Computer Simulations and the Changing Face of Scientific Experimentation

Computer Simulations and the Changing Face of Scientific Experimentation

Edited by

Juan M. Durán and Eckhart Arnold

CAMBRIDGE SCHOLARS PUBLISHING

Computer Simulations and the Changing Face of Scientific Experimentation,
Edited by Juan M. Durán and Eckhart Arnold

This book first published 2013

Cambridge Scholars Publishing

12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-4438-4792-5, ISBN (13): 978-1-4438-4792-6

To Kassandra Pomper, for her support and companionship
during all the stages of this book
—Juan

To Rolf Ital who spoiled me by buying me my first computer
for my twelfth birthday
—Eckhart

TABLE OF CONTENTS

INTRODUCTION

William Aspray recorded the first use of computers for scientific usage between 1952 and 1957. By June of 1952, the IAS computer was finished and ready to be tested; although it needed some extra months for repair and general maintenance, the major issue for the team of scientists and engineers was to understand the new instrument. The digital computer, built and designed on solid theoretical foundations, presented a significant challenge; namely, it was necessary to dedicate some extra time to learn the operation of the machine, identify appropriate algorithms, and determine the range of mathematical applications within the computer's capacity (1990, 155). By the time the computer became a more knowledgeable and reliable instrument, scientists and engineers began to use it with great success in specific scientific applications. By 1954, the calculation of the energy band structure of iron that would test the theory of ferromagnetism became the first scientific application to run on a digital computer (1990, 159).

In the years following 1954, the digital computer proved to be a fundamental tool for the development and advancement of scientific understanding. Today, despite their short history, computers are leaving an indelible mark on numerous and disparate scientific disciplines such as particle physics, astronomy, behavioral science, psychology, sociology, and economics. Arguably, there is virtually no scientific discipline that has not been involved, in one way or another, with the digital computer. This durable presence extends widely along the uses and needs of scientific practice. For instance, the numerical experiment of calculating the energy band structure of iron qualifies, in contemporary parlance, as a computer simulation. The main topic of this book is precisely to address the uses of and needs for computer simulations in contemporary scientific practice. In this context, computer simulations are discussed from a philosophical, historical, and scientific point of view.

Nowadays, there is a renewed interest in understanding the role that computer simulations play in scientific practice. Do computer simulations belong with the calculator and the test tube, or do they belong higher in the epistemic hierarchy, closer to theories and experiments? Are they just scientific models implemented on the digital computer, or do they represent a novel way of doing science? Given the centrality of the issue, it is not surprising to find that there have been many attempts to theorize about the nature of computer simulations as experimental devices. Admittedly, these questions have been around for quite some time. As early as 1967, Naylor, Burdick and Sasser, define a computer simulation as:

> A numerical technique for conducting experiments with certain types of mathematical and logical models describing the behavior of an economic system on a digital computer over extended periods of time (…) The principal difference between a simulation experiment and a 'real world' experiment is that with simulation the experiment is conducted with a model of the economic system rather than with the actual economic system itself (1967, 1316).

It is astonishing to note the similarity of this quotation with more contemporary literature on the topic. Current philosophical inquiry also engages in similar efforts, such as distinguishing between a computer simulation and a 'real world' experiment, or exploring the methodological implications of implementing a scientific model as a computer simulation.

Yet, despite these few similarities, much of the contemporary philosophical investigation is simply not the same as in the late 1960s. From a historical perspective, the introduction of silicon based circuits, and the subsequent standardization of the circuit board significantly helped the industry and the growth in the computational power of computers. Such growth in speed of calculation, size of memory, or the number of programming languages forcefully challenged the established ideas and encouraged the seeking of new questions and answers.

One of the leading questions on this issue has been whether computer simulations stand for a new way of conducting scientific practices, or if they simply represent another computational method subsidiary of experimentation. The work of Rohrlich (1990) sets the grounds in this direction. He argues, computer simulations do provide a qualitatively new methodology for the physical sciences, lying somewhere intermediate between theoretical physics and empirical methods of experimentation.

However, Frigg and Reiss (2009) deliver the most pressing contemporary discussion on the philosophical relevance of computer simulations. The authors understand computer simulations in the context of the philosophy of models and, as such, with no significant distinctions from other uses of modeling in experimental practice. Humphreys (2009) answers their skepticism by indicating that the way the argument is presented is misleading, for it illuminates only computer simulations from the perspective of a philosophy of models. To Humphreys' mind, computer simulations raise questions that cannot be answered by a familiar philosophy, but rather need to be addressed at face value.

Beyond the specific contribution that this discussion can offer to the philosophical study of computer simulations, there is general agreement that computer simulations raise important questions for the general philosophy of science. One interesting example is the search for general criteria that distinguish computer simulations from experiments. Such a search has ramifications on studies about the epistemic power of computer simulations, the ontological and epistemological status of simulation data, the importance of new methodologies involved in the design and building of a computer simulation, and similar questions.

