



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Computer simulations and experiments

Citation for published version:

Massimi, M & Bhimji, W 2015, 'Computer simulations and experiments: the case of the Higgs boson', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, vol. 51, pp. 71-81. <https://doi.org/10.1016/j.shpsb.2015.06.003>

Digital Object Identifier (DOI):

[10.1016/j.shpsb.2015.06.003](https://doi.org/10.1016/j.shpsb.2015.06.003)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics

Publisher Rights Statement:

© Massimi, M., & Bhimji, W. (2015). Computer simulations and experiments: the case of the Higgs boson. *Studies in history and philosophy of modern physics*. 10.1016/j.shpsb.2015.06.003

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Computer simulations and experiments: the case of the Higgs boson

Michela Massimi and Wahid Bhimji

Abstract

Simulations have been at the center of an important literature that has debated the extent to which they count as epistemologically on a par with traditional experiments. Critics have raised doubts about simulations being genuine experiments, on the ground that simulations seem to lack a distinctive feature of traditional experiments: i.e., the ability to *causally interact* with a target system. In this paper, we defend the view that simulations are indeed epistemologically on a par with traditional experiments. We first identify three possible ways of understanding the *causal interaction claim*. We then focus on the use of simulation in the discovery of the Higgs boson to show that in this paradigmatic case, simulations satisfy all three possible readings of the causal interaction claim.

1. Introduction. Simulating and experimenting

Over the past few years, computer simulations have attracted the increasing attention of philosophers of science working on models and the epistemology of experiments. This is a blossoming research field, where pressing issues about the calibration, validity, and reliability of computer simulations are raised in areas as sensitive as climate science, for example. But what is so special about computer simulations? Are computer simulations the twenty-first century face of experimentation? Do computer simulations enjoy a distinct—and more debatable—status compared to ordinary experiments? This is the key *epistemological* question that we address in this paper.

While the systematic use of computer simulations—from climate science to high-energy physics—is undeniable, the epistemology of simulation is more of a contentious issue. Computer simulations seem a new kind of experimental activity in need of a distinctive epistemology (for some early work in the field, see Humphrey 1991, 1994). Yet critics have drawn sober conclusions about the allegedly special epistemic role of computer simulations (for a survey, see Frigg and Reiss 2009). How do computer simulations differ then (if at all) from ordinary experiments?

The recent detection of the Higgs boson is one of the most illuminating examples of how computer simulations have become an integral part of experimentation in high-energy physics. It also raises interesting questions about the epistemic status of simulations and their interplay with ordinary experiments. Our paper has a twofold aim. First, we illustrate the use of simulations in two key features of the discovery of the Higgs boson: namely, the background determination (necessary for identifying the occurrence of a novel particle), and the interpretation of the novel particle as the Higgs boson. Second, we look at the case of the Higgs boson to explore three possible ways of understanding an important claim, the so-called *causal interaction claim* (CIC henceforth). CIC has recently been invoked to justify the epistemological priority of ordinary experiments over computer simulations. Our final goal is to show that in the case of the Higgs boson no suitable qualification of CIC licenses the epistemological conclusion that simulations do not count as genuine experiments because they lack causal interaction with the target system.

A wide and narrow notion of simulation can be found in the literature.¹ According to a widely accepted notion, simulations are continuous with computations. The term simulation is interchangeable with machine-aided computation. In climate science, for example, simulations are routinely employed as computational techniques that allow predictions in situations where no experiment can be performed (i.e. to predict the mean surface temperature changes within 0.5 degrees in the next fifty years under a certain emission scenario, to use the example of Steele and Werndl 2013, p. 613). In this context, it is natural to think of simulations as falling within the remit of computations, rather than proper experiments. In this first wide sense, simulations are synonymous with computations. Leaving here aside the thorny and somehow orthogonal issue of what a computation is, and what might count as a computation (e.g. does the human mind perform computations?), one can distinguish a second notion of simulation.

According to this second notion, simulations are representations. As we see it, this second notion—albeit conceptually and logically distinct—is downstream of the wide notion of simulation as computations. For computations have after all a *representational function*. Computations performed by a computer can *represent* (e.g. a computer game representing real life situations) as much as computations performed by the human mind can *represent* (e.g. mental representations). This second notion of simulations *qua* representations has been recently brought to the fore by Parker (2009, pp. 486–7), who defines computer simulations as “a time-order sequence of states that serves as a representation of some other time-ordered sequence of states”. Parker does not present this second notion of simulation as downstream of the first wide notion of simulation-*qua*-computations. But what both simulations-*qua*-computations and simulations-*qua*-representations share is their contrast with ordinary experiments, which Parker defines as “an investigative activity that involves intervening on a system”. On Parker’s definition, simulating is a way of *representing* a target system, while experimenting is a way of *intervening* on a target system. Thus, on a wide (twofold in our view) notion of simulation (i.e. *qua* computation, or *qua* representation) simulations are at a distance from experimenting.

But there are other domains of inquiry where a narrow notion of simulation has found its way. According to this narrow notion, simulations are in fact continuous with experimenting. This third narrow notion of simulations *qua* experimental activity—and even *qua* complementary experimental practices *in their own right*—has proved helpful in areas where the traditional boundaries between experimenting and theorizing (or between data production and data interpretation) is typically blurred. High-energy physics is an example. The very nature of experimenting in high-energy physics challenges traditional philosophical notions of experimenting (as a form of intervening) *versus* theorizing (as a form of representing). It is not just manipulation or intervention that prove inadequate to describe what goes on in these experimental situations. But theorizing itself enters bottom-up in the very possibility of designing experiments and models of data capable of extracting meaningful signals from the ‘background’ of events.

Consider a typical scattering experiment in high-energy physics, whereby a beam of particles (say, electrons or protons) is smashed against a target (say, electrons, protons or others). Smashing notwithstanding, the core and focus of the experimental activity does not reside in the collision itself, but in analyzing its decay products. Novel discoveries happen when scientists compare the ‘observed’ data (produced by the collision) with the ‘expected’ background (predicted by the relevant model for the scattering in question), and spotting eventual unexpected phenomena (in the form of a non-foreseeable plot of

¹ We thank an anonymous reviewer for drawing our attention to this distinction and for forcing us to spell out the following threefold notions.

data, for example). Interpreting the non-foreseeable plot of data points as evidence for a new particle requires, in turn, a model that can interpret the spread of the plot, its height, and so forth, as evidence for a kind of particle with a certain mass, average life-time, and decay products, compatible with the expected background.

At no stage in this complex chain of events, is there a clear-cut division between intervening *versus* representing. And even at the simple experimental level of particle phenomenology, understanding the nature of the collision, its decay products, and being able to identify new phenomena involves systematics, i.e. the use of computer-aided techniques and theoretical assumptions to control background noise, potential sources of errors, and model data to extract meaningful signals from thousands of events resulting from the collision. It is in this context that the aforementioned narrow notion of simulation has emerged as a complementary experimental practice in its own right, continuous with the computer-aided modeling techniques that are such an integral part of the experimental landscape of high energy physics. In the rest of this paper, we concentrate our attention to this third narrow notion of simulation-qua-experimental activity, and to some of the pressing epistemological questions it poses.

The distinction between a wide and a narrow notion of simulation can somehow be found in Winsberg (2009), who has distinguished between what he calls simulation_R and simulation_A. While simulation_R is co-extensive with Parker's definition of simulation as a kind of representation, simulation_A refers to computer simulations as "a kind of activity on a methodological par with, but different from, ordinary experimentation. (...) The contrast class for simulation_A is ordinary experiment; there are ordinary experiments, on the one hand, and there are computer simulations and analog simulations, on the other" (ibid., p. 583). So, once more, we should ask what distinguishes simulations from ordinary experiments. Winsberg (2010, p. 71) defends the thesis of the epistemological priority of experiments over simulations on the ground that the amount of knowledge needed for model-building relevant to simulation depends—to a large degree—on experiments and observation.² So understood, the epistemological priority thesis asserts that the reliability of model-building methods used in computer simulations crucially depends on our experimental history. Good experimental knowledge is required to build reliable computer simulations. As such, experiments come first over simulations. Let us call this first way of thinking about the epistemological priority thesis EPT₁. EPT₁ is a claim about the reliability of our scientific knowledge, and its ultimate experimental foundations. Good simulations require reliable scientific knowledge, and reliable scientific knowledge rests on solid experimental grounds. An alternative, more poignant way of expressing EPT₁ would be to say that simulations are not the product of a priori knowledge. As such EPT₁ expresses a view that would be hard to deny, a view that we share, fully endorse, and will not discuss any further in the rest of this paper.

