Epistemic entitlements and the practice of computer simulation

John Symons and Ramón Alvarado
University of Kansas
(Draft 5/16/18)

Introduction

Arguably, computer simulation has become a third major branch of science, supplementing traditional theoretical and experimental research. Computer simulation permits investigation into topics where cognitive, ethical, political, or practical barriers would otherwise prevent scientific inquiry (See Symons and Boschetti 2013, 809). In important and controversial matters such as nuclear weapons testing, climate science, and studies of the behavior of epidemics, to take just a handful of cases, computational simulation is often our only viable means of investigation. Thus, we see an increased focus on how to determine the appropriate level of credence that we ought to assign to the results of computer simulations (Boschetti et. al 2012). This is not only a concern for the scientific and philosophical communities. Understanding the epistemic status of the scientific claims produced using computer simulations presents an increasingly pressing practical challenge for public policy and commercial decision-making (See Vallor 2017).

We begin with the assumption that computer simulation has made positive contributions to scientific progress in a variety of fields.[1] What we mean by computer simulation is broadly inclusive. On our view, computer simulation is the scientific or engineering practice of studying the behavior of actual or possible natural systems using software implemented in electronic devices. This software often (but not always) relies on or approximates the behavior of mathematical models of the natural systems of

[1] Astrophysics, climate science, epidemiology, fluid mechanics, and particle physics are some of the scientific disciplines in which computer simulations have played a crucial role advancing the state of our knowledge (Oreskes et al., 1994; Oreskes, 2004; Norton and Suppe, 2004)

interest. The algorithm coded into the software itself is intended to implement a function that takes as an input a representation of some initial state of the natural system and gives as an output the state of the system at some later time. The simulation repeats this operation on the outputs of previous iterations of the operation, thereby creating a sequence of representations that can shed light on the unfolding dynamics of the natural system.

In this paper, we explore the conditions under which one is justified to treat simulations as authoritative sources of knowledge. The question of when it is rational to trust computer simulations involves a range of distinguishable questions. Prominent among these include the following:

- How should we judge the results of simulation in areas of science where empirical evidence and data are relatively scarce?
- What justifies our credence in areas where we have abundant data but where no settled theory dominates the scientific domain?
- With respect to computational methods themselves, can we be confident that the software generating a simulation is free of error?
- Given the complexity of modern simulations and the increasing importance of machine learning in science, to what extent should we expect human understanding or humanly graspable explanations to play a role in science?

Taken in isolation, these are not new questions for philosophers of science. However, the epistemology of computer simulation brings them together in novel ways (Frigg and Reiss, 2009; Humphreys, 2009).[2] Moreover, the way we respond to these questions is likely to shape the practice of science and engineering moving forward.

While we cannot answer all of these questions in detail in the space of a single paper, they all revolve around the question of our entitlement to trust computer simulations. We argue for a relatively conservative position, namely that a high-level of

[2] Philosophical treatments of the epistemology of computer simulations are in an early stage of development and work on this topic so far has been covered well by Eric Winsberg in his *Stanford Encyclopedia of Philosophy* article (2015).

trust in the results of computer simulations is only appropriate when simulations are grounded in well-curated data, established scientific theory, empirical evidence, and good engineering practices. Though these norms may seem obvious, they are not always followed in practice and their is not always emphasized in philosophical discussions. Simulations that are not grounded in these ways, we argue, may turn out to be interesting sources of insight or inspiration, but they should not be accepted as guides in decision-making where potentially harmful or expensive risks are involved.[3]

Our position is motivated by reaction against an excessively pragmatist strain in discussions of computer simulation and modeling. We argue that accounts of epistemic entitlement for computer simulation simply cannot be grounded in straightforwardly pragmatic terms. Crude forms of pragmatism should not be the philosophical response to the problems of the epistemology of computer simulation for reasons we will explain here.

