true.

2 Common features of simulations and experiments

Although the understanding of computer simulations as “computer experiments” [Gramelsberger, 2010] or as a “third way” of doing science will be criticised in this article, it must be admitted that these characterizations of computer simulations are well motivated by a large number of sometimes striking similarities between computer simulations and experiments. In the following I first examine the similarities between simulations and experiments and then their alleged differences. The differences will be discussed at greater length than the similarities, because it is the claim of this article (disputed by others) that there exist several insurmountable differences between simulations and experiments, while it is not denied that there are indeed many common features.

The common features of simulations and experiments are the following:

1. *Methodological structure*: Both simulations and experiments operate on an object to learn something about a target system. The object must in some way or other be representative of the target system. In the case of an experiment the object can also be identical with the target system or be an instance of the target system (see below, section 3, point 2). Here, with *object* the entity on which a computer simulation or an experiment operates is meant. With *target system* the entity in nature about which we want to learn something through a simulation or an experiment is meant.² An example for an experiment where the object is not identical with the target system would be a ripple tank that is used to study the nature of light waves. An example for an experiment where the target system is identical or at least an instance of the target system would be a pendulum that is used to study gravity. In the case of computer simulations the object is always a representation of the target system but never the target system itself. The only possible exception would be a simulation that is not conducted to learn anything about some target system in nature, but merely to study the properties of the model that is implemented in the simulations. In this case the simulation does not have a target system.
2. *Controlled Environment*: Both simulations and experiments run in a controlled environment. However, in a simulation all causally effective factors are controlled by the simulation setup. In an experiment all but one specific factor, the factor under experimentation, is controlled. The fact that in order to run a computer simulation all parameters must have determinate values has been described as their “semantic saturation” by Barberousse, Franceschelli and Imbert [2009, p. 572]. One could say that from the point of view of nature an experiment is semantically saturated, too, but not from the point of view of the human experimenter.
3. *Interventions*: Computer simulations just as experiments allow interventions on the object [Parker, 2009, p. 487] [Morgan, 2003, p. 223]. In fact, the easiness of intervening in the model and monitoring the effects is one of the advantages of computer simulations over material experiments.
4. *Evaluation tools*: Computer simulations apply tools that were formerly thought of as typical for experimental data analysis like visualisation, statistics or data mining [Winsberg, 2010, p. 33]. This, again, emphasizes the similarity between both types of scientific procedure.
5. *Error management*: Similar techniques of error management are used for both simulations and experiments. Among these are:
 - (a) Validation of the setup (or the apparatus) against cases with known results.
 - (b) Testing for the responsiveness on interventions.
 - (c) Replicating the results.

² In this article the term “target system” is strictly confined to empirical target systems.

- (d) Testing for the conformance of the results with undisputed theoretical and phenomenological background knowledge.

Regarding the last two points, one might wonder what sense these make in the case of simulations, since the exact replication of simulation results becomes trivial if both the program code and the system specification are given. And conformance of the results with background theory becomes trivial if the theory is built into the simulations [Winsberg, 2010, p. 44]. However, replicating a simulation in a different system environment, say, with a different simulation package or another set of math libraries may indeed be useful in order to ascertain that the simulation results do not depend on the idiosyncrasies of a particular computational environment or infrastructure. For experiments, too, replication under varied conditions is considered to confirm the experiment's results more than replication under exactly the same conditions [Franklin and Howson, 1984]. Only, in contrast to computer simulations replication under the same conditions still bestows some additional inductive confirmation to an experiment. Contradictions of the computer simulation's results with the background theory might reveal implementation errors or unanticipated consequences of approximations and simplifications. Just like with experiments there is "constant concern for uncertainty and error" [Winsberg, 2010, p. 34] in simulations.

6. *Unanticipated results*: Both simulations and experiments allow us to learn something new and potentially surprising about their object [Morgan, 2003, p. 224]³ and, if the object is truly representative, also about the target system.
7. *Partial autonomy from theory*: Just as it has been described for experiments by Hacking [Hacking, 1983], simulations "have a life of their own" and are in part "self-vindicating" [Winsberg, 2010, p. 45]. Winsberg [2001, p. 447] has described simulations as "downward", "autonomous" and "motley". At least autonomy and being motley can equally well be ascribed to experiments.
8. *Comparable epistemological challenges*: Simulations and experiments share the challenge of bridging the gap between their object and the target system, or to put it differently, between the laboratory setup on the one hand and the real world outside the laboratory on the other hand [Arnold, 2008, p. 174/175]. Again, there are a subtle differences in this respect between simulations and experiments: In the case of a simulation the gap to be bridged is that between the numerical representation and the represented target system. In the case of the experiment the question is rather if the behaviour that the target system exposes under laboratory conditions can be transferred to real world instances of the target system.