But there is another way of thinking about the epistemological priority thesis, which has also attracted attention in the recent literature, and which hinges instead on the alleged causality and materiality of ordinary experiments. Let us call it EPT₂. It is to this second way of thinking about the epistemological priority thesis that we want to focus our attention on in this paper. EPT₂ has appealed to the "materiality" of ordinary experiments as an argument for the epistemic priority of experiments over computer simulations. For example, Guala (2005, pp. 214-5) has argued that while in ordinary

² "for epistemic agents like us, experiments are epistemologically prior to simulations. In both simulations and experiments, you need to know something to learn something. But the knowledge you need in a simulation is always quite abstract and sophisticated, and it usually depends on things you have learned from a long history of experiment and observation. That is because we do not commit ourselves to the reliability of model-building principles unless they have been tested against experiments and observations" Winsberg 2010, p. 71.

experiments we encounter the same “material causes” which are at work in a target system; this is not the case with computer simulations, where the relationship between the simulation and the target system is purely formal and abstract. Along similar lines, Morgan (2003, p. 217) has stressed the non-materiality of computer simulations as an argument for the epistemic power of ordinary experiments over simulations. While experiments are made of the same stuff of target systems, thereby allowing scientists to draw justified inferences from the experimental results back onto the target system under investigation. This is not the case with computer simulations. Hence, Morgan’s conclusion (2005, p. 326) that “ontological equivalence provides epistemological power”. More recently, Giere (2009) too has made a similar point against Morrison’s (2009) take on computer simulations as a kind of experimenting and measuring. In Giere’s words (*ibid.*, p. 61): “the epistemological payoff of a traditional experiment, because of the causal connection with the target system, is greater (or less) confidence in the fit between a model and a target system. A computer experiment, which does not go beyond the simulation system, has no such payoff”. Morrison’s argument (2009; and forthcoming ch. 6 and 7) for computer simulations being epistemically on a par with ordinary experiments is based on the same function that models play in ordinary experiments and in simulations *qua* measuring instruments (in the interpretation of measurement outputs). Against Morrison, Giere has raised two distinct points: the first against the premise of Morrison’s argument; the second against her conclusion.

Against the premise of Morrison’s argument, Giere argues that it is not the case that models function as measuring instruments. For example, in the classical experiment with the pendulum to measure the Earth’s gravitational constant, it is not the model of the pendulum that acts as a measuring device (although it does enter into the abstract representation of how the pendulum works and what correcting factors are needed for the exact measurement of the gravitational constant). Instead, the real measuring instrument is the pendulum itself “which interacts causally with the Earth’s gravitational field. The models used to correct the measured period of the pendulum are abstract objects which do not interact causally with anything” (Giere 2009, p. 62).

Against the conclusion of Morrison’s argument, Giere goes further to suggest that not only do we have no reasons for treating computer simulations on a par with ordinary experiments, but that doing so would in fact be “a bad idea”:

It is central to the notions of experiment and measurement, as traditionally understood, that it involves interaction (however indirect) with a target system and, moreover, that the result of this interaction is an appraisal of the representational adequacy of the models involved. Simulation experiments are lacking this central feature of experimentation. Calling computer experiments, “experiments”, thus suggests that running a simulation somehow provides a basis for evaluating the representational adequacy of the simulation models involved when it does not. This is dangerous in that it can mislead consumers of simulation results (and maybe even some practitioners) into thinking that their simulation models are epistemologically better founded than they in fact are (*ibid.*).

Morrison (forthcoming, ch. 6) has replied to Giere’s argument in her latest extensive treatment of computer simulations, where she argues that even in traditional experiments, the epistemological work is not necessarily done by any causal connection to the target, but instead by modeling assumptions playing an “integral role in what we take to be the causal information gleaned from experiment”. Morrison illustrates her arguments with the case study of the Higgs boson, as a paradigmatic example of how computer simulation “provides the foundation for the entire experiment. To put the

point in a slightly more perspicuous way, simulation knowledge is what tells us where to look for a Higgs event, that a Higgs event has occurred, and that we can trust the overall capability of the collider itself. In that sense the mass measurement associated with the discovery is logically and causally dependent on simulation”.

Thus, EPT_2 appeals to the materiality of the interaction at work in ordinary experiments to advance three distinct, yet related, claims about experiments being epistemically prior to simulations:

- A. *The causal interaction claim* (henceforth, abbreviated as CIC): computer simulations lack a distinctive feature of traditional experiments, namely the ability to causally interact (directly or indirectly) with the target system (Giere’s first argument).
- B. *The representational adequacy claim*: because computer simulations lack (direct or indirect) causal interaction with the target system, running a simulation does not warrant the representational adequacy of the simulation model itself (Giere’s second argument).
- C. *The downward path inferential claim*: because experiments allow us to intervene and manipulate on the same material causes of the relevant target system, experiments allow us to make justifiable inferences about the target system in a more robust way than simulations can do (Guala and Morgan’s arguments).

Claim A. implies claims B. and C. (not the other way around). It is our confidence in the *causal interaction claim* about ordinary experiments that justifies the *representational adequacy claim*, and the *downward path inferential claim*. Thus, the overall burden of the proof for EPT_2 lies with claim A. itself. The *causal interaction claim* needs be qualified to bear the weight required for assessing the alleged epistemological priority of ordinary experiments over simulations.

In the rest of this paper, we review and assess claim A. by following Morrison’s strategy and looking at computer simulations used for the detection of the Higgs boson. In Section 2, we identify and review three possible ways of understanding the *causal interaction claim* (CIC) in ordinary experiments. In Section 3, we describe the three-stage process of computer simulations in HEP, and we illustrate in some detail the use of simulations in two key aspects of the discovery of the Higgs boson. While we side unequivocally with Morrison in defending simulations as epistemically on a par with experiments in the context of the discovery of the Higgs boson, we also add some qualified provisos about the limits of computer simulations in the discovery. Thus, we won’t be arguing for Morrison’s stronger conclusion that the discovery of the Higgs boson is “logically and causally dependent on simulation”. However, we will indirectly defend Morrison’s argument against Giere by showing how there is no distinct sense in which the causal interaction claim can be deployed to conclude that simulations do not count as experimental practices in their own right in the discovery of the Higgs boson. We return to the *causal interaction claim* in the final Section 4, where we conclude that no suitable qualification of CIC (as outlined in Section 2) can be found to license the epistemological conclusion that simulations do not count as genuine experiments.

2. Getting clear about the *causal interaction claim*

There is something intuitively right about the *prima facie* epistemic dissimilarities that many philosophers have noted between simulations and experiments. While simulations seem to enjoy the status of representations, experiments are investigative activities involving interventions and manipulations. The idea of manipulating and causally intervening on a target system seems key to ordinary experiments, but not to computer simulations. Yet trying to spell out exactly how this causal feature—distinctive of ordinary experiments—works out in practice proves rather elusive. In what follows, we take up this challenge, by trying to identify three possible ways of thinking about the *causal interaction claim* (CIC). How should we understand the claim that ordinary experiments involve (direct or indirect) causal interactions with the target system?