In addition, we argue that epistemic entitlements for computer simulations cannot be recast in terms of the principles governing the behavior of individual human agents in everyday epistemic contexts. To invoke everyday epistemic practices to justify our reliance on computer simulation is to fail to recognize the stringent standards governing our best scientific inquiry. We are particularly critical of the tendency in the epistemology of computer simulation to take Tyler Burge's Acceptance Principle as a starting point for an account of epistemic entitlement in this context.

Our positive proposal is very straightforward. We argue that simulations by themselves are empty. They are grounded only insofar as they rely on established theory, good engineering practice, well-curated data, and empirical evidence.

1. Why Pragmatism Fails as an Account of the Epistemology of Computer Simulation

Our discussion of epistemic entitlements for simulations begins with a thought experiment involving a mechanical oracle. As we shall see, the oracle produces reliable

[3] Policy regarding large scale institutional interventions, life-critical systems assessment, pharmaceutical tests and the outcome of otherwise untested medical procedures fall within this category.

predictions but does not grant its users understanding or explanation concerning the system whose states it predicts or concerning its own inner workings. The differences we note here between the mechanical oracle and the actual practice of simulation in contemporary science and engineering should show the limits of pragmatist and Burge-style approaches to the epistemology of computer simulation. We will explain, by contrast with the oracle, where we should look in order to understand how trust in actual simulations is and ought to be warranted.

One implication of our account will be that some prominent uses of "simulation" in the social sciences, for example, much of what is called agent-based modeling, may provide interesting sources of insight but will not count as trustworthy guidance with respect to policy or other kinds of decision-making. [4] We should not, and (for the most part) do not treat simulations as oracles. As we shall see, our trust in simulations is grounded in much more complex practices than simply deferring to a record of predictive success.

1.1 The Tale of the Mechanical Oracle

Imagine, that once upon a time, there was a civilization in which people had outsourced the task of predicting and controlling their natural and social environments to black boxes with blinking lights. The story of how these boxes appeared was lost in the mists of history, but some said that the boxes appeared suddenly, like the obelisks in Arthur C. Clarke's *2001: A Space Odyssey*. Unlike Clarke's obelisks, the boxes with blinking lights were talkative and responded to questions. At first, people might have approached them with trepidation; naturally, they would have been unsure of whether the boxes could be trusted. However, over time, those who trusted the boxes gained an advantage over those who did not. Eventually, the practice developed whereby almost everyone asked the boxes questions with the expectation of receiving predictively accurate answers.

While the mechanical oracles became sources of guidance and prediction for their users, the inner workings of these machines remained a mystery. When users asked the

[4] Nicole Saam (2017), for example, thinks that given the obstacles posed by epistemic opacity, only those computer simulations that are generated and treated as *thought experiments* (as opposed to laboratory experiments) should be considered in the social sciences.

boxes for an explanation of their inner workings, the boxes simply refused (or were unable) to provide one. Users *assumed* that the boxes obeyed the usual physical laws and that their inner workings were in accord with mathematical principles of computation, but this assumption was simply part of their shared commonsense metaphysical background theory rather than having anything directly to do with what was known specifically about the artificial intelligences in the blinking boxes. Neither the details of the software engineering process that produced the oracles nor the machine learning algorithms that seemed to support their operation were accessible to the users of the artifacts.

The consensus among users was that the boxes rarely made serious errors with respect to important matters like the paths of hurricanes, the likelihood of epidemics, or the behavior of the economy. They were, by all relevant measures, predictively accurate. For the sake of argument, let us suppose that the users of the oracle are correct in their judgment of its predictive power. And thus, users posed questions to the oracle, took past successes as evidence that future answers were reliable, and granted the boxes significant influence in decision-making. Eventually, it was widely agreed that ignoring the oracle in matters involving important decisions would be unethical.