From September 21st to September 23rd 2011, the interdisciplinary workshop "Computer Simulations and the Changing Face of Scientific Experimentation," sponsored by the University of Stuttgart and the Stuttgart Research Center for Simulation Technology (SRC SimTech), brought together philosophers, historians, sociologists, and scientists into a common discussion with the purpose of revisiting some of the questions here mentioned, and addressing the new challenges that computer simulations pose to scientific practice.

We have divided this book into three mutually related parts. Part One (Theory) is dedicated to the theoretical understanding of the relation between simulations and experiments in the current philosophy of science. Part Two (Practice) fleshes out some of the theoretical conceptualizations presented in Part One by illustrating case studies from current scientific research on computer simulations. These case studies highlight the shift from experiments to computer simulations that is observed in current scientific practice, and describe the patterns of interaction between simulation methods and experimental methods in current scientific research. Part Three (History) broadens the perspective by offering case studies on the historical development of "computer experiments" as a research method.

The first part of the book is dedicated to the diversity of views among philosophers regarding existing distinctions between computer simulations and experiments, the epistemic power of computer simulations, and the new methodologies that they represent.

In the first contribution ("What Are Data About?"), Paul Humphreys calls our attention to the discussion about the status of data produced by a computer simulation. His paper focuses on the content of data produced, instead of the source that produces such data. According to the author, the origins and modes of production of these data show that the empiricist point of view is no longer an attainable position in the philosophy of science. This argument derives its force from what the author calls ‘causal-computational instruments’; that is, an instrument that relies on a causal process that links the data source in nature with the measurement, but that also requires further post-processing for rendering reliable data. In Humphrey’s mind, then, such causal-computational instruments cannot be interpreted in the same way as Hacking discusses microscopes, where a realist interpretation of the images is justified by the independent access to the same phenomenon through different observational instruments. The decisive point here is that the data delivered by a causal-computational instrument, like a CT scan, are the result of deliberate engineering. Depending on the particular purpose, say, whether the data is meant to be “read” by a human agent or further processed in the computer, the appearance of the engineered data may differ considerably. In order to determine its representational content, it is therefore central to take into account the origin of the data as well as the engineering steps by which it is formed (and transformed). Causal-computational instruments, then, pose a significant challenge for philosophers interested in traditional problems of empiricism, realism, and the notion of data.

If Humphreys reminds us that there is a considerable amount of engineering involved in the production of the empirical data by causal-computational instruments, Anounk Barberousse and Marion Vorms (“Computer Simulations and Empirical Data”) attack the problem from the opposite side; that is, by examining whether the data produced by a genuine computer simulation can, with any good reason, be considered empirical data. Starting from the assumption that empirical data are about physical systems, Barberousse and Vorms challenge the opinion that the data produced by computer simulations cannot be new or surprising. It is frequently assumed that computer simulations, because they rely heavily on pre-existing theoretical background knowledge of the simulated objects, are less capable of producing genuinely novel and surprising insights about their target system than observations or traditional experimentation. The authors support the claim that this assertion is mistaken with the example of computer simulations of deterministic chaos.

While this conclusion emphasizes the capacity of computer simulations to produce empirical data that are as novel and surprising as that of experiments or observations, Eckhart Arnold points out the differences that remain between simulations and experiments as scientific methods (“Experiment and Simulations: Do They Fuse?”). Most notably, he argues that the results produced by computer simulations cannot go beyond what lies in the deductive closure of their premises. According to Arnold, a simulation, unlike a material experiment, cannot be employed as an *experimentum crucis*. The chapter therefore contains an elaborated criticism of some, in Arnold's opinion, misguided philosophical conceptualizations of computer simulations. With respect to the borderline between simulations and experiments, however, one question remains that is not so easily dismissed: How can a measurement that involves the computational refinement of its data properly be distinguished from a computer simulation that makes use of input data of empirical origin? To this question, Arnold gives a tentative answer based on the *measuring a cause by its effect pattern*, a pattern that is typical for many traditional measurement methods already.

The contribution by Juan M. Durán ("The Use of the 'Materiality Argument' in the Literature on Computer Simulations") continues the discussion on the differences between computer simulations and experiments, but this time from a meta-critical point of view. Durán's main concern is to unpack the underlying rationale that has been guiding the argumentation in current literature. By addressing the so-called "materiality argument" present in three different conceptualizations, the author shows that there is a common argumentative structure that inevitably shapes the final epistemological evaluation of computer simulations. Specifically, Durán presents what he calls 'the materiality aftermath,' a meta-criticism that exposes the rationale underlying the arguments in the current literature on simulations. In the author's mind, 'the materiality aftermath' is the result of the philosopher's ontological commitment to computer simulations, from which epistemological consequences are drawn. The author believes that adapting the philosophical investigation to this rationale leads to a conceptual corset in the inquiry of the epistemology of computer simulations. Durán's conclusion is sober, and aims at endorsing the philosophical investigation on computer simulations as neither restricted by, nor limited to, ontological commitments, but rather addressed at face value.