Thus, there is a considerable number of important features which computer simulations and experiments have in common or with regards to which they appear to be at least very similar. But the similarities would justify placing simulations into the same category as experiments only if there are not, at the same time, any fundamental differences between simulations and experiments. Therefore, in the following, the alleged differences between simulations and experiments will be examined.

³ Mary Morgan makes a subtle difference between simulations that "may surprise the experimenter" and relatively more material experiments that "may yet confound the experimenter" [Morgan, 2003]. In contrast to Morgan, I do not believe that there is a difference regarding the kind and strength of "surprise" that computer simulations on the one hand side and material experiments on the other hand side may elicit from the experimenter. Rather, the difference she hints to that the relatively more experimental procedure contains real, and therefore to her estimate potentially "confounding", empirical data is in this article taken care of by considering the gathering of new empirical data as a distinguishing feature of experiments in contrast to simulations (see point 5 in section 3 below).

3 Distinguishing features of experiments and simulations

Although there are many striking similarities, there are also quite a few differences between simulations and experiments. The question is if any of these differences is epistemically relevant and important enough to place simulations and experiments into different epistemological categories. A difference would be epistemically relevant if, due to this difference, experiments allow us to gain knowledge about the real world that cannot, even in principle, be gained by simulations. If such a difference exists then we can – and in fact have to – distinguish between simulations and experiments.

The following candidates for distinguishing features of simulations and experiments have been suggested:

1. *Materiality*: Quite a bit of confusion has arisen about the claim that it is the materiality of experiments (as opposed to the “virtuality” of computer simulations) that distinguishes experiments from simulations. Philosophers that hesitate to follow this distinction point out that computer simulations also operate on physical, and thus “material” systems, because a computer is a physical system [Hughes, 1999, p. 139] [Parker, 2009]. Sometimes even the view is discussed that running a program on a computer is just an experimental trial of the computer as a physical system [Winsberg, 2010, p. 36, 54]. Clearly such an extreme view cannot be upheld as is shown by Barberousse, Franceschelli and Imbert [2009]. At the same time, although it is discussed in the literature, it appears doubtful whether such an extreme position has seriously been maintained by any philosopher at all. Winsberg ascribes this view to Humphreys [1994] by maintaining that Humphreys argues for the view that simulations are experiments “on the grounds that when a program runs on a digital computer, the resulting apparatus is a physical system.” [Winsberg, 2010, p. 36]. But ascribing this position to Humphreys is most certainly a mistake,⁴ because in the article that Winsberg refers to, Humphreys explicitly says: “It is this simulation by numerical methods of physical process that has led to the term ‘numerical experimentation’... It thus occupies an intermediate position between physical experimentation and numerical mathematics” [Humphreys, 1994, p. 112], which does not sound like the position that Winsberg ascribes to him. The misinterpretation already appears in Winsberg [2003, p. 114]. Possibly following Winsberg [2003], though with some hesitation, the same misinterpretation of Humphreys [1994] is made by Barberousse, Franceschelli and Imbert [2009, p. 559/560]. Notwithstanding this misinterpretation their paper has the merit of making clear that what matters in a computer simulations is not the “material” process, but what happens on the semantic level, i.e. what the bits and bytes mean that are processed in the computer.

The author coming closest to the view falsely ascribed to Humphreys by Winsberg is Wendy Parker, who maintains that “Computer simulation studies ... are .. material experiments in a straightforward sense”, because the “system directly intervened on during a computer simulation study is a material/physical system” [Parker, 2009, p. 495]. But even this formulation does not imply that Parker considers a computer simulation as a trial *of* the computer as a physical system.

A somewhat less extreme view is taken by Morrison, who “locates the materiality not in the machine itself but in the simulation model” [Morrison, 2009, p. 45]. With this wording, however, her denial that materiality makes a difference rests on a rather liberal use of the word “material” which blurs the distinction between the representation of a material object and the material object itself. Her argument therefore leaves the fact undisputed that computer simulations work with a numerical and in this sense non-material representation of the target system while “material” experiments examine a physical object directly. Just as Parker [2009], Morrison is still right in so far as the

⁴ This has been confirmed to me by Paul Humphreys.

materiality of experiments does not generally render the results of experiments more credible than those of simulations [Morrison, 2009, p. 54f.].