1.2 Are We Oracle-believers?

It is tempting to analogize the situation of the oracle-believers and our own relationship to at least some prominent computer simulations. At first glance it seems that in several important contemporary decision-making contexts we already depend heavily on systems that fit the profile of our imagined mechanical oracles. Risk assessment algorithms, for example- which compute the many factors that may play a role in an individual's probability of repeated criminal behavior- or data analysis techniques used to deploy policing resources in certain neighborhoods are often deeply opaque to those that use them (Symons and Alvarado, 2016; See also O'Neil, 2016). It is often the case that we do not have sufficient understanding of the workings of these computational methods to evaluate each of the epistemically relevant steps in its operation. Paul Humphreys described systems of this kind as "epistemically opaque" (2009). Likewise, given the complexity of modern software, we cannot fully survey the operation of computational simulations, nor can we be sure that they operate correctly all the time (Symons and Horner, 2014: Horner and Symons, 2014). Little surprise then that reliance on software intensive systems in science has struck some thinkers as questionable and perhaps unscientific (Newman, 2015). If we are ignorant of the method by which the simulation generated its answer, then indeed it seems that we do not have rational basis to believe the answers. But this intuitive reaction is based on hastily accepting the analogy between the mechanical oracles of our thought experiment and actual computer simulations.

In scientific and commercial contexts it certainly seems as though we have already come to rely heavily on software intensive systems. Given that science and commerce are, for the most part, successful, it looks as though our faith in simulation is not unreasonable. However, we are not in the same position as the oracle-believers of our thought experiment. We will argue, for instance that the oracle-believers are not engaged in *science* in any standard sense of the word. Part of the task of the epistemology of computer simulation is to explain the difference between the scientist's position in relation to epistemically opaque computer simulations and the believers' relation to their mechanical oracles in the story above.

Scientists and engineers are often motivated by pragmatic considerations. Given short-term commercial interests and other pressures it can be especially tempting to adopt a broadly pragmatic attitude towards simulation. However, if we restrict ourselves to pragmatic considerations alone we will not be in a position to explain the differences between us and the oracle-believers. As we will see below, pragmatic reasons should be distinguished from properly epistemic reasons.

Given something like the pessimistic meta-induction with respect to the truthfulness of science over the course of history, if our criterion for trust is the history of past successes or failures, then the oracle-believers are certainly more entitled to believe the blinking boxes than we are to believe our latest, best science. The pragmatist must defer to the oracle, since his criterion for deciding between theories is their relative predictive accuracy. While it is certainly true that predictive power is an important desideratum in choosing between theories or models, it is not the only dimension of scientific inquiry (See Symons and Boschetti 2013). Predictions can be accidentally true, predictions can be trivially true (Everyone alive today will be dead in 300 years), or predictions could be true by magic or clairvoyance. In none of these cases would we ordinarily call the sources of these predictions scientific.[5]

In cases like the oracle, in which the system generates reliably good predictions, most of us would wonder *why* the oracle works so well. Indeed, it is often the case, in the history of science that natural regularities and correlations become the *target* of explanation. Thus, for the scientifically-minded the mysterious predictive power of the oracle would become the explanandum. For the pragmatist, since predictive power is the only significant criterion of success, explaining predictive power could only be justified in terms of the pursuit of some additional, enhanced, predictive power. By stipulation, the oracle in our example has maximal predictive power. Therefore, in the case of the oracle, there is nothing more for a pragmatist to do or to ask for. The inability for the pragmatist to shed much light on the difference between the oracle and ordinary scientific practice is indicative of the weakness of pragmatism as a philosophical position, particularly so in the philosophy of science.

However, this is not merely a matter of pragmatism's failure as a satisfying epistemology. By failing to capture the actual factors by which we rationally come to trust simulations, we risk neglecting the practical processes by which rational trust in science that depends on computational simulation ought to be cultivated in practice.

2. The Reasonableness of Acceptance

One way that philosophers of science have suggested explaining *why* we trust computer simulations appeals to the ways that epistemic entitlements work in other less controversial forms of inquiry (See for example Barberousse and Vorms, 2014). For example, it might make sense to trust simulations in the same way that we trust our perceptual faculties and our memories, or in the same way that we trust expert testimony. We are entitled to base beliefs on the evidence of our senses or the testimony of experts in spite of not having full access to either the underlying workings of the senses or the

[5] A position of this kind leads Quine to concede that he would have to count genuine clairvoyance as science. (1974) While genuine clairvoyance would be extremely interesting and useful, it would be a phenomenon that is largely orthogonal to science.

education and abilities of the expert. It seems that, for all ordinary purposes, one is warranted in believing that the universe is over one billion years old without knowing more than a few paragraphs worth of cosmology. Similarly, we can trust that we are correctly identifying the smell of freshly baked bread even though we might have no understanding of the operation of the olfactory system.