The contribution by Pío García and Marisa Velasco ("Exploratory Strategies: Experiments and Simulations") turns the discussion to a notion of 'exploratory strategy' applicable to computer simulations. Particularly, the authors analyze exploratory strategies in experiments and computer simulations, and elucidate the methodological and epistemological role in both domains. Their proposal, then, consists first in drawing some distinctions between computer simulations and experiments. Second, the authors make explicit the concept of 'exploratory strategy,' establishing a further distinction between exploratory experiments and other types of experiments. This second step allows them to present their own proposal as a different way to approach the epistemic and methodological aspects of scientific practices, particularly, computer simulations. Some relevant cases of experimental and simulation activity are considered in the context of 'exploratory strategies.'

In the second part of the book, the focus is shifted from the abstract and theoretical philosophical discussion to the analysis of concrete examples. The first of these papers is the study of simulations of cardiac electro-physiology by Annamaria Carusi, Blanca Rodriguez and Kevin Burrage ("Model Systems in Computational System Biology"). Their case study concerns multi-scale models of cardiac electro-physiology. These models represent a challenge from a technical as well as a philosophical point of view. Defying any sharp distinction between simulations and experiments, the authors claim that "the basic unit of analysis when considering questions of the validation and epistemic warrant of computational methods in systems biology" is the model-simulation-experiment-system (MSE). In particular, the target system cannot be understood simply as a given reality, rather it is co-constructed with the MSE system. The construction of the target domain is inevitable because the validation data need to be comparable to the MSE system. However, the term 'construction' must not be misunderstood as implying a relativistic understanding of science in this context. The validation experiments remain independent in the sense that they do not make use of any data that have been used for model construction.

Anne Marcovich and Terry Shinn's contribution ("Computer Simulation and the Growth of Nanoscale Research in Biology") explores three links between computer simulations and nanobiology research. First, they show that there is a correlation between nano-related biology publications in the early 1990s and the introduction of computer simulations in scientific practice. Second, computer based research contributes to the cognition of nanobiology through the creation, organization, and consultation of databases. Finally, the authors show that "simulation molecular graphics generate images that are informationally and analytically rich, and that offer a fundamental input into novel forms of epistemology." Their contribution shows not only how the academic agenda is strongly driven by the introduction of new technologies, but also how computer simulations can provide a genuine understanding of their simulated target system, requiring a novel form of epistemology.

In their contribution, Lucía Ayala and Jaime Forero-Romero ("Computer Simulations in a Cosmological Context") discuss the case of testing hypotheses in cosmology. Physical cosmology represents a special case in the natural sciences with regard to the available methods for testing a hypothesis. Since direct experiments are excluded, observations and simulations must carry out this testing function. In their contribution, the authors discuss the special case of numerical simulations as an essential tool for understanding the observed large-scale structures in the Universe. This discussion is followed by a description of the limitations of simulations in understanding such large-scale structures. For instance, the physical nature of computer simulations becomes a limitation. As the authors point out, time, data storage, and data transfer rates are restricted. Ultimately, theory, observations, and simulations work together and, with their different potentials and limitations, mutually complement each other in contemporary astronomy.

Muniza Rehman traces the latest developments in the use of simulations and experiments in the pharmaceutical industry ("Experimentation and Simulations in the Pharmaceutical Industry"). Rehman places simulations between traditional experimentation and theoretical accounts. To the author's mind, two kinds of simulation studies are common in the pharmaceutical industry: Computer-assisted trial designs (CATD) and computer-simulated clinical trials (CSCT). The former are employed to study the experimental design of clinical studies, before they are conducted. The latter are used to estimate the outcome of clinical trials, potentially rendering some of these trials unnecessary and thus reducing the number of clinical trials that actually have to be conducted. Some philosophers have disputed that simulations provide a true novelty over traditional modes of modeling and theoretical exploration. Nevertheless, given how strongly the use of computer simulations has affected the practice of drug testing in the pharmaceutical industry, Rehman concludes that from this perspective simulations are indeed a *sui generis* activity in a Humphreyan sense.

The third and last part completes the book with historical case studies. Wolfgang Brand ("Designing the Membrane Roof of the Munich Olympic Stadium using Supercomputers") presents a historical case study of the deployment of the first supercomputers in architecture and civil engineering. The events around the design of the tent-shaped membrane roof of the Munich Olympic Stadium for the 1972 Olympic Games demonstrates how physical models of constructions enable technologies for the construction of naturally shaped buildings. It is argued that the 1960s mark the period in which the usage of high performance computers triggered the change toward architectural design processes. The technology available had already reached a state where model building was no longer necessary. It is shown how two groups using different methods on the same computing infrastructure designed the roofs inspired by the ideas of Frei Otto. They developed wide-spanning lightweight structures consisting of pre-stressed cable nets covered by transparent tiles. The group of John H. Argyris relied on the Finite Element Method, which he co-invented. While another group headed by Klaus Linkwitz used least-square fitting and developed the new Force Density Method, all influenced by geodesic methods. Both attempts were successful and led to the landmark Olympic Stadium in Munich, as we know it today.