Where does all this leave us? According to Winsberg [2010, p. 62] the property of “materiality” is too vague to mark the difference between simulations and experiments [Winsberg, 2010, p. 62]. However, this vagueness seems to be mostly due to the fact that different authors are using the term “materiality” in a different sense. If “materiality” is strictly confined to that what happens on the physical level then it can serve as a distinguishing feature because computer simulations represent their target system on a semantic level which experiments do not. But even then it is not the fact that computer simulations represent the target system on a semantic level as such that forbids them to preempt the epistemic function of experiments in all possible cases, but that as a consequence of this fact they cannot operate on the target system directly.

2. *Operating directly on the target system:*

Although, as has just been argued, materiality can be appealed to in order to draw a line between simulations and experiments, an epistemically more relevant difference consists in the fact that “in a simulation one is experimenting with a model rather than the phenomenon itself” [Gilbert and Troitzsch, 2005, p. 14]. Or, in other words: Simulations do not operate directly on the target system.

It is true that – as Margret Morrison asserts [2009, p. 43] – in many experiments the measuring procedures already assume a model of the target system. But while the outcome of a computer simulation of a target system is exclusively determined by the model of the target system embedded in the simulation, the outcome of an experimental measurement is not determined by the model of the target system alone. Rather, it is determined by the measurement device, which, as the case may be, presumes a model of the target system [Morrison, 2009, p. 43] and the target system itself. It is also true that not all experiments operate directly on the target system. If one experiments with an electrical harmonic oscillator in order to learn something about a mechanical oscillator [Hughes, 1999, p. 138], then this experiment does not operate on the “phenomenon itself”. However, in order to make a difference between the two categories of computer simulations and experiments it suffices that some experiments operate directly on the target system and no computer simulation operates directly on the target system. While the latter is obvious it has been called into question by Eric Winsberg, whether there really are any experiments of which it can be said – without further qualification – that they operate directly on the target system. As Winsberg explains: “It might be argued, of course, that Mendel’s peas and Galileo’s chandelier are instances of the systems of interest . . . few of Galileo’s contemporaries would have thought of his chandelier as a ‘freely falling object.’ Some, conceivably, might have doubted that cultivated plants are an instance of natural heredity” [Winsberg, 2010, p. 52/53]. However, Winsberg’s examples do not show that experiments do not operate directly on the target system, but they highlight a completely different problem, namely, whether the inductive inference from a particular instance to the general case is warranted. Even if in the two examples a hypothetical sceptic might deny the validity of inductive reasoning, both examples remain examples of experiments that operate directly on the target system. The most that can be concluded from them is that operating directly on the target system is not much of an epistemic advantage for experiments in cases where we are in doubt about what kind of inductive inferences from the experimental results are warranted.

Thus, operating directly on the target system is a feature that distinguishes at least some experiments from computer simulations. And, what is more, it is a feature that extends the epistemic

reach of experiments beyond that of simulations. Operating directly on the target system therefore marks an epistemically relevant difference between experiments and simulations.

3. *Epistemic primacy*: As another distinguishing feature of experiments and simulations the epistemic primacy of experiments has been suggested. Material experiments, it is assumed, “have a potentially greater epistemic power than nonmaterial ones” [Morgan, 2003, p. 221]. However, the claim of an epistemic primacy of experiments has strongly been disputed [Parker, 2009]. It is impossible to claim a general epistemic superiority for experiments because there are processes that we know enough about to calculate, but which – at the same time – are difficult to trace experimentally. Under this condition, experiments compare poorly to computer simulations of the same process. Good examples for this are computer simulations of chemical reactions [Senn and Thiel, 2009], which we can calculate because they follow the laws of quantum mechanics.