The most intuitive example concerns experiments that *measure* physical quantities. At the simplest level, experiments designed to *measure* some physical quantity involve *direct* causal interaction by comparing the target system with a unit of measure taken as the standard or canon for measuring that particular physical quantity. That is how we measure the length of objects with meter sticks; or, the passage of time with atomic clocks; the weight of objects using the kilogram. In all these cases, we *measure* physical quantities by *directly* interacting with the relevant target systems via the relevant units of measure: e.g., how many times the meter stick fits into the length of the object we are measuring; how many seconds—ticked by an atomic clock—fill the time interval we are interested in measuring; how many 1-kg weights should be placed on a scale to balance the weight of the relevant object we intend to measure. *Direct causal interaction* so understood belongs to the realm of metrologists, craftsmen, and merchants. Key is our ability to *compare* the target system with the relevant unit of measure *by juxtaposing them* (leaving here aside problems caused by the SI *Grand Kilo* shrinking over time; meter rods subject to length contraction in relativity; and the all-pervasive difficulty of finding a universally accurate unit of measure that historians and philosophers of metrology have abundantly described).³

More broadly, direct causal interaction is also core to engineering and classical mechanics: from Archimedes' lever, to Galileo's inclined planes, Foucault's pendulum (in Giere's reading of it), and Stevin's hydrostatics, among others. We *measure* motions and forces by direct causal interaction with, and manipulation of physical objects that instantiate them. Obviously, these are just two examples, the most intuitive ones, of how measuring involves direct causal interaction with the target system. *Measuring* may in fact involve different *degrees of causal interaction* that we as agents entertain with the target system. In decreasing order of strength, we can identify three possible (and certainly not exhaustive) degrees of causal interaction at work in different kinds of experimental situations, all designed to *measure* physical quantities. These degrees capture three possible modes through which we—qua *epistemic agents*—advance knowledge claims about the world by entering into three different kinds of causal interaction with it:

- (CIC₁) Experiments involve *direct* causal interactions with the target system when a physical quantity is *calibrated* by direct comparison with observed data. Calibration serves two distinct purposes. Not only does it refine and *tune* the value of a physical quantity to match the observed data. But, in so doing, it also tests the reliability of the instrument producing those observed data, in the light of background knowledge we may have about how the data should look like.

³ On the SI Grand Kilo, see Riordan (2013); on the measurement of time with atomic clocks, see Tal (forthcoming); and, on the more general problem with measurements and standards of accuracy, see Teller (2011), (2013).

(CIC₂) Experiments involve *quasi-direct* causal interactions with the target system when the experimental apparatus is designed to *track* how a physical quantity may interact with another, suitably chosen. In these situations, the causal interaction between, say, physical quantity x and physical quantity y , is what allows us to determine x . In other words, x manifests itself only via causal interaction with y . Thus, although the direct causal interaction is between x and y in nature, we come to know about physical quantity x via a *quasi-direct* causal interaction, i.e. by tracking how x *causally behaves* with another quantity y .

(CIC₃) Experiments involve *indirect* causal interactions with the target system when we *infer* an entity against relevant experimental background. These are the experimental situations where scientists *causally infer* the existence of an entity as the best explanation for novel signals with respect to a well-understood background. Understanding the experimental background and *gaining control of it* is then pivotal for our ability to make *causal inferences* about new entities.

Let us briefly review each case, by starting with some general observations about (CIC₁). As Allan Franklin (1997, p. 32, ft. 2) has pointed out, there is an important difference between *measuring* and *calibrating*: the result of a measurement is not usually known in advance, while it is in calibration, where the validity of experimental results is assessed. Thus, calibration illustrates the two-way-street nature of CIC₁. In ordinary experiments, we do not just use an instrument to *tune* a physical quantity by comparing it with observed data; but we also assess the validity of our experimental results (and ultimately, the reliability of our instrument) by comparing the experimental results with known phenomena. In Franklin's words: "If your spectrometer reproduces the known Balmer series in hydrogen, you have reason to believe that it is a reliable instrument. If it fails to do so, then it is not an adequate spectrometer" (ibid., p. 32). Scientists may calibrate the spectral lines of the hydrogen series (say, their respective frequencies) by comparing the observed data with the known Balmer series. In doing so, they assess also the reliability of the spectrometer that has produced those observed data. Thus, calibration as a paradigmatic instance of CIC₁ delivers a pervasive, and historically influential image of experiments as involving *direct causal interactions* on the world. This image tacitly motivates the widespread feeling about the epistemological priority of experiments over computer simulations in giving us access to the physical world and make justified claims about it (recall claim C., above). Indeed, under this first reading of the causal interaction claim, one might argue that computer simulations are not epistemically on a par with ordinary experiments because, strictly speaking, a simulation produces no 'observed' data but only simulated data. Moreover, simulations presuppose knowledge of the same phenomenon they simulate; hence Giere's remark that running computer simulations is not a way of checking the representational adequacy of the simulation itself (claim B. above).

Yet other kinds of experiments involve a somewhat less direct form of causal interaction with the target system. Causal interaction with the target system comes in degrees. Our second class of experiments involves what might be called a *quasi-direct* causal interaction with the target system. Consider the Fizeau–Foucault experiments in 1849–50 to measure the speed of light in various media (air, water, glass). Foucault's experimental set-up consisted of a spinning mirror device, powered by a turbine, whereby a beam of light could be split and sent across a set of concave and spinning mirrors so that the time-interval of its path through the instrument in various media could be measured when the incoming beam (and its slight shift due to a shift of the spinning mirrors) was observed through a microscope. On this basis, Foucault concluded that the speed of light in vacuum was close to 298,000km/sec, but much less in other

media such as water. Or, consider Joule's paddle-wheel experiments in the 1840s, through which the interchangeability of mechanical work and thermal energy was first established, and the *joule* introduced as the unit of work or energy. The experiment, consisting of a cylinder filled with water and a thermometer, and a system of pulleys activating paddle-wheels inside the cylinder, was designed to *track* how the amount of mechanical work spent to activate the pulleys and the paddle-wheels converted into thermal energy (measured by the increase of temperature in the water). In both these examples, the experiments were designed to measure a physical quantity—such as the speed of light, or mechanical work—by *tracking* (via the microscope and the spinning mirrors; or via the paddle-wheel system and the thermometer) causal interactions between the physical quantity under investigation and another suitable physical quantity (i.e., resistance in media such as water, in the first example; and thermal energy, in the second example).

Thus, this second class of experiments involves what might be called a *quasi-direct causal interaction* with the target system. The experimental apparatus is not designed to interact directly with the relevant quantity (as in calibration), but instead to *track* causal interactions between a physical quantity x and another physical quantity y . Causal interactions are important in this kind of experimental situations because the relevant physical quantity x can only be determined via its causal interaction with another relevant quantity y . The epistemic job is done by the causal interaction between the two physical quantities in nature, and experiments of this kind are designed to elicit the manifested causal behavior of the relevant quantity. Thus, under this second reading of the causal interaction claim (CIC₂), one might argue that computer simulations are not epistemically on a par with ordinary experiments, because a simulation operates in an artificial, virtual environment, where data input and software packages determine the sequence of simulated events, and as such they do not *track* any causal interaction in nature. One may try to respond that the simulated sequence of events *tracks* the real sequence of events in nature. But such a response would beg the question by leaving unexplained how a virtual sequence of events can possibly track a real one, if not by sheer methodological assumption.

There is, finally, a third class of experiments we want to consider here. Coming closer home, these are the experiments that characterize most of twentieth-century particle physics, and involve only *indirect* causal interactions with the target system by *causally inferring* an entity as the best explanation for novel signals with respect to a well-understood background. Take, as an example, Anderson's discovery (Anderson 1933) of the positron in the photographs of cosmic ray tracks in a Wilson cloud chamber. This experiment was the forerunner of an entire tradition in experimental particle physics, where new entities predicted by a theory (in this case, Dirac's hole theory) are experimentally detected or *inferred* as the best explanation for novel 'signals' or 'events' with respect to a background of well-known and well-understood signals and events. The trail of ions in Anderson's photographic plate, which revealed the existence of the positron, had the same curvature of electrons (given their mass-to-charge ratio and known causal behavior in a magnetic field), but opposite direction (indicating positive, rather than negative electric charge). In this case, the experiment causally inferred a new particle (with a certain mass and electric charge) as the best explanation for the novel signal with respect to a well-understood background of signals, which Anderson had control of (e.g. by adjusting the strength of the magnetic field, for example). In other words, Anderson could *infer* the positron because he knew that electrons causally behaved in a certain way in a Wilson cloud chamber. Thus, in CIC₃ causal knowledge of the background is paramount. *Indirect causal interaction* here means that the experiment allows scientists to *causally infer* an entity against a well-understood background, which

they can control. Once again, simulations seem to fall short of CIC_3 , because simulations do not allow scientists to gain control of the experimental background, in the sense of being able to intervene and manipulate the experimental background (e.g. change the intensity of the magnetic field) for the purpose of *causally inferring* the presence of a new entity.