A somewhat different perspective on simulations is introduced by Michael Resch ("What's the Result? Thoughts of a Center Director on Simulations"). As head of the high-performance computing center in Stuttgart, Resch addresses the technological procedures (and their limitations) by which simulations are implemented and executed on the computer. In this respect, Resch proposes an addition to Winsberg's (2010) layered model of simulations, which also includes numerical schemes, program structures, programming models, and hardware architectures. All of these influence the capabilities as well as the limitations of the simulation approach. Resch, then, embeds his 'prototypical workflow' into a broad philosophical perspective, covering the question of verification and validation, as well as the need for rendering simulation results comprehensible to human beings. The latter issue does not only concern the specialist user of simulations, but also is of interest for society –as the example of climate simulations may illustrate.

We hope that readers from different humanistic and scientific fields that concern themselves with computer simulations find the broad perspective of our book useful. The editors would like to thank the University of Stuttgart and the SRC SimTech for financial support that made the workshop possible. This book is a publication of the papers presented at that workshop. We are in debt to the participants for making the workshop a successful event. Most of all, we would also like to thank all the authors that, with their excellent contributions, made this book possible.

Juan M. Durán and Eckhart Arnold

References

Aspray, William, *John von Neumann and the Origins of Modern Computing*, The MIT Press, 1990.

Axelrod, Robert, *The Evolution of Cooperation*, Basic Books, 1984.

Barnerousse, Anouk et al. "Computer Simulations as Experiments", *Synthese*, 169, (2009): 557-574.

Ceruzzi, Paul E., *A History of Modern Computing Computer*, The MIT Press, 1998.

Frigg, Roman and Reiss, Julian. "The Philosophy of Simulation: Hot New Issues or Same Old Stew?" *Synthese*, 169, (2009): 593-613.

Giere, Ronald, "Is Computer Simulation Changing the Face of Experimentation?" *Philosophical Studies*, 143, (2009):59-62.

Humphreys, Paul W. "The Philosophical Novelty of Computer Simulation Methods", *Synthese*, 169, (2009): 615-626.

Morgan, Mary, "Experiments versus Models: New Phenomena, Inference and Surprise" *Journal of Economic Methodology*, 12, (2005):317-329.

—. "Experiments without Material Intervention" In *The Philosophy of Scientific Experimentation*, edited by Hans Radder, 216-235. University of Pittsburgh Press, 2003.

Morrison, Margaret, "Models, Measurement and Computer Simulation: The Changing Face of Experimentation" *Philosophical Studies*, 143, (2009):33-57.

Naylor, Thomas H. et al. "Computer Simulation Experiments with Economics Systems: The Problem of Experimental Design" *Journal of the American Statistical Association*, 62, (1967): 1315-1337.

Parker, Wendy, "Does Matter Really Matter? Computer Simulations, Experiments and Materiality" *Synthese*, 169, (2009):483-496.

Rohrlich, Fritz, "Computer Simulation in the Physical Sciences", *PSA: Proceedings of the Biennal Meeting of the Philosophy of Science Association*, (1990): 507-518.

Simon, Herbert A., *The sciences of the Artificial*, MIT Press, 1996.

Winsberg, Eric, *Science in the Age of Computer Simulation*, University of Chicago Press, 2010.

PART I:

THEORY

CHAPTER ONE

WHAT ARE DATA ABOUT?

PAUL W. HUMPHREYS
UNIVERSITY OF VIRGINIA

Empiricism is no longer a tenable position in the philosophy of science. As a result, it is worth examining what it was that made empiricism such an attractive position for so long and if anything valuable can be salvaged from that tradition. The debates about empiricism usually contrasted knowledge obtained from observation and experiment on the one hand with knowledge obtained from theories on the other. Because computer simulations are firmly entrenched as a third mode of pursuing scientific inquiry, one way to explore what made empiricism important is by contrasting data that are provided by experiments and observations with data that are generated by computer simulations. In doing so, my paper will indirectly address one of the original philosophical issues about computer simulations: in what ways, if at all, do computer simulations differ from scientific theories on the one hand and experiments on the other?[1] Early in the discussions, claims were made that simulations had some kind of intermediate status between theory and experiment, while also standing as sui generis methods. More recently, claims have been made that simulations can be used in place of material experiments under certain circumstances.[2] Although it is true that there are similarities between certain aspects of simulations and experiments, pointing out analogies between laboratory experiments and computer simulations, such as the ability to manipulate variables and control for confounders, does not address one of the central epistemological questions that arise once

---

[1] This issue was present almost from the inception of computer simulations. Explicit attitudes towards it can be found in (Ord-Smith and Stephenson, 1975, 3), (Rohrlich, 1991), (Humphreys, 1994).
[2] See Norton and Suppe (2001), Winsberg (2003), Parker (2009).

simulations are brought into the picture. This question is: Are the data produced by computer simulations different in kind and in content from experimental and observational data, and from data generated by traditional scientific or mathematical theories? If we have reasons to agree that they are different in a scientifically relevant way, then that is one dimension along which simulations occupy a distinct scientific niche. Empiricists have usually treated the issue of the empirical source of data within a methodological context such as confirmation, verification, or falsification. I shall focus instead on content. One important question is this: what is it for a datum to have empirical content?[3] An answer to that question is provided in (Humphreys forthcoming); here we can make additional progress by addressing a related question: what are various kinds of data about?