Perhaps something similar is the case for computer simulations. Let's call this the analogical approach to entitlements. Computer simulations are enormously complicated and in general they operate in ways that are not surveyable by ordinary human minds. Computer simulations are also epistemically opaque in some cases. Under these circumstances, some argue, it might still be reasonable to trust computer simulations given the assumption that everything is working smoothly. When this line of reasoning appears in the philosophical literature, it frequently draws on insights from Tyler Burge's work. In particular it takes as a starting point Burge's view of how human beings ordinarily maintain a posture of acceptance in epistemic matters. Just as we ordinarily trust our senses, we also ordinarily tend to accept what other people say unless we have reason to disbelieve them.

> "Justification in acquiring beliefs from others may be glossed, to a first approximation, by this principle: *A person is entitled to accept as true something that is presented as true and that is intelligible to him, unless there are stronger reasons not to do so*. Call this the *Acceptance Principle*. As children and often as adults, we lack reasons not to accept what we are told. We are entitled to acquire information according to the principle-without using it as justification-accepting the information instinctively."(1993 p. 467)

Burge points out that this ordinary disposition to accept the testimony of others is a necessary condition for acquiring language, and for a range of other social phenomena.[6] As far as science is concerned, failure to act in accordance with the Acceptance Principle makes rational collaborative projects of inquiry impossible. On this view, testimonial

---

[6] Burge's account of the Acceptance Principle (as he acknowledges) is very similar in spirit to the Principle of Charity, as it figures in Quine (1960) and Davidson (1973). The principal difference between these principles is the role that Burge's notion of preservation of content plays in his account.

sources such as published research, standard pieces of scientific equipment, and expert opinion should be trusted by default as a precondition for the possibility of ongoing inquiry.

For Burge, acceptance is the epistemic default position for human beings and it grounds what he calls an *apriori entitlement* whereby a person is entitled to accept a proposition that is "presented as true and that is intelligible to him unless there are stronger reasons not to do so, because it is prima facie preserved (received) from a rational source, or resource for reason; reliance on rational sources- or resources for reason-is, other things equal, necessary to the function of reason" (1993 p.469). Given that the Acceptance Principle serves as a necessary precondition for the function of reason and therefore of science, it has been natural for philosophers influenced by Burge to imagine that a priori warrants can ground computer simulations. This is particularly the case if one sees computational methods as mere extensions to existing mathematical methods.

Clearly there are occasions in which one is entitled to a belief independently of one's subjective grasp of the epistemic rights or warrants supporting that belief. We can be entitled to believe, for example, what the weather forecaster has told us independently of the state of our knowledge of meteorology (Dretske, 2000; Williams, 2000; Adler, 2015). Given the epistemic entitlements described by Burge, we are frequently exempted from such justificatory practices as citing evidential support or of giving reasons more generally (Wright, 1985; Burge, 1993; Lackey, 1999; Dretske, 2000; Davies, 2004; McGlynn, 2014). In Sections Four and Five we will discuss when we are legitimately exempted from justificatory practice in more detail.