Yet, a certain kind of epistemic primacy can still be claimed for experiments on behalf of their more direct relation to the empirical world. Experiments are more directly related to the empirical world because the object of an experiment is part of the world, whereas the object of a computer simulation is always a numerical representation of a model of some part of the world. The difference is of minor importance if the object and the target system are not the same, because then the crucial question is whether the object adequately represents the target system, which in this case (and in this case only) is the same question for computer simulations and experiments. But the fact that experiments operate on real-world objects becomes very important when fundamental hypotheses need to be tested. A hypothesis is fundamental if neither the hypothesis nor its negation are implied by known facts and known natural laws. Because it is in principle impossible to test fundamental hypotheses with computer simulations, one can reasonably say that experiments are “epistemically prior” [Winsberg, 2010, p. 71] to computer simulations. It should be noted that the status of being a fundamental hypothesis may change over time. For example, Kepler’s laws were fundamental at the time of their discovery. But later they could be derived from Newton’s theory of gravity. Still, at any point in the history of science there exist some fundamental hypotheses that in virtue of their being fundamental cannot be tested by a computer simulation but need to be tested by material experiments.

Thus, the fact that experiments can be employed to test fundamental hypotheses while computer simulations cannot, is another distinguishing feature.

4. *Different justifications*: As a further criterion for distinguishing computer simulations and experiments Eric Winsberg has suggested that the difference lies in “how researchers *justify* their belief that the object can stand in for the target.” According to Winsberg “What distinguishes simulations from experiments is the *character of the argument given* for the legitimacy of the inference from object to target and the *character of the background knowledge* that grounds that argument.” [Winsberg, 2010, p. 63] However, if simulations need to be justified differently from experiments, then this implies that there is a more fundamental difference between simulations and experiments of which their different mode of justification is merely a symptom. If this was not the case, why would there be the need for a different justification in the first place? This more fundamental difference seems to be no other than the following:
5. *Generating new empirical data*: It lies in the very nature of computer simulations that they cannot generate new empirical data. Here, a piece of data is considered *new empirical data* if its existence is neither part of nor implied by our previous knowledge and if it contains information about the world. According to this definition, for a piece of data to count as empirically generated, it is not sufficient to contain information *about* the world alone. The data must also in some sense

(i.e. by measurement or observation) have been gathered *from* the world. It is presupposed as a fundamental fact here that data about the world which are not logically implied by our existing stock of knowledge can only be obtained by empirical observation.

In contrast, the expression “new knowledge” is used here already if no human agent has been aware of it so far (and in that sense hasn’t known it), notwithstanding whether it is implied in our previous knowledge or not. Thus, one can reasonably say that computer simulations provide us with new knowledge about the world, even though all of the knowledge that simulations can provide is already contained in the theories and other explicit or implicit assumptions that enter into the construction of the computer simulations, like modeling assumptions, local theories and (old) empirical data.

A difficulty when discussing this point is that the terminology is by no means fixed and can thus easily be misunderstood. It must therefore be emphasized that “generating new empirical data” is strictly understood in the sense of obtaining new data from the empirical system under study. The situation can be illustrated by a simulation of H-tunnelling on the basis of quantum mechanics which has been conducted by Goumans and Kästner [2010].⁵ In this example computer simulations lead to the conclusion that tunneling of H contributes to H₂-enrichment in outer space. This is new knowledge because without the simulation results there would be much less reason to consider this true, although it could of course be conjectured. Yet, this simulation did neither gather nor generate any new empirical data. The data it produced is derived from quantum mechanics plus further modeling assumptions. One can say that quantum mechanics is itself derived – by means of inductive and abductive reasoning – from the physical experiments that lead to the development of quantum mechanics. Thus, in a sense, they indirectly form part of the input of the simulation. But this does not make the simulation any more empirical. Only if the simulation results would be data about those empirical systems from which the input data was taken, the simulation could be understood as the evaluation component of a highly refined empirical measurement procedure (see section 4.4 below). Since this is not the case here, it really is a simulation, i.e. a non-empirical procedure of deriving knowledge about particular empirical systems.

The position maintained here that simulations merely explore the consequences of existing knowledge is denied by Winsberg who, referring to another example, says: “To think it is true is to assume that anything you learn from a computer simulation based on a theory of fluids is somehow already ‘contained’ in that theory. But to hold this is to exaggerate the representational power of unarticulated theory. It is a mistake to think of simulations as tools for unlocking hidden empirical content” [Winsberg, 2010, p. 54]. However, save for the qualification that what we can learn from a simulation is not only what is contained in the theory but what is contained in the theory plus further implicit or explicit modeling assumptions, Winsberg is simply wrong here. There just are no other sources from which simulations can draw in order to produce their results.

Thus, another clear and important distinguishing feature of experiments is that experiments generate empirical data and can therefore provide us with new empirical information while simulations cannot.