A further, bigger worry needs also be addressed. One may wonder the extent to which these three senses of CIC really capture the kind of causal claim that Giere, Guala and Morgan have been appealing to. For it would seem that even under CIC_1 , which is the closest among the three to the spirit of Giere’s causal interaction claim, the kind of causal interaction at play is in fact a formal causal dependency between the observed data and the expected data. A critic may retort that this begs the question against Giere–Guala–Morgan “materiality” requirement, whereby it is the actual measuring instrument (e.g. the pendulum) that physically interacts with the target system (e.g. the Earth’s gravitational field). Comparing, adjusting, tweaking experimental data with respect to expected data (as in any calibration process captured by CIC_1) does not cut any ice against the material, causal-physical interaction that matters to the Giere–Guala–Morgan argument. And a fortiori, if CIC_1 does not do, neither can CIC_2 or CIC_3 , which are weaker versions of CIC. Or, so the argument goes.⁴ At the heart of this objection lies once again a deeply entrenched view about *measuring* as an *experimental activity* involving intervening and manipulating on a target system. On this view, it is the materiality of the experimental device that causally-physically interacts with the materiality of the target system to deliver the measurement outcome. It is the mass of the pendulum that interacts with the mass of the Earth; or it is the electromagnetic waves coming from Supernovae Ia that interact with the lenses of the Blanco Telescope in Chile; or the friction caused by the paddle-wheels interacting with water molecules that ultimately bear the onus of the causal-physical interaction claim. Calibrating, tracking, inferring—the kinds of causal interactions captured by the three aforementioned senses of CIC—seem to miss this key distinctive “material” feature of measuring and experimenting. And, the argument goes, it is this distinctive “material” feature that ultimately does the job, when it comes to measuring and experimenting.

While we accept that the material feature—so strictly understood—is indeed absent in $CIC_{1-2,3}$, we question the conclusion that it is in fact this material feature that bears the burden of the job when it comes to measuring and experimenting. This conclusion follows from a further premise in Giere’s argument concerning the dichotomous division between intervening and representing, whereby it is the materiality of the experimental device (e.g., the mass of the pendulum) that physically interacts with the Earth’s mass, whereas the models used to correct the measured period of the pendulum are abstract objects, which do not interact causally with anything and as such do not do any real job in the measuring process. But—we argue—this dichotomy between the materiality of the experimental device vs. the models / laws involved in the construction and functioning of the device itself is questionable; and accepting this dichotomy would simply beg the question for Giere’s argument.

The dichotomy is questionable for two main reasons. First, materiality (so understood) is not a *sufficient condition* for experimenting and measuring. My hand exercises a causal-physical interaction when I push the cup of coffee on the table with a given force; but without the use of an appropriate model (featuring Newton’s second law), such causal-physical interaction *by itself* does not amount to any measurement of any physical quantity. Even Franklin’s legendary electric kite experiment—perhaps one of the most flamboyant examples of causal-physical interaction in the history of physics—did

⁴ We thank an anonymous reviewer for constructive criticism on this point.

not actually measure electric current but simply concluded that lightning was made of an electric fluid passing from the cloud to the rod. To measure electrical current, ammeters rely on Ohm's law, which says that the current passing through a conductor between two points is proportional to the potential and inversely proportional to the resistance between the two points. Causal-physical interactions can be found where one wishes to find them in nature, without them amounting to anything like measuring a physical quantity. Train wheels rotating along tracks; kites flying in the air; stones rolling down a hill are just three examples of how nature is teeming with causal-physical activities. What makes some of these causal-physical interactions worthy candidates for measurements (e.g. the pendulum for the Earth's gravitation field; wind tunnel for air resistance; electric current passing through a conductor), while most of them (e.g. rolling stones, flying kites, etc.) are not? Obviously, something needs to be said here by defenders of the "materiality" condition, as it is simply not the case that materiality *per se* can bear the weight that Giere's argument seems to place on it, without the proper support from models and laws that select relevant aspects of the causal-physical interaction, model it in a law-like way and make it conform to general explanatory patterns, which in turn function as measuring procedures (i.e. it is because of the pendulum conforming to a suitable instantiation of Newton's second law that it can be made to function as a measuring device for the Earth's gravitational field; it is because of electricity obeying Ohm's law that ammeters can be built to measure the electric current; and so on).

Second, materiality (so understood) is not a *necessary condition* either for measuring. Consider Marc Lange's example of a double pendulum (Lange 2013, p. 502-3), where the four equilibrium configurations can be explained not by appealing to the causally-physical interaction among the forces acting on the two bobs (with masses m and M), but (alternatively) by appealing to the topology of the configuration space (i.e. points on a toroidal surface and the effects of possible distortions of the torus). Lange refers to this as an example of non-causal, distinctively mathematical explanation where the four equilibrium configurations for the double pendulum can be explained not in terms of the causal forces acting on the bobs, but by appealing to Newton's second law as "the framework within which any force must act", where Newton's second law is described as being "modally more necessary even than ordinary causal laws...or constitutive of the physical task or arrangement at issue" (ibid., p. 506). Distinctively mathematical explanations such as this one show once more how it is possible to provide an account of the pendulum and its configuration points that abstracts from the network of causal relations in the world. More to the point, it shows how—pace Giere's argument—there is nothing special or privileged about the materiality condition in Giere's own example of the pendulum that license the conclusion that it is indeed this material feature that bears the burden when it comes to measuring and experimenting. On the contrary, the possibility of offering a non-causal, distinctively mathematical explanation of the pendulum and its functioning shows how crucial models (and the laws of nature they instantiate) are in the actual process of experimenting and measuring. While it is possible to abstract from the network of causal forces acting on the pendulum, no explanation of the pendulum and its working would be possible without resorting to Newton's second law (with its constitutive, or modally necessary status, to use Lange's words). And what Newton's second law does is to express a framework for capturing causal dependencies in nature, a framework that can be adapted to capture a variety of different phenomena and specific forces in nature.

Going then back to our original question, are critics correct in claiming that computer simulations lack the ability, distinctive of ordinary experiments, to *causally interact* with the target system? So far we have identified three possible readings of the causal interaction claim (CIC) and the respective ways in which each seems to license the

thesis of the epistemological priority of experiments over simulations (EPT_2). We have also further clarified how the kind of causal dependency captured by $CIC_{1-2,3}$ (in terms of calibrating, tracking and inferring) encompasses three central notions at work in measuring and experimenting across a variety of situations. And we have explained why the materiality condition *per se* is neither a necessary nor a sufficient condition for experimenting and measuring, and it must be supplemented by proper models and their associated laws of nature, without which no meaningful, law-like, explanatory pattern can be identified in the variety of causal-physical interactions existing in nature.

In the next Section, we turn our attention to the paradigmatic use of computer simulations in the detection of the Higgs boson. We briefly present some salient features of how computer simulations are used in this context: (a.) for *background determination*, which is key to the experimental detection of the particle; and (b.) for the *interpretation* of the particle thus detected as the Higgs boson. In Section 4, we return to the causal interaction claim and argue that with respect to its formulation along the lines of $CIC_{1-2,3}$, there are no distinctive differences between experiments and simulations as used in the detection of the Higgs boson. More precisely, our goal is to show that computer simulations used in ATLAS meet or satisfy all three characterizations of the causal interaction claim so that the epistemological priority of experiments over simulations on the basis of the causal interaction claim (i.e. EPT_2) does not go through in this paradigmatic case.

3. Simulating and experimenting: the case of the Higgs boson

In analyses such as the discovery of the Higgs boson, high-energy particle physics makes extensive use of simulated data. For example, the ATLAS collaboration at the Large Hadron Collider (LHC) has simulated more than 10 billion collision events using a worldwide network of more than 100,000 computing cores. The typical stages involved are illustrated below in Fig. A and the key aspects relevant to the discussion in this paper are also described below (for more detail see ATLAS (2010)).