3. Computer Simulations are not Transparent Conveyers

Following Burge's application of the Acceptance Principle to the epistemology of computer-assisted mathematical proofs, Anouk Barberrouse and Marion Vorms (2014) argued that the epistemic warrants supporting our reliance on computer simulations are not intrinsically constituted or enhanced by appeal to empirical evidence. They argue that the warrants behind our trust in computational methods can and do take the form of Burge-style entitlements. Just like other devices whose role, like that of memory or perception, is to transmit and preserve intact information, computational methods that manipulate mathematical content appropriately are called *transparent conveyers* (Burge, 1993; Barberousse and Vorms, 2014). In a transparent conveyer if the propositions being conveyed are justified a priori, for example, then these justifications will not change from a priori to empirical. [7] Following Burge, they call this *content preservation*. [8] From this perspective, a warrant for belief can be a priori even if the only way one can attain the warrant depends on some contingent fact about the world, such as the fact of having a brain or some particular perceptual capacity. The fact that a human being needs a brain to do arithmetic is irrelevant to the justification of an arithmetical proposition. One implication of this position, is that a cognitive capacity such as memory "is no more intrinsically an empirical faculty than it is a rational faculty. Its function in deductive reasoning is preservative" (Burge, 1999)[9]. In other words, when memory does what it is supposed to do it conveys information without altering it.

However, unlike perception and memory, or even the human tendency to trust others, the introduction of computational methods into scientific inquiry has been the product of artifacts and practices whose reliability has involved a gradual process of hard work on the part of scientists and engineers (Winsberg, 2010). As Evelyn Fox-Keller points out, contemporary uses of computational methods in science are the product of massive collaborative efforts since the Second World War involving trial and error approaches to practical problems (Fox-Keller, 2003).

As Eric Winsberg argues, this historical development includes an array of distinguishable evidential benchmarks and other features such as considerations of fit,

---

[7] This is sharp contrast to the view, called the materiality thesis (Morgan, 2005; Parker, 2009), which stipulates that the devices by which information is manipulated bear significant epistemic import.

[8] Though they acknowledge that there is a substantial difference between computer assisted mathematical proofs, such as the ones Burge focused on and complex computer simulations in terms of content preservation, they justify the aprioricity of a scientist's entitlement to trust a simulation in virtue of a second strategy, which we will inspect in detail in section three: trusting computer simulations, they argue, is like trusting expert testimony. Though expert testimony may be fallible, for this view, casting a general doubt on the practice absent specific reasonable doubt can be seen as irrational.

[9] Similarly, when we rely on our senses we grant that when they are working the way they are supposed to they transmit information without altering it. That is, as explained above, they are transparent conveyers. Thus, though one can acknowledge their fallibility, in the absence of a plausible reason to doubt their well-functioning, it is rational to rely on the senses (Burge, 1993)

calibration, and even extra-theoretical and extra-mathematical engineering practices (Winsberg, 2010). These practices involve both hardware and software innovations that are well-known in practice but are seldom discussed in the epistemology of science. From architectural considerations for optimal processing to mathematical discretization, any trust assigned to computer simulations involves facts about the history of their successful deployment but also the appropriate management and assessment of uncertainties, errors and calibration procedures (Ruphy, 2015). These sanctioning processes, and therefore the computer simulation itself, should not be regarded as transparent in the sense required by Burge as we will explain below.

Consider discretization techniques. Many computer simulations are the result of a process that includes the transformation of differential equations in a mathematical model into manageable step-by-step algebraic expressions that represent approximate values for specific finite spatiotemporal states of a system.[10] Discretization techniques involve determining practical ways of implementing calculations in manageable chunks. This transformation/translation procedure introduces epistemically relevant decisions on the part of the modeler that are distinct from the original mathematical model. That is, the reason a specific set of equations is chosen in the *mathematical* model has nothing to do with the challenge of approximating the model via a computer simulation. By contrast, the choice of specific discretization techniques in a computer simulation will be responsive to the practical necessity of implementing a model in a digital device. In fact, when it comes to discretization techniques, the decision to select one technique over another is often a matter of engineering trade-offs (Winsberg, 2010 p.12, 23).