To sum up the discussion, we find three important distinguishing features of experiments: 1) Experiments can provide us with new empirical data, computer simulations cannot. 2) Some experiments operate directly on the target system, computer simulations never do. 3) Experiments can be used for the testing of fundamental hypotheses (*experimentum crucis*), computer simulations cannot.

⁵ I am indebted to Johannes Kästner from the Institute of Theoretical Chemistry at the University of Stuttgart for explaining this simulation to me.

4 Borderline cases

The above described distinguishing features between simulations and experiments justify making a difference between the categories of simulations and experiments. At the same time they represent optional features of experiments that do not become acute in every single instance. Therefore, there exist instances of experiments that do not differ in any epistemically relevant sense from simulations which means that the two categories overlap. In the following, it will be argued that this does not imply that there are single instances where the distinction is blurred up to a point at which it cannot be determined any more whether some scientific procedure is a simulation or an experiment. Rather, a distinction can always be made, only in some cases the distinction is not epistemically relevant. In these cases computer simulations can (at least in principle) act as replacement for experiments.

In order to clarify the situation, I will first give examples of extreme cases of experiments that do not fall into the overlap region of simulations and provide examples of experiments that fall squarely inside the overlap region. Then, the borderline cases of simulation-like experiments and experiment-like computer simulations will be described on a more abstract level. Finally, I discuss the question of hybrid simulation-experiments, i.e. scientific procedures like magnetic resonance imaging where empirical raw data is post-processed by simulation-like procedures so that on the first sight it appears hard to tell whether it is a simulation or an experiment. Nevertheless, it is claimed, a reasonable distinction remains possible.

4.1 Experimentum crucis and analog simulations

The fact that the categories of simulation and experiment overlap in cases where none of the three distinguishing features become relevant, may lead to the mistaken conclusion that simulations and experiments cannot be clearly distinguished or separated as research methods or that it is at least hard to tell the difference [Morrison, 2009, Parker, 2009]. The confusions arising from this mistaken conclusion can best be clarified by considering the border case of experiments that fall clearly outside the overlap region of the two categories and the other border case of experiments that obviously fall inside the overlap region.

The border cases that need to be considered in this respect are on the one hand the *experimentum crucis* that can never be replaced by a computer simulation, and the *analog simulation* that can in principle always be replaced by a computer simulation on the other hand.

An example for an *experimentum crucis* is Young's double-slit experiment to demonstrate the wave nature of light [wikipedia, 2010]. At the time when it was conducted it would for principle reasons have been impossible to replace this experiment by a computer simulation, since the outcome of a computer simulation would depend on which of the competing theories, wave theory or corpuscular theory, had been built into the simulation.

The opposite case of an analog simulation can best be highlighted by a quotation by John von Neumann, because it describes this case in ideal-typical form:

“The purpose of the experiment is not to verify a proposed theory but to replace a computation from an unquestioned theory by direct measurement. . . . Thus wind tunnels are used . . . as computing devices . . . to integrate the nonlinear partial differential equations of fluid dynamics.” (quoted by Winkler et al. [1987, p. 28] and Winsberg [2010, p. 35])

In the particular case that von Neumann speaks of, the experimental setup performs the function of an analog computer. It is trivial that in this case the experiment can – without any loss of epistemic

power – be replaced by a computer simulation if the computers are powerful enough. It is important to keep in mind that this example covers only one particular kind of experiment.

4.2 Simulation-like experiments

The last example gives rise to the notion of *simulation-like experiments*. Simulation-like experiments can be understood as experiments where none of the above mentioned distinguishing features becomes relevant. These are experiments that do not operate directly on the target system but on an object representing the target system, in which case also the empirical data they provide is not data about the target system but about the object that is used in the experiment to represent the target system. Or, if they do operate on the target system, it is merely for convenience and the data they provide is not expected to be any different from that which would have been predicted by models of the target system based on known theories. And finally, simulation-like experiments can only be experiments which are not employed to test fundamental hypotheses.

The distinction between simulation-like experiments and other types of experiments may not be very strict. This holds in particular if it is made a requirement that simulation-like experiments do not operate on the target system directly but, like computer simulations, substitute the target system with a representation of it. Does a scale-model, for example, operate on the target system or not? It could be argued yes, because the scale-model is of the “same stuff” and the same natural laws take effect on it. But it could also be argued no, because a scaled down model is not the real thing itself.