What are Data?

In order to remain as ontological neutral as possible about data, I take a datum to be the value of a variable. The term ‘variable’ will be used here in a way that is neutral between items such as a mathematical function that represents a property and the property itself.[4] This dual use carries with it certain dangers because the role of representations in computer simulations is crucial, but where appropriate I shall explicitly note which use is in play. The variable can be scalar and discrete valued, which covers situations in which the datum concerns a qualitative monadic property such as “is red,” or it can be vector and continuous valued, capturing relational features such as “has velocity v with respect to frame F.” Other possibilities can be accommodated. I shall not distinguish between atomic and non-atomic data because nothing that follows depends upon making that distinction. Finally, although the expression “data” often carries the force of something given, something fundamental, those connotations must be rejected. Data can be the result of processing, transformations, and interpretation, and we can and often do question the data.[5]

---

[3] There are other epistemological issues about simulations and experiments including: 1) the a priori versus the a posteriori content of data from each, 2) the empirical versus the formal content of data from each, and 3) the relative rates of reliability as truth generators for data from each.

[4] The variable can be purely formal and hence represent nothing.

[5] The distinction between data and phenomena drawn in Bogen and Woodward (1988) and other papers is compatible with the definition of a datum given here, although their emphasis on the causal production of data perhaps indicates a narrower use of the term “data” than is considered here.

Data can be assessed on their own terms, without regard for what generated them and an important aspect of our definition is that it does not mention the origins of a datum, allowing data to originate from computational processes, from experiments, from theory, and perhaps other sources. Yet, if we are to address the question of what the data are about we must solve the *inverse inference problem*. The inverse inference problem consists in making, and providing a justification for, an inference from the data to its source. In the debates between scientific realists and empiricists, the issue is usually cast as one of whether terms in scientific theories that purport to refer to unobservables genuinely refer, but it can be recast as the problem of what objects' and properties' existence can justifiably be inferred from the empirical data. In these terms, inferring the existence of Saturn's rings and their properties from what is observed through a low powered telescope is an inverse inference problem, as is inferring the existence of a virus from an electron microscope image. The converse of the inverse inference problem, the direct inference problem, is the problem of what data will be available given the existence of the source.

To see how addressing the inverse inference problem and assessing the content of data make a difference to how we evaluate data, consider the traditional division between empiricism and rationalism.[6] For empiricists, data that are the result of direct perceptual experience, or on a slightly more liberal agenda, data that are a result of observations by elementary equipment that include the human senses, are the most desirable and, for many, are the only source of genuine knowledge. The reasons for this desirability vary. On the one hand there is a widely shared belief that the origins of the data in the causal world make their content more desirable than the content of data whose origins are whatever produces a priori knowledge. On the other hand the empiricists' starting point was the content of the datum and not its origin, an orientation that deliberately left open the possibility that the external world might not be the source of the empirical datum and might not even exist, leading either to a lack of commitment to the existence of that data's origins, as in constructive empiricism, or to an outright denial, as in idealism.

[6] What constitutes empiricism and rationalism is, and probably forever will be, a matter of scholarly dispute. I am using the terms here as surrogates for broad epistemic attitudes that I assume most philosophical readers will recognize. For the record, I subscribe to the view that although it is often a matter of historical contingencies when a particular philosophical issue is raised and becomes the subject of focused discussion, the issue itself transcends those historical contingencies.

Empiricists have granted privileged status to observational data for a number of other reasons. One was that data about directly observable entities seem to have the certainty that was lacking in data that were about unobservables. This certainty was a reason for refusing to provide a solution to the inverse inference problem and it was the intrinsic content of the data to which any certainty attached. A second reason was that data about observables were supposed to act as a theory independent basis for deciding between rival theories. On what might now seem to be rather naive grounds for taking the intrinsic content of a datum to be theory-independent, this gave privileged status to such content, and a bonus was that by avoiding making inverse inferences, no theory was needed in that capacity either. A third reason was that empirical data were taken to be the only reliable source of information about contingencies existing in our world; a priori methods were incapable of that degree of specificity. For various reasons, all of which are plausible, the first and second of these reasons no longer have the force once attributed to them and in light of the well-known arguments formulated by Quine (1951), the distinction between the a priori and the a posteriori is now seen to be a much more difficult distinction to make than was originally imagined.

It is a different issue that lies behind some of the difficulties in assessing the status of data from simulations and experiments. The issue is the extent to which inverse inference problems need to be solved in order to decide what the data are about. One set of solutions to this problem, following the empiricist tradition, attempts to attribute content to the data without taking into account their origins. This approach starts with data and avoids making inferences about their origins as far as is possible. If the content of data from simulations and from experiments is equivalent under this approach, then data from the two classes are inter-substitutable. Thus, if a simulation of independent tosses of a coin with parameter *p* is based on an accurate model of a sequence of tosses of a real coin with that degree of bias, the data from the simulation can replace the data from the experiment and we can ignore the origins of the data.[7] Another set of solutions suggests that the origins of data in material systems make those data about something different than data coming from a computer simulation and so inferences that are often not easy to justify are required to use data from simulations in place of data from experiments. These are complicated issues and I can only sketch a solution here, but the overall view is that the origins and mode of production of data must be taken into account.