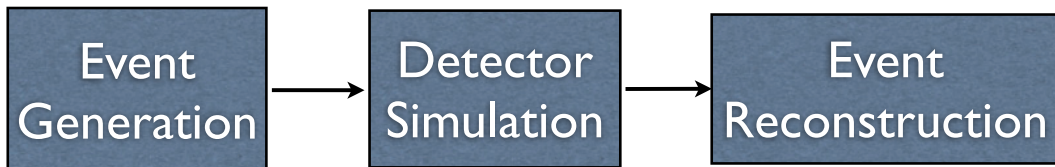


Figure A. The three-stage simulation process in high-energy particle physics

Event generation uses theoretical models to predict the quantities of different types of particles that will be produced in the high-energy collision (such as the collision of two protons at the LHC); their kinematic properties (energy, momentum etc.) as well as the particles they decay into. We provide specific examples of the models used later in our discussion of the Higgs analysis. These theoretical models provide probability distributions such as the number of times a particle of a given energy would be expected

at the LHC. From these probability distributions, a statistically representative sample of such particles can be drawn using *Monte Carlo* methods. The output of this process is a number of *events* each containing a list of particles and their properties.

Detector Simulation takes the list of particles resulting from event generation and uses a very precise description of the particle physics experiment to determine how the experimental set-up would react to the passage of those particles. For the ATLAS detector this is done using the Geant4 software package. This includes a description of all the material, sensors, and electronics and how they would react to different types of particles, on the basis of past experimental measurements. The output of this process is a series of detector signals or readings close to those obtained when such particles traverse the physical ATLAS detector.

In this stage there is also a process that adds additional hits to the previous output that would be present in a real LHC collision because of other collisions occurring at the same time. For the ATLAS experiment these hits also come from simulation.

Event reconstruction runs on these simulated detector signals applying the *same* software used on those obtained from the real detector, with the aim of reconstructing what particles were initially present. The output of this process is the same kind of measurements, such as the momentum of the reconstructed particles, available for the experimental data collected at the LHC. Using the same software ensures that the same experimental constraints apply in reconstructing simulated and real collisions. For example, reconstructing a particle track on the real physical detector requires joining up a series of detector hits. This process is ambiguous and may give the wrong answer because of the presence of many particles and hits. By using the same software for simulated “hits”, the same ambiguity, and potentially incorrect determination of initial particle occurs both in the simulation and real collision data. The reason for this is to produce a simulation that is as imperfect as the experimental data themselves: if the experiment would get the wrong track, the simulation should do so as well. The goal is to create the best possible representation of experimental reality (i.e. of what was physically detected). One reason to proceed in this way is to ultimately use the simulation to “remove” the ambiguity from the experimental data. For example, if it is seen in simulation that the experimental software can find only 50% of the real tracks, then the real, experimentally detected number of tracks should be doubled to the amount that would have been actually produced.

This general three-stage process is used for simulations across a wide-range of experiments in high-energy particle physics, and provides a template for thinking about how simulations enter in experimental designs and data analyses. We now turn our attention to the specific example of the Higgs boson discovery at the ATLAS experiment. This case study nicely illustrates the above three-stage process at work, and we focus on two specific examples of how computer simulation is used to detect the new particle.

3.1 Use of Simulation in the detection of the Higgs boson

In what follows, we concentrate on two important uses of simulation in the recent detection of the Higgs boson at ATLAS:

1. **Background determination:** simulated data is used to determine the nature and amount of other, already known particles that would have occurred in the collision so that a novel particle can be identified as a new signal against this background of known signals.

2. **Interpretation:** simulated data is produced for the Higgs boson as predicted by the Standard Model and alternative hypotheses. In this way, it can be determined, by comparison with the new particle identified in stage (1.), whether the new particle is indeed a Higgs boson, and what its properties are.

3.1.1 Background determination: finding the Higgs in the 4-lepton channel

The discovery of a new particle comes from the observation of an excess of events after all other known physical processes (“background”) have been taken into account. In the search for the Higgs boson on the ATLAS experiment this was demonstrated initially mainly in the “channels” where the Higgs boson decays to a final-state of two photons or four leptons.⁵ The usual variable in which this excess is illustrated is in the invariant mass of these final-state particles, which corresponds to the mass of the hypothesized Higgs boson. The distributions for the Higgs boson decay to four leptons (m_{4l}), and to two photons ($m_{\gamma\gamma}$), from the latest public ATLAS analyses, with the full data collected in 2011–2012, are shown in Figures B and C, respectively.

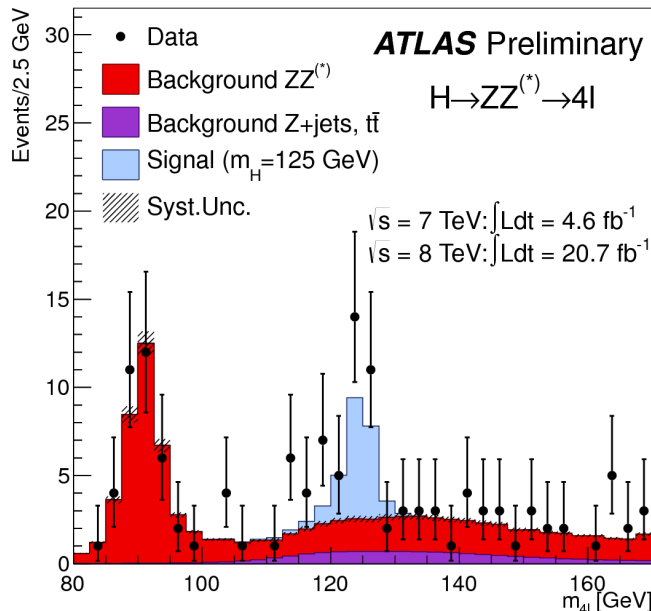


Figure B: The distribution of the invariant mass of the Higgs decay to four leptons in the ATLAS detector (ATLAS 2013b). The black points are a count of the total events seen in each invariant-mass range in the actual physical experiment. The red and purple histograms are the expected number of background events determined from simulation as outlined in the description here below. The light blue histogram shows the expected distribution for a Higgs boson of a mass of 125 GeV determined from simulated data.

⁵ On the ATLAS experiment the Higgs boson is searched for in "channels", defined by the Higgs boson's decay to other particles such a two photons, two Z-bosons (which further decay to four leptons), two W-bosons, two b-quarks, or two tau-leptons. Of these, "two photons" and "four leptons" are the channels that most contributed to the first discovery and are now used to measure the mass of the Higgs boson, on which we focus here.

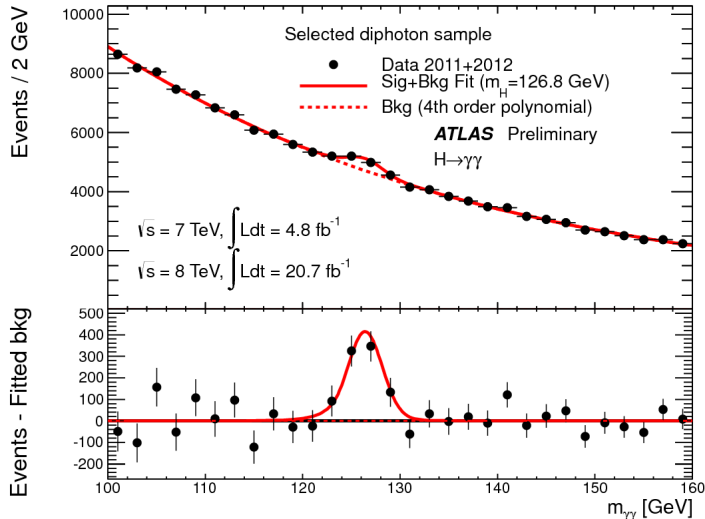


Figure C: The invariant mass of the Higgs to two photons decay in the ATLAS detector (ATLAS 2013a). The black points are a count of the total events in each invariant mass range seen in the physical experiment. The red lines are statistical fits to those points, where the dotted one is a fit to the points that are not near the observed bump and are therefore background. The bottom plot shows the value obtained when the value of the dotted red line is subtracted from the value of the black data point for each x-value.