On the other side, the distinction between simulation-like experiments and computer simulations appears to be reasonably clear. Computer simulations operate on a symbolic representation of their target system, but the representation of the target system in simulation-like experiments is not symbolic. Thus, simulation-like experiments and computer simulations can be distinguished, although this difference may not be epistemically relevant in the above defined sense. For, if the symbolic representation in the computer is good enough, computer simulations can replace simulation-like experiments.

It is important to keep in mind that even though some experiments are simulation-like and therefore not easily distinguished from simulations, it would be wrong to take this as an indication that simulations in turn are generally like experiments.

4.3 Experiment-like computer simulations

The opposite case to simulation-like experiments is that of computer simulations that are *experiment-like*. The motivation for considering computer simulations as experiment-like is that some computer simulations resemble experiments in many different ways. As has been noted earlier (section 2) there are indeed many similarities between simulations and experiments and, naturally, these are strongly emphasized by that part of the philosophical literature on simulations that tends to liken simulations with experiments [Morrison, 2009, Parker, 2009, Winsberg, 2010]. Again, not all of these similarities may become relevant in every single case of a simulation. But as a matter of terminological convention we can call those simulation studies *experiment-like* which are conducted in a way that mimics experimental studies and which therefore show many of these similarities. This is often the case with simulations that are used as surrogates for experiments when it is either too costly or ethically unacceptable or for other reasons impossible to conduct a material experiment. The previously quoted example of H-tunnelling under conditions that hold in outer space is a good example for an experiment-like computer simulation, because it can be regarded as a substitute for an experiment to determine the

H-tunnelling rate which is practically not feasible, because the tunnelling rates are too slow to be determined experimentally at temperatures which are as low as in outer space [Goumans and Kästner, 2010, p. 7351].

Experiment-likeness if understood in this sense describes a similarity of simulations to experimental procedures on the phenomenological level. It does not make a simulation any more empirical if it is experiment-like. Therefore, no matter how experiment-like some simulations are, there will remain some experiments (namely those, where it matters that they have an empirical content) that will never be fully resembled nor replaced by any simulation.

4.4 Hybrid simulation-experiments

A more complicated case is that of hybrid methods which fuse simulations and empirical methods such as experiments or measurements. It is the existence of these methods that Morrison [2009] mobilizes in order to dispute a clear cut distinction between simulations and empirical measurements. An example would be magnetic resonance imaging [Lee and Carroll, 2010], because before the raw data obtained from the electromagnetic signals emitted by the previously stimulated protons of the body is turned into an image it runs through various highly sophisticated computations some of which are not quite unlike a computer simulation. This raises the question whether the procedure as a whole is more like a measurement and thus an empirical procedure or more like a computer simulation or, maybe, something different that resembles neither of these. The last answer represents the stance that Mary S. Morgan [2003] takes. If this answer is true then this would mean that the distinction between simulations and experiments (or measurements for that matter) is inevitably blurred.

In order to answer this question a few words need to be said about the nature of measurement. Any measurement procedure that makes use of a more or less sophisticated apparatus, and in particular those measurements that involve computations, produces data as its output that for the sake of the apparatus or the computations involved is best described as *refined data*. Now, it is still reasonable to classify *refined data* (in contrast to *raw data* which has not gone through any intermediate processing or computation steps) as empirical data or empirically measured data as long as the data describes the state of system from which the measurement was taken at the time of the measurement. Thus, if we measure the humidity and temperature today in order to compute the humidity and temperature of tomorrow this is not a measurement but a prognosis based on a measurement. But if we measure the volume and weight of a piece of metal and then compute its density, we can reasonably speak of an (indirect) measurement of the density, the density being the *refined data*, while the weight and volume is the *raw data*.⁶

But this means that even a highly sophisticated hybrid method such as magnetic resonance imaging can be classified as empirical measurement, because the refined data it produces is data about the same system from which the raw data was gathered and is at the same time dependent on the content of the raw data. In this particular case, the systematic link results from the fact that the refined data

⁶ In the literature on simulations also other definitions of hybrid methods can be found. Mary S. Morgan, for example, speaks of “the hybrid ‘simulation’ form, which mixes mathematical and experimental modes” [Morgan, 2003, p. 225]. For her a simulation is already a hybrid if it uses empirical input data, in which case she uses the term “virtually experiments” (see her table 11.3 on page 231 of her paper). In the beginning of her paper she gives the example of two simulations of a bone where the model of the bone is based on different kinds of empirical measurements. For her, both are hybrid methods which she locates between experiments and simulations. According to the stance adopted in this article both are simulations because the output data describing the strength of the bone is not causally responsible for the input data, which captures the geometrical structure of the bone.