[7] This is the position taken by Kästner and Arnold (2012) in which well-confirmed theories play a central role. See also Winsberg (2009)

Simulations and Experiments

In recent years, there has been considerable discussion about whether computer simulations can serve as a replacement for material experiments. Those who have argued for relevant similarities between simulations and experiments tend to emphasize methodological considerations.[8] Barberousse, Franceschelli and Imbert (2009) (hereafter BFI) have drawn an important distinction between two types of data, data_E and data_A. BFI define data_E as being ‘of empirical origin, namely produced by physical interactions with measuring or detection devices’ (2009, 560). It seems clear from this definition, and also from the examples used to illustrate the definition, that data_E are data produced by purely causal instruments. In contrast, data_A are about a physical system. BFI note that data_A may be produced by data_E ‘but also via other processes, among them analytical or numerical pen-and pencil-computed solutions of systems of equations representing the target systems, and computer simulations’ (2009, 560) In the present context I shall take simulations to be like traditional pencil and paper solutions in the sense that they are drawing out consequences of formal representations.[9]

The distinction drawn by BFI is important and helpful and their insistence that it is the representational aspects of computer simulations that constitute the dividing line between experiments and simulations is exactly right, but we shall need to see how the distinction plays out in the realm of causal-computational instruments (see section “Causal-computational instruments” below). The distinction also opens up some important

[8] In this paper, ‘simulation’ refers to a digital computer simulation and ‘experiment’ refers to a laboratory experiment. In the latter, all known relevant variables except for explicitly specified independent variables are controlled and the manipulations of the independent variables are epistemically transparent in the sense that the causal effects of the manipulations on those variables are known. The point of the experiment is then to identify the causal effects of the independent variables on the dependent variables. The situation with a single independent variable and a single dependent variable is a special case.

[9] This assertion is consistent with my position (Humphreys 1994, 2004) that the physical implementation of computer simulations places constraints on simulation methods that are not present in traditional a priori mathematics and that epistemic opacity, including the need to make inductive inferences in place of deductive inferences, is usually present. Despite some claims in the literature to the contrary, I have never endorsed the view that running simulations on material computers is itself a reason to justify substituting data from simulations for data from experiments. Numerical experiments are significantly different from material experiments.

philosophical questions. One is how to interpret data that have transformations applied to them after their origination. Suppose we grant that we can correctly specify what counts as a measurement or a detection device.[10] Then, in the case of $data_E$, consider what happens when a representation of an empirical datum has a formal transformation applied to it. Suppose that we have a square divided into two so that the left hand side is white and the right hand side is black. An imaging device (consider a digital camera for simplicity) takes a photograph of the square and forms a digital visual image that duplicates the original square. The image is unquestionably a representation of the square and the data, which are a spatial array of black and white pixels, are about that square. It is easy to perform a formal transformation on that data set so that all of the pixels on the right hand side are transformed from black to white and all of the pixels on the left hand side are transformed from white to black. What is this second image a representation of and what is it about? Exactly the same image could have been obtained by a purely causal process by using a mirror to produce the left-right inversion. So one answer to these questions is that it is a mirror image of the original square, hence a representation of it and that the $data_A$ are about the original square. Now consider the case where only the formal transformation of the right hand side from black to white is carried out. The resulting image is a completely white square. What is this a representation of and what is it about? A variety of answers are plausible. To preserve consistency with the first and second cases, it seems we should give the same answers: The purely white square is a representation of the original square and the data are about it. Yet, it is such an extremely poor representation that one wonders in what sense it counts as a representation at all.

To see more clearly what is at issue, suppose that we have a digital photograph of a couple, Jack and Jill, against a white background. A computer algorithm removes the pixels representing Jack, replacing them with white pixels, leaving only a visible image of Jill. This image contains $data_E$ according to the above definition, and this is surely correct, but what is it about? Most people would say that the image, which consists of a spatial data array, is about Jill. Clear enough, although this answer deviates from the criteria we used for the black and white squares example. So now consider a parallel example in which the original photograph is of Jill alone, but an algorithm transforms white pixels into a colored array that is a representation of Jack. What is this new image about? It is a representation of Jack and Jill, and it is therefore about Jack

[10] This is not at all easy and I shall not attempt to solve the problems here.

and Jill, although it is a photograph of Jill only and so, on an origins view of data content, about Jill only.[11] The pixels that make up the image of Jill are all $data_E$ as well as $data_A$, whereas the pixels that make up the image of Jack are $data_A$ only.