In these plots the black points that are labeled “data” are frequency counts of the data actually observed in the ATLAS experiment. So the y-axis is a count of the total number of events seen in the experiment for each mass value shown on the x-axis. The red and purple histograms in figure B labeled “background” show the frequency of data that would have been expected to be produced from other known physical processes that are not Higgs boson production. A difference can be clearly seen between the “data” and the expected “background”. It is this difference that establishes that a previously unknown particle has been produced. Observation of the new particle is sometimes highlighted by a subtraction of this background from the observed data (as is done in the bottom plot in figure C), where the values of the background histograms are subtracted from the data points so that the number of events shown corresponds to only the number of Higgs candidate events. To actually establish the existence of a new particle, physicists calculate the probability that the background alone, without a new particle, could cause the observed data—the probability for this in the ATLAS data for the 4-lepton channel is now less than 1 / billion (ATLAS 2013b: 17-18). Evaluation of this probability relies on knowing the number of background events as much as it requires the number of observed data events. Therefore the experimental observation of a new particle depends crucially on the background determination as well as on the observed data.

For the four-lepton case the background is determined from simulation. The simulated events are produced following the three-stage process illustrated in figure A. The Monte-Carlo event generation stage uses theoretical models for the non-Higgs production of Z bosons from either quark-antiquark annihilation and gluon-gluon production, either alone or with additional quarks. These are then passed through a simulation of the whole detector and the masses of the particles reconstructed in the same way as for the observed data, so they can be plotted as the red background in figure B. In order to understand the background well (and reduce statistical errors), many thousands of events may be produced. However, this frequency is not the amount of background (i.e. height of the red and purple histograms) that should be used in figure B.

Instead what should feature in Fig. B is the amount that one would expect in the real experimental sample collected. Therefore, the histogram needs to be *normalized*. This can be done using a theoretical model such as that used in the event generation part of the simulation, which gives a measure of the cross-section, or rate of production at the LHC: this is how the red background in Fig. B is normalized.

However, for the purple histogram in figure B (i.e. the background processes where a Z boson is produced with additional quarks, (Z+jets or ll+jets), or where top quarks are produced ($t\bar{t}$)) a notably different procedure is followed. This difference is mentioned in ATLAS (2013b: 8): “the level of the irreducible ZZ(*) background is estimated using MC simulation normalised to the theoretical cross section, while the rate and composition of the reducible ll + jets and $t\bar{t}$ background processes are evaluated with data-driven methods”. *Data-driven* methods are performed by looking at the distribution of other variables, with a slightly different selection of events, so that there could be no Higgs boson but more of these types of background (often termed “control regions”). For example, Figure D shows such a control region for the H-to-4l analysis, focusing on the distribution of the mass of the two highest momentum leptons. This plot contains only background, it should not have any Higgs signal in it, so the simulated background can be compared directly with the observed data. A statistical fit is performed to adjust the amount of the simulated data for each background to match the observed data. The amount that the simulated data needs be adjusted by is taken as a *scale factor* that is also applied to these backgrounds when making the plot in figure B. So each bin of the purple histogram in figure B is multiplied by the scale factor determined from this control region before extracting the Higgs signal.

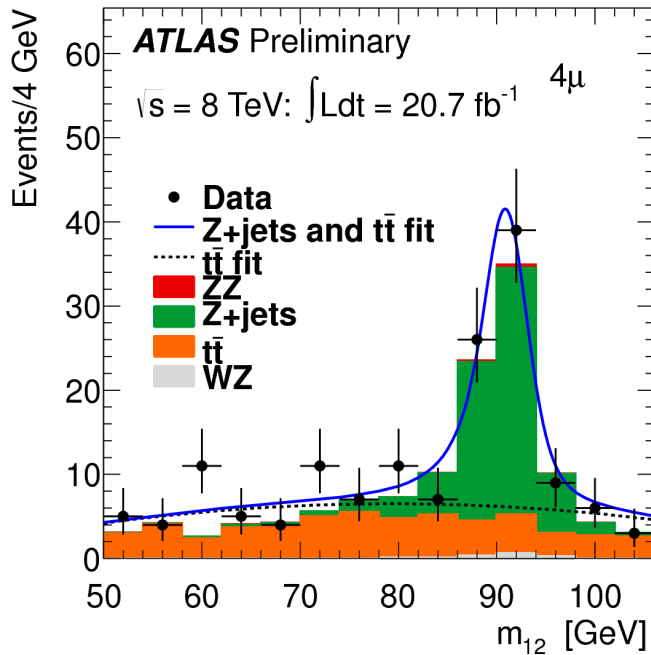


Figure D: Control region of the Higgs to 4-lepton analysis, with a statistical fit to the distribution of the mass of the two highest momentum leptons used to obtain the amount of Z +jets and $t\bar{t}$ background.

Comparison between simulated and experimentally observed data in control regions is a form of *validation* that gives confidence in the simulation. It is particularly required if the underlying physical process being simulated is complex. In some cases, such comparisons are used to *calibrate* or *tune* the simulation, whereby parameters of the underlying physical model used for event generation are changed to ensure a better agreement. For example, the fraction of strange quarks produced in the proton-proton collision is a parameter of the model that can be tuned by comparing distributions of simulated data with that observed for physical quantities such as the number of final state particles containing strange quarks.

However, in this case for the Z +jets and $t\bar{t}$ background, the physical model is not changed. Instead the scale factor derived from comparison with the experiment is applied directly to the final distribution of the simulated data. Therefore, the resulting data used for the background measurement is a hybrid of simulated and experimental data.

Compare and contrast the four-lepton case with Figure C, which shows the background determination for the case when the Higgs boson decays to two photons. Here the background is determined by performing a statistical fit to the experimentally observed distribution. In this channel there are large numbers of observed events and therefore the statistical errors are small. Furthermore the background as shown in figure C has a simple shape where the numbers of events steadily decrease with increasing mass. Because of this feature, it is possible to reliably extrapolate from the number of observed events that do not have a mass around 125 GeV to the amount of the background underneath the Higgs signal “bump” at 125 GeV. This extrapolation is checked with simulated data to ensure that the simulation gives the same distribution. However, by contrast with the 4-lepton case, the background determination in the two-photon decay channel *does not directly require simulated data at all*.

This example therefore shows a case where simulations in particle physics can be replaced by experiments. The approach described above for the 2-photon channel that uses experimental data instead of simulated, could also be done in the 4-lepton case. Instead of using simulated data for the red and purple histograms in Figure B, one could experimentally measure the observed amount of events that do not have a mass around 125 GeV and extrapolate that amount beneath the Higgs signal “bump” at 125 GeV in Figure B.

This then invites the question as to why physicists use simulation instead of using the experimental approach just described for the 4-lepton decay channel. The reason for preferring simulation over experiment is essentially to improve the accuracy of the background determination; or, in other words, to reduce the uncertainty. Reducing uncertainty is very important because, as mentioned above, to actually establish the existence of a new particle, physicists calculate the probability that the background alone, without a new particle, could cause the observed data. This probability depends on the number of background events and signal events but also on the uncertainty associated with those numbers. So the experimentalist must choose the most accurate method for the most significant result—just as one would choose the best calibrated and most precise ruler to perform a length measurement.

In the 4-lepton case, there are much fewer events observed than in the 2-photon case: i.e. the number of events in the histogram in Figure B is far fewer than the number of events in figure C. This is also reflected in the large error bars on the black data points on Figure B, while those in Figure C are so small as to not be visible. Because of that error on each data point, a statistical fit to extrapolate the background under the “bump” in the 4-lepton channel would have a large uncertainty—there are many possible extrapolation curves that would be consistent with the black data points within the large

error bars. Therefore, in this case, it is preferable to use simulations. More simulated data can be generated, simply using more computer time, to reduce the uncertainty in the red and purple histograms, while more experimental data can only be collected with more proton-proton collisions at the LHC. However, the simulated data points have an error bar too, shown as the shaded area labeled ‘Syst. Unc.’ in Figure B. This is the ‘‘Systematic Uncertainty’’ driven by, for example, the level of understanding in the theoretical models underlying the simulation. Therefore, in some cases such as the 2-photon decay mode, an experimental approach may give a smaller overall uncertainty, and so it may be preferable.