represents a structure (that of the part of the human body which are displayed on the screen) that is causally responsible for the raw data. Because of this causal link the refined data can be reconstructed from the raw data. The assessment of refined data as empirical data and of hybrid methods as empirical methods despite their containing a theoretical component is motivated by the following reasons:

1. The content of the refined data depends on the content of the raw data and changes if the content of the raw data changes depending on the particular refinement process.
2. Hybrid methods can have all of the features that distinguish experiments from simulations as described above (see section 3): They can provide us with new empirical data, they can operate directly on the target system and they can therefore be used to test fundamental hypotheses.
3. In a sense there are no pure empirical methods, anyway. Even an observation with our barest senses is partly determined by the “apparatus” of our sense organs. But if this is true, then we have little reason to consider the measurements we make with the help of a sophisticated artificial apparatus as less empirical than the observations with our sense organs. At the same time, the distinction between purely theoretical methods (human reasoning, calculations or simulations etc.) and empirical methods (including hybrid methods) is fairly clear cut.

Therefore, the existence of hybrid methods that combine measurement with simulation leaves the distinction between simulations and experiments intact.

5 Summary and conclusion: Computer simulations as a tools for drawing conclusions from prior knowledge

Regarding the epistemic status of computer simulations, the above reasoning leads to the conclusion that computer simulations are in no way an empirical method of science but that they clearly belong on the theoretical side of science. This result is not surprising, because, after all, computers are logical machines that only allow us to draw the consequences of given premises. Therefore, the result of a computer simulation can have no more empirical content than can logically be derived from the input data and the modelling assumptions of the simulation.

Computer simulations can be likened to experiments, because for certain types of experiments – like the wind tubes that von Neumann refers to (see section 4.1) – their empirical character is of no particular importance. But other than experiments, computer simulations can under no circumstances gather new empirical data, operate directly on an empirical target system or be employed for the testing and possible falsification of fundamental natural laws.

As a consequence, it is at best a metaphorical way of speaking if computer simulations are described as “computer experiments” [Gramelsberger, 2010]. And it is misleading if computer simulations are described as a “third way of doing science” [Axelrod, 2006, p. 90] [Rohrlich, 1990, p. 507] [Küppers and Lenhard, 2005], if by this it is meant that computer simulations stand somewhere between the theoretical side and the empirical side of science, because computer simulations are a purely theoretical device. At most it can be conceded that the description as a “third way” and the metaphor of a “computer experiment” capture certain phenomenological aspects of computer simulations which concern the way simulation studies are set up, conducted and evaluated. In this respect computer simulations may indeed appear experiment-like. For they allow us to try things out and they may lead, just like experiments, to surprising and unanticipated results. But computer simulations do so only in an artificial environment in the computer and not in an empirical environment.

This has important consequences for the epistemic status and, in particular, the justification and validation requirements of computer simulations. When Wendy Parker concedes that “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’ [rather than with a computer simulation]” [Parker, 2009, p. 494], then this is true, but the emphasis is misplaced. For, it is in fact only when we already know very much about a target system in terms of comprehensive and empirically well-confirmed background theories that we can safely rely on computer simulations as a substitute for experiments. It is therefore no surprise that computer simulations are most successful in those areas of science where we have powerful background theories, like quantum chemistry for example.

By the same token, the fact that computer simulations are a merely theoretical tools renders understandable their lack of success in those areas of science that have to do without empirically well-confirmed background theories such as the social sciences, where the lack of empirical content and proper validation of simulations is frequently lamented [Heath, Hill and Ciarello, 2009]. There is a danger of underestimating the paramount importance of empirical validation of computer simulations, if they are conceived of as some kind of experimental tool. It is therefore best to think of computer simulations not as “virtual experiments” but as tools that allow us to evaluate the consequences of theories and prior knowledge.