We have here a familiar set of philosophical issues. $Data_A$ can be about the causal sources that give rise to $data_E$. Emphasis on the causal origins of the data, typified by causal theories of reference and perception, lead to one set of answers regarding what $data_A$ are about. But what $data_A$ are about can have nothing to do with the relevant $data_E$ and the interpretation is imposed by the intentions of an interpreter of the data. We thus need to say more regarding what the $data_A$ are about. Under the causal view, for an individual datum we can plausibly say that it is about whatever gave rise to that datum regardless of the accuracy of its content. In this there is an echo of the causal theory of reference in that all of the descriptive content of a piece of referential apparatus can be wrong and yet that apparatus can successfully refer. Thus, rather than begin with the datum itself, we begin with a realist attribution of the existence of the source. The inverse inference to that source is underdetermined, but this is an additional complication that is unavoidable and I set it aside here.[12] The underlying problem here is this: when philosophers still believed in pure observations, the idea was that such things gave us direct access to what was being observed. In contrast, we required inverse inferences to know what the referents of theoretical terms were. That view about direct access seems quite naive now, but we can retain one element by highlighting the fact that there is a causal pathway connecting the observation with the entity observed. Yet, we lose that causal pathway not only with simulations but also with a widely used class of imaging devices. The point here is that what data are about is a vexed and complicated issue that is intimately tied to an adequate theory of reference. BFI were right to draw our attention to this aspect of the simulations versus experiments debate. The origins of the data, whether material or not, are insufficient to determine the content of $data_A$. So let us generalize the concept of $data_E$ to $data_O$ where $data_O$ are data generated either by causal or computational sources. Here the 'O' indicates that the origin of the data be included in a specification of the data.

[11] Definitions of 'photograph' stipulate that the image must have been formed by electromagnetic radiation (usually visible light) falling on some recording device.

[12] This is not to suggest that underdetermination problems in inverse inference methods are unimportant. Both theoretically and in practice solutions to these problems must be found.

Causal-Computational Instruments

To help clarify matters, it is useful to consider a particular type of instrument, those that I shall call causal-computational instruments. Almost all discussions of scientific instruments implicitly restrict themselves to what I shall call non-computational instruments.[13] By a non-computational instrument I mean that the instrument receives some physical process as an input, the instrument causally interacts with the input to transform it, the instrument's output is another physical process, and none of these processes or interactions is a computation.[14] All of the familiar scientific instruments discussed in the philosophical literature are of the non-computational kind: optical telescopes and microscopes, magnetometers, oscilloscopes, and so on.[15] In the last fifty years or so, a potentially different class of instruments has been developed that I shall call causal-computational instruments. These take physical processes as inputs and at some point in the operation of the instrument, they convert physical states into digital representations that undergo computational transformations before producing the instrument's output.[16] Of course, these causal-computational instruments have causal aspects not only because of their inputs but because the implementation of the computations is carried out by causal processes. Yet, causal-computational instruments fall into a class intermediate between purely causal instruments and computer simulations because inferences and representations play a crucial role in their operation but, unlike pure simulations, the causal inputs to the physical device also play a central role in the interpretation of the output.

---

[13] One of the few exceptions is Israel-Jost (2011).

[14] For our present purposes, what counts as a computation will involve only those in the class of Turing computable discrete functions. This rules out the view that all physical processes are computations and provide the basis for a principled distinction between computational and non-computational instruments.

[15] I am in this paper excluding the human perceptual system as an example of a scientific instrument because it is too difficult to disentangle interpretations of the datum from the causal processes that lead to the datum, although in a more general context there are epistemological advantages to viewing the human perceptual apparatus as simply another instrument that produces data.

[16] Many traditional instruments now use digital displays for their outputs but that does not by itself introduce a computational element into the instrument. Although the distinction is perhaps not easy to make completely clear, an instrument in which the display types are antecedently fixed does not count. Under computational theories of vision, parts of the human perceptual system may count as a computationally enhanced instrument.

I shall take as examples of causal-computational instruments the category of medical imaging devices that includes computed tomography (CT) and positron emission tomography (PET) instruments.[17] Although the physical operation of specific types of instruments is crucial for understanding how they produce data, most of the philosophical points I make here generalize from the specific examples discussed. A generic diagram of scientific instruments is given in Figure 1-1.

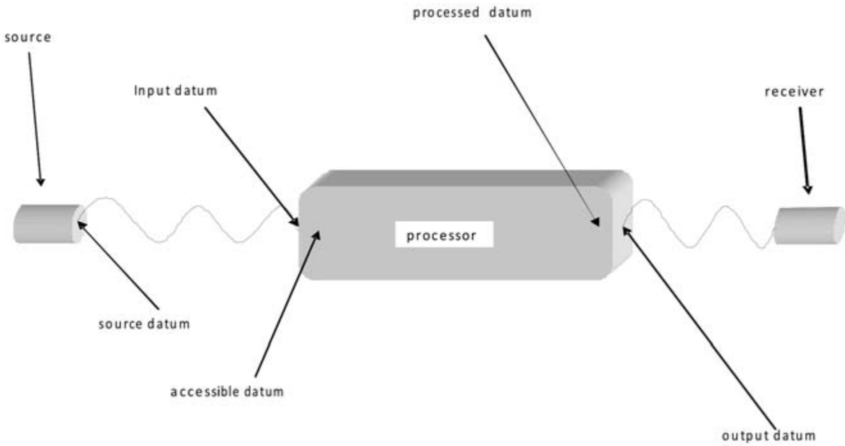

Figure 1-1

What I have called the processor can be either a purely causal transformation device, such as a telescope lens, or a computational device. The generic case that I consider has the source as an object with a single spatially varying quantitative property represented by a continuous or discrete function $f(s)$ on the space $\mathbb{R}^2$ or $\mathbb{R}^3$. Values of f are the source data. For concreteness, take as the running example the situation in which f represents the intensity of X-rays in a spatial region or the spatial distribution of some radioactive biological marker, where the spatial region includes some target such as a human body. The task is then to estimate the mathematical form of f or specified values of f using the receiver data. The input data are often the result of complex physical