To sum up, the degree to which the discovery of a new particle depends on simulation varies considerably. The amount of background, crucial for the discovery of the new particle, can be taken from either experimental measurements, or from simulation. In the 4-lepton case, physicists treat simulated data as equivalent to an experimental measurement precisely because of the uncertainty described above. But the process can also, in principle, not depend directly on simulation, as is the case for the ATLAS analysis of the Higgs decaying into two photons. The background may also be a hybrid of simulated data and experimental observation (as in the case of the $ll+jets$ background to the Higgs to 4-lepton analysis) further illustrating the impossibility of drawing a sharp epistemic distinction between simulations and ordinary experiments. The experimental detection of a new particle depends as much on the background determination, which is often simulated (as in the 4-lepton case), as it does on the observed data points. So simulating data is an integral part of the experimental discovery of the Higgs boson, and scientists’ choice as to whether to use simulations or experiments shows once more the interchangeable role they play in the context of high-energy physics research, pace EPT₂.

3.1.2 Interpretation: Measuring the Higgs

After a new particle had been detected on ATLAS, attention immediately switched to establishing whether this was indeed the Higgs boson predicted within the Standard Model (SM) of particle physics, and to measuring its properties such as its mass, which was not predicted by the theory. To accomplish this task, use was made, once again, of simulated data, following the same three-stage process of figure A. This time, the model used in the event generation stage is that of the Higgs boson decaying to the channel under investigation (for example 4-lepton or 2-photon) with the ‘‘coupling’’ (and therefore rate) predicted by the SM. These models take the mass of the Higgs boson that is not predicted by the SM as a parameter with different possible values, so that many simulated samples with different masses are generated.

These simulated samples are compared with the experimentally observed data again using the distribution of the mass of the four leptons shown in figure B. In that figure, the Higgs simulated data is the light blue histogram labeled ‘‘signal’’. The various samples generated with different masses are compared with the observed experimental data to establish the extent of the agreement. This can be done even by eye, but in practice a statistical fit is performed varying the normalization and the mass of the sample to match the shape (produced by the simulation) to the black data points in figure B. The normalization, or height of the best-fit signal histogram, measures the ‘‘signal strength’’ and can be compared to theoretical predictions for the couplings of the SM Higgs boson. The mass value of the simulated sample that best fits the black points provides the measurement of the mass. For the mass measurement, the expected shape and normalization from simulation are also extrapolated (‘‘morphed’’) to mass values in-

between the generated samples. This extrapolation is performed in order to have a continuous estimation of the shape and normalization for all mass values, rather than only at the discrete values input, without having to generate many additional samples.

It is perhaps interesting to note that again in this case the process could also be achieved without simulation. The mass, and signal strength could be measured directly from the x and y values of the peak of the black points observed in data with some theoretically or experimentally derived correction and uncertainty. So again simulation is a key component of the analysis conducted, and it can be used interchangeably with experimental and theoretical methods in the detection and identification of the Higgs boson.

4. Some concluding remarks. What can we learn from the Higgs case?

What to make of the causal interaction claim in the light of the above discussion? As we saw in Section 1, the epistemological priority of ordinary experiments over computer simulations has—among others—been argued for on the ground that experiments involve causal-physical interactions with the target system (as per EPT_2). In Section 2, we have clarified what might be at stake in the causal interaction claim, and offered three possible readings of it. The discussion of the computer simulations involved in the detection of the Higgs boson in Section 3 puts us in a good position to go back to those readings, and assess whether they may or not apply to this specific case. In the case of the Higgs boson, we have seen how computer simulations go hand-in-hand with experiments in the ordinary sense (i.e. the ATLAS material experiment with particle collisions from the LHC) to deliver the final experimental results. We have, in particular, seen how computer simulations are key to both establishing the existential claim that “there is a new particle”, and to further qualifying that claim as “there is an Higgs boson”.

More to the point, the use of computer simulations in the Higgs case seems to undercut the epistemological priority of experiments over simulations (EPT_2) allegedly delivered by CIC. The computer simulations at work in the Higgs case *satisfy* all the three readings of CIC that we identified in Section 2. Recall the first reading:

(CIC₁) Experiments involve *direct* causal interactions with the target system when a physical quantity is *calibrated* by direct comparison with observed data. Calibration serves two distinct purposes. Not only does it refine and *tune* the value of a physical quantity to match the observed data. But, in so doing, it also tests the reliability of the instrument producing those observed data, in the light of background knowledge we may have about how the data should look like.

Computer simulations in the Higgs case clearly satisfy this first reading of CIC. Indeed, calibration plays a vital role in the whole process of the detection of the Higgs boson (for an extensive treatment, see Morrison, forthcoming, ch. 8). Physical quantities of the underlying physical model used in the event generation of the simulation get calibrated or tuned by comparison with observed data. For example, the fraction of strange quarks produced in proton-proton collisions is a quantity that can be tweaked and tuned to provide the best fit for the data points in a range of distributions of physically observed quantities, such as the number of final state particles containing strange quarks. In other cases, what gets tuned or calibrated is not a physical quantity itself, but instead the scale factor by which we need to adjust the simulated data to match the observed data, without changing the underlying physical model and its quantities.

More to the point, the reliability of the instrument (say, the spectrometer) that calibration licenses in ordinary experiments, translates into the reliability of the simulation that calibration licenses in the Higgs case. Remember the way the scale factor was determined in Fig. D, from a control region for the Higgs decay into $4l$, which focused on the distribution of the mass of the two highest momentum leptons. In that case, the simulated background is compared with observed data, normalized, and calibrated accordingly (before being inputted into the purple histogram in Fig. B). The reliability of the computer simulation (or, to use a more precise terminology, the *validation* of the simulation) used in the production of Fig. D is assessed via comparison between observed data points (and their statistical fit) and what our physical model says about $Z+$ jets and top-antitop background, in an analogous way in which the reliability of a spectrometer is assessed by comparing experimental data about the hydrogen spectrum with the known Balmer series. Calibration as a paradigmatic instance of the *direct causal interaction claim* is a way of establishing the representational adequacy, and hence the validity, of our results in ordinary experiments, no more than in computer simulations in the Higgs case. Thus, there does not seem to be a principled epistemic distinction between CIC_1 in ordinary experiments and in computer simulations in the Higgs case. Calibration plays exactly the same role in sanctioning the reliability of ordinary experimental set-ups and of computer simulations (as per CIC_1).

How about the second reading of CIC?

(CIC_2) Experiments involve *quasi-direct* causal interactions with the target system when the experimental apparatus is designed to *track* how a physical quantity may interact with another, suitably chosen one. In these situations, the causal interaction between, say, physical quantity x and physical quantity y , is what allows us to determine x . In other words, x manifests itself only via causal interaction with y . Thus, although the direct causal interaction is between x and y in nature, **we come to know about physical quantity x via a *quasi-direct* causal interaction with the target system by tracking how x *causally behaves* with y .**

Here the relevant causal interactions are not the ones between the experimental set-up and the target system itself as in CIC_1 , whereby the same reliability of the experimental set-up gets sanctioned. Instead, the relevant causal interactions are those going on *in the world* as it were (say, between the speed of light and the refracting medium in the Fizeau-Foucault experiment; or between thermal energy and mechanical work in Joule’s experiments). The role of the experimental set-up is to reliably *track* those causal interactions. Hence, we qualified CIC_2 as involving a *quasi-direct* causal interaction. Now, like in the Fizeau-Foucault experiment, in the Higgs case too the primary causal interactions are the proton–proton collisions at the LHC. And like in the Fizeau-Foucault case, detecting the relevant physical quantity (be it the speed of light, or the mass of the Higgs boson), involves being able to *track* how our physical quantity interacts with another one. For the Higgs mass, this is done by looking at how the Higgs boson decays into particles, whose total invariant masses and rates of production are *causally dependent* on the (to-be-measured) Higgs mass. These are the relevant causal interactions occurring in nature. ATLAS computer simulations are designed to track those causal interactions by producing various simulated samples with coupling rates predicted by the Standard Model and different mass values, and by “morphing” between them to give a continuous parameterization of the signal shape.