References

- Arnold, Eckhart. 2008. *Explaining Altruism. A Simulation-Based Approach and its Limits*. Heusenstamm: ontos Verlag.
- Axelrod, Robert. 2006. *Simulation in the Social Sciences*. Vol. Handbook of Research on Nature Inspired Computing for Economics and Management. Volume 1: Nature Inspired Computing and Social Sciences Hershey (Pennsylvania): IGI Global chapter 7, pp. 90–100.
- Barberousse, Anouk, Sara Franceschelli and Cyrille Imbert. 2009. “Computer simulations as experiments.” *Synthese* 169:557–574.
- Franklin, Allan and Colin Howson. 1984. “Why do scientists prefer to vary their experiments?” *Studies in the History and Philosophy of Science* 15(1):51–62.
URL: <http://www.sciencedirect.com/science/article/pii/0039368184900293>
- Gilbert, Nigel and Klaus G. Troitzsch. 2005. *Simulation for the Social Scientist*. McGraw-Hill.
- Goumans, Theodorus P. M. and Johannes Kästner. 2010. “Hydrogen-Atom Tunneling Could Contribute to H_2 Formation in Space.” *Angewandte Chemie, International Edition* 49:7350–7352.
- Gramelsberger, Gabriele. 2010. *Computereperimente. Zum Wandel der Wissenschaft im Zeitalter des Computers*. Bielefeld: transcript Verlag.
- Hacking, Ian. 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*. Cambridge University Press.
- Heath, Brian, Raymond Hill and Frank Ciarello. 2009. “A Survey of Agent-Based Modeling Practices (January 1998 to July 2008).” *Journal of Artificial Societies and Social Simulation (JASSS)* 12(4):9.
URL: <http://jasss.soc.surrey.ac.uk/12/4/9.html>
- Hughes, R. I. G. 1999. The Ising model, computer simulation, and universal physics. In *Models as Mediators. Perspectives on Natural and Social Science*, ed. Mary S. Morgan and Margaret Morrison. Cambridge University Press chapter 5, pp. 97–145.
- Humphreys, Paul. 1994. Numerical Experimentation. In *Patrick Suppes: Scientific Philosopher. Volume 2. Philosophy of Physics, Theory Structure and Measurement Theory*, ed. Paul Humphreys. Vol. 2 Kluwer Academic Publishers pp. 103–118.
- Küppers, Günter and Johannes Lenhard. 2005. “Computersimulationen: Modellierungen 2. Ordnung.” *Journal for General Philosophy of Science* 36:305–329.
URL: <http://dx.doi.org/10.1007/s10838-006-2348-1>
- Lee, Daniel C. and Timothy J. Carroll. 2010. *Magnetic Resonance Imaging*. Vol. Practical Signal and Image Processing in Clinical Cardiology Dordrecht, Heidelberg, London, New York: Springer Verlag chapter 16, pp. 251–273.
- Morgan, Mary S. 2003. Experiments without Material Intervention. Model Experiments, Virtual Experiments, and virtually Experiments. In *The Philosophy of Scientific Experimentation*, ed. Hans Radder. University of Pittsburgh Press pp. 216–233.
- Morrison, Margaret. 2009. “Models, measurement and computer simulation: the changing face of experimentation.” *Philosophical Studies* 143:33–57.
- Parker, Wendy S. 2009. “Does matter really matter? Computer simulations, experiments, and materiality.” *Synthese* 169:483–496.
- Rohrlich, Fritz. 1990. Computer Simulation in the Physical Sciences. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association. Volume Two: Symposia and Invited Papers*. University of Chicago Press pp. 507–518.

- Senn, Hans Martin and Walter Thiel. 2009. "QM/MM Methods for Biomolecular Systems." *Angewandte Chemie. International Edition* 48:1198–1229.
- wikipedia. 2010. "Double-slit experiment."
- URL:** http://en.wikipedia.org/wiki/Double-slit_experiment
- Winkler, Karl-Heinz A., Jay W. Chalmers, Stephen W. Hodson, Paul R. Woodward and Norman J. Zabusky. 1987. "A numerical laboratory." *Physics Today* 40:28–37.
- URL:** <http://static.msi.umn.edu/rreports/1987/89.pdf>
- Winsberg, Eric. 2001. "Simulations, Models and Theories: Complex physical systems and their representations." *Philosophy of Science* 68 (Proceedings):442–454.
- URL:** <http://www.jstor.org/stable/3080964>
- Winsberg, Eric. 2003. "Simulated Experiments: Methodology for a Virtual World." *Philosophy of Science* 70:105–125.
- Winsberg, Eric. 2009. "A tale of two methods." *Synthese* 169:575–592.
- Winsberg, Eric. 2010. *Science in the Age of Computer Simulation*. Chicago and London: The University of Chicago Press.