[17] The principal use of PET scans is for imaging of brain tumors, epilepsy, strokes, and Alzheimer's disease. Magnetic resonance imaging (MRI) devices use different methods than do PET and CT devices.

processes within the system that must be modeled in order to infer both the general form and the specific values of *f*.

Both CT and PET instruments construct a two dimensional (sometimes three dimensional) image from a sequence of one dimensional projections. The construction process, which is inescapably computational, involves a set of inverse inferences from the receiver data to the source of that data. A number of different mathematical techniques are used for these inferences (here I shall discuss one of the most frequently used methods, filtered backprojection). Although such inferences run against the primary direction of causation from input to output, this does not violate the causal component of these instruments. Similar inverse inferences are made in purely causal instruments, such as refracting telescopes, to the conclusion that the image at the eyepiece is an image of the source object.

In two dimensional computerized tomography instruments, X-rays, collimated to lie in a plane, traverse the object to be imaged and impinge on detectors on the far side of the object.[18] Each detector receives a one dimensional projection of the target object along a given ray, and the computational algorithms combine all the projections around a 180º arc to construct a two dimensional image of a cross section of the target. The energy of the X-rays is attenuated by traveling through the object, and the degree of attenuation depends upon the densities of the materials through which the X-ray is traveling. Although Hounsfield's CT prototype used matrix inversion methods, these are no longer used to recover the values of attenuation coefficients because there is a relatively high level of noise in the projections and this can cause instabilities in direct inversion techniques. In addition, the large amount of data collected makes the computational load on matrix inversion methods infeasible. The choice of mathematical techniques is thus affected by both technological constraints, and the fact that the physical system does not satisfy the idealizations needed for matrix inversion to be effective. Instead, backprojection algorithms or iterative methods are used.

The backprojection methods that make inverse inferences from the detected intensities to the attenuation coefficients use inverse Radon transforms.[19] The basic idea is that the total attenuation along a ray is the sum of the attenuations in each pixel, and the backprojection method adds back the attenuation in each voxel by performing a line integral along the direction of the ray. By taking rays in many different directions, the 2-D

---

[18] For simplicity, I take the X-rays to be parallel rather than distributed in a fan-shaped beam.

[19] I note that some of these mathematical techniques had been developed previously for use in astronomical imaging using radio telescopes.

matrix of pixels can be reconstructed. But the bare mathematical method assumes that the physical processes are idealized in certain ways and in order to eliminate artifacts one needs to know how the image was constructed.

In order to argue for the view that the generating conditions of the data must be known, consider an argument that Ian Hacking (1983) has used in favor of entity realism. The argument goes like this for the case of microscopes. It is sometimes possible to observe the same structure with the aid of microscopes that use different, independent, physical processes, such as ordinary optical microscopes, fluorescent microscopes, interference microscopes, polarizing microscopes, and so on. Hacking argues that it would be incredible to assert that there was no common physical structure that was giving rise to these common observations from different instruments: "If the same structure can be discerned using many of these different aspects of light waves, we cannot seriously suppose that the structure is an artifact of all the different physical systems" (Hacking, 1983, 147). This argument is flawed because it does not properly take into account the fact that the observed structure is deliberately engineered.[20] We can easily see this in the case of the medical imaging techniques discussed here. Consider the example of a sinogram, which is a representation of the raw data produced from a CT scan with the frame of reference attached to the detectors and which rotates around the target object. The intensity of the radiation received at a detector is plotted against the angle of rotation of the radiation source relative to a fixed baseline in the object's frame of reference.[21] To almost all readers of this essay, sinograms do not represent anything familiar. However, when inverse Radon transforms are applied to the pixels constituting the sinogram, it is transformed into something familiar, such as an image of a human skull but that familiar image has its 'obvious' representational structure imposed by choices of the instrument designers. The intentional content is useful to us because of the perceptual apparatus of human observers, but for a computer, the sinogram is at least as useful a representational device and results from a coordinate transformation between the two frames of reference. We could say that the sinogram and the familiar image of the skull are both in an equivalence class of representations where the equivalence relation is determined by a set of

[20] For arguments that this point also applies to traditional causal instruments, see (Humphreys, 2004, 33-37). One difference between causal and causal-computational instruments in this regard is the relative ease with which images can be constructed in the latter instruments.

[21] For images of sinograms see Webb (2003).