Thus, the simulation *tracks* causal interactions in nature by using a theoretical model (Standard Model) to represent those causal interactions, and by supplementing the missing info in the model (i.e. the Higgs mass value) with a range of simulated samples

until a continuous parameterization of the signal shape can be reached, as a genuine *measurement* of the Higgs mass. In the Fizeau-Foucault case no more than in ATLAS, the overall experimental set-up is designed to *track* causal interactions in nature. The difference is that while in the Fizeau-Foucault experiment a beam splitter, a set of concave mirrors, and a turbine powered spinning mirror was sufficient to track causal interactions between the speed of light and the refracting medium, through which the speed of light could be determined; in ATLAS computer simulations are essential to determine the value of a physical quantity, such as the Higgs mass, by tracking its causal behavior via a variety of simulated samples with different mass values and coupling rates. Although the primary causal interactions between proton–proton collisions and the respective Higgs decay products go on *in the world* (more precisely, inside the LHC in Geneva), ATLAS *tracks* those causal interactions by comparing found experimental data (i.e. events / invariant mass, namely the black dots of Fig. B) with a range of simulated mass samples. Once again, there does not seem to be a principled epistemic distinction between CIC_2 in ordinary experiments and in computer simulations in the Higgs case. The Fizeau-Foucault experiment and ATLAS can be regarded as two extremes of a continuum, whereby computer simulations have replaced turbine-powered spinning mirrors in tracking causal interactions in nature.

We come finally to CIC_3 , which says:

(CIC_3) Experiments involve *indirect* causal interactions with the target system when we *infer* an entity against relevant experimental background. These are the experimental situations where scientists *causally infer* the existence of an entity as the best explanation for novel signals with respect to a well-understood background. Understanding the experimental background and *gaining control of it* is then pivotal for our ability to make *causal inferences* about new entities.

In the positron case, like in the Higgs case, physicists *causally infer* an entity as the best explanation for a novel signal with respect to experimental background. Gaining control of the experimental background is what allows scientists to identify a novel signal in a sea of background noise and other irrelevant signals. The difference once again is that while in Anderson’s case the experimental background consisted of a photographic plate of cosmic rays with electron tracks (whose curvature Anderson could control by adjusting the strength of the magnetic field and knowing how a certain mass-to-charge ratio would causally behave in that field); in ATLAS, the background typically consists of a hybrid of experimental data *and* simulated data, as in the 4l decay channel illustrated in Fig. B. Indeed, most of the experimental work goes towards designing effective ways for determining and controlling the experimental background, against which the hypothesized particle can then be reliably inferred.

As we saw in Section 3, in ATLAS computer simulations prove indispensable to determine the *nature* of the experimental background—e.g. how many different kinds of particles are produced in the collisions and their respective decays (for example ZZ , Z +jets; top–antitop quarks, to mention just those in the decay channel of 4 leptons); the *shape* of the background—e.g., what is the distribution of each kind of particle for the invariant mass of the decay product in GeV; as well as the *normalization* of the background—e.g. the heights of the shapes of different colors in Figure B (the overall number of events/particles of each kind being produced for a given invariant mass). Thus, in this case too, computer simulations at ATLAS are no different from ordinary experiments: indeed they are an integral part of the experiment, and of its very same ability to intervene and control the background, against which a novel signal can be

causally inferred. The thesis of the epistemological priority of ordinary experiments over computer simulations (EPT₂) seems to lose its bite also along CIC₃ lines.

Does this mean that the detection of the Higgs boson and its properties necessarily depend on computer simulations? As our discussion in Section 3 about the 2-photon decay channel shows, there are in fact different avenues all leading to the same experimental results about the Higgs and its mass. Some (4-lepton decay) rely more heavily on simulations than others (2-photon). Thus, there is a sense in which the detection of the Higgs boson need not be entirely entangled with the fortune of computer simulations, and their reliability. In other words, we would like to add a proviso to Morrison's claim that the detection of the Higgs boson is logically and causally dependent on computer simulations. This proviso notwithstanding, computer simulations are on a par with / interchangeably used with experiments at ATLAS, whenever considerations about uncertainty in the background determination make it necessary.

Ultimately, different decay channels, with their respective experimental and technological set-ups, respond all to nature and experimental data. But responding to the "tribunal of experimental data" is not what is at stake in the *causal interaction claim*. As we hope to have shown in this article, what is at stake in the causal interaction claim is, at best, not immediately transparent. Our first goal was to get clear about the claim in order to assess its epistemic force, when it comes to the alleged epistemological priority of ordinary experiments over simulations (EPT₂). Our second goal was to review a timely and paradigmatic case study, where computer simulations are routinely used as an integral part of a larger experimental set-up to make possible causal interactions claims of various kinds.

Acknowledgements

We thank the editors and the reviewers for very constructive comments on earlier versions of this paper.

Figures

Figures B-D © Copyright 2013 CERN reproduced under CC-BY-3.0 license

References

- Anderson, (1936) "The production and properties of positrons", *Nobel Lecture*, 12 Dec. 1936.
- ATLAS Collaboration (2010), *The ATLAS simulation infrastructure*, European Physical Journal C 70.3 823-874.
- ATLAS Collaboration (2013a), *Measurements of the properties of the Higgs-like boson in the two photon decay channel with the ATLAS detector using 25 fb⁻¹ of proton-proton collision data*, ATLAS-CONF-2013-012
- ATLAS Collaboration (2013b), *Measurements of the properties of the Higgs-like boson in the four lepton decay channel with the ATLAS detector using 25 fb⁻¹ of proton-proton collision data*, ATLAS-CONF-2013-013
- Franklin, A. (1997) "Calibration", *Perspectives on science* 5, 31–80.
- Frigg, R. and Reiss, J. (2009) "The philosophy of simulation: hot new issue or same old stew?", *Synthese* 169, 593–613.

- Giere, R. (2009) “Is computer simulation changing the face of experimentation?”, *Philosophical Studies* 143, 59–62.
- Guala, F. (2005) “Models, simulations, and experiments”. In L. Magnani and N. Nersessian (eds.) *Model-based reasoning: science, technology, values* (Kluwer), pp. 59–74.
- Humphreys, P. (1991) “Computer simulation”. Proceedings of the 1990 Biennial Meeting of the Philosophy of Science Association, A. Fine et al. (eds.). East Lansing, MI: The Philosophy of Science Association. Vol. II: 497:506.
- _____. (1994) “Numerical experimentation”. In *Patrick Suppes: Scientific Philosopher*, P. Humphreys (ed.), Kluwer, pp. 103–24.
- Lange, M. (2013) “What makes a scientific explanation distinctively mathematical?”, *British Journal for the Philosophy of Science* 64, 485–511.
- Morgan, M. (2003) “Experiments without material intervention: model experiments, virtual experiments and virtually experiments. In *The philosophy of scientific experimentation*, H. Radder (ed.), University of Pittsburgh Press, pp. 216–235.
- _____. (2005) “Experiments versus models: new phenomena, inference, and surprise”, *Journal of Economic Methodology* 12, 317–29.
- Morrison, M. (2009) “Models, measurement and computer simulation: the changing face of experimentation”, *Philosophical Studies* 143, 33–57.
- _____. (forthcoming) *Reconstructing Reality: Mathematics, Models and Simulations*, OUP.
- Parker, W. (2009) “Does matter really matter? Computer simulations, experiments, and materiality”, *Synthese* 169, 483–96.
- Riordan, S. (2013) “The Making of the Kilogram: 1789–99”, PhD dissertation.
- Steele, K. and C. Werndl (2013) “Climate models, calibration, and confirmation”, *British Journal for the Philosophy of Science* 64, 609–35.
- Tal, E. (forthcoming) “Making Time: A study in the epistemology of measurement”, *British Journal for the Philosophy of Science*
- Teller, P. (2011) “Two models of truth”, *Analysis* 71, 465–72.
- _____. (2013) “The concept of measurement-precision”, *Synthese* 190, 189–202.
- Winsberg, E. (2009) “A tale of two methods”, *Synthese* 169, 575–92.
- _____. (2010) *Science in the Age of Computer Simulation*, The University of Chicago Press.