Discretization techniques, a fundamental aspect of computer simulations, undermine the possibility that computer models can serve as transparent conveyers insofar as they alter the nature of the epistemic justification of the content being manipulated. And this is so even when one grants that they are working as intended. In fact, it is a feature of how discretization techniques are applied. Further, even if the reasons to trust the results of the equations in the original mathematical model are

[10] One can think of discretization as trying to approximate a circle by drawing one regular polygon after another with more sides each time starting form a square. Of course a square is a terrible circle, but a polygon with millions of sides may be visually indistinguishable for practical purposes. Nevertheless at each point, one is not drawing a continuous curve but rather a series of straight lines at an angle from each other.

grounded in well-established theory, or are supported a priori on purely mathematical grounds, the introduction of discretization techniques involves an engineering element that alters the justificatory considerations involved in a computer simulation of the mathematical model. Thus, in a very basic sense, computer simulations cannot count as transparent conveyers given Burge's characterization because justificatory elements distinct from the ones warranting the original content do in fact enhance, decrease or constitute the epistemic warrants of the manipulated content (McEvoy, 2008).[11] Furthermore, computer simulation involves independent epistemic warrants at different stages. Consider the following. Even if the mathematical model is fully warranted and works as intended and even if the discretized version of the model also works as intended there is no reason to think that we have reason to trust the latter *because* of the former.

When a discretized model is ultimately implemented in a device, for example, it requires yet another epistemically relevant transformation in the process of coding. In coding, considerations of fit, trust and/or reliability of a given algorithm will depend on independent factors from those involved in the discretization process. Unlike discretization techniques that involve established techniques and theories, code is often the result of highly idiosyncratic problem-solving approaches. Coding practices are often error-prone. Consider for example the way many significant software bugs are 'patched'. That is, they are not 'fixed' per se, as in actively engaging with their malfunctioning components or erasing an erroneous block of code. Rather, code is added that 'patches' the original code by superseding previous functionality. These patches almost always introduce their own new bugs, making the process of assessing the reliability of the software even harder than it already was. In the process of patching, as is often observed, software, which is at the core of any and all processes of computer simulations, tends to grow over time merely because it can (Holzmann, 2015). This is especially the case now that memory has become so inexpensive in modern computing. The 'true' command in Unix, for example, which originally consisted of an empty file with nothing to execute grew to nearly 23,000 bytes from 1979 to 2012 (2015). The inevitability of errors in the

[11] McEvoy in his response to Tymoczko and Kitcher on the aprioricity of computer assisted mathematical proofs concedes as much by saying "What determines whether a proof is a priori is the type of inferential processes used to establish the conclusion of that proof. If the method of inference for any of the steps in the proof is a posteriori, it is a posteriori" (2008 p.380).

practice of coding is an empirical reason to decrease trust in any content manipulated via computational methods. Thus, even if one grants that a computer simulation seems to work as intended, it cannot be regarded as a transparent conveyer in Burge's sense.

In a similar vein as the views illustrated above and in contrast to our own, Beisbart (2017) argues that warrant transmission takes place from one stage in the simulation process to the next. On his view, the results of computational methods can be said to provide knowledge in virtue of a few inferential transfers from one level of propositional content to another. That is, at each step, there is a result, a proposition to be considered: from target phenomenon to mathematical model, from mathematical model to computerized model, from computerized model to computational implementation, and from single implementation to iterations (Beisbart 2017 p.160-162). At any given step our trust is justified if we assume that that our methods are the adequate ones and we assume they work as intended. In other words, we are warranted in inferring from one stage to the next if we grant that each level is somewhat/somehow warranted to begin with. For Beisbart an agent is "inferentially justified in believing a propositional result constructed from a computer simulation if she is justified in believing the dynamic equations used to feature the system under scrutiny and if she is justified to think that the simulation works as intended" (2017 p.171). However, for the agent to be justified in believing that the simulation works as intended, all that is required is that the "epistemic agent has sufficient reason to believe that it does so" (p.169). This "sufficient reason" in turn implies, the *assumption* that at every inference step, from mathematical model to final display, certain types of errors (round-off errors, modeling errors, hardware failures) are "excluded or are sufficiently small" (p.161-163). At every step however, trust is warranted with the proviso that there are no significant epistemic challenges (2017,162). As Beisbart describes matters, a scientist is ultimately justified in believing a proposition *q* derived from a computer simulation because what the computer itself has done is drawn inferences from the propositions of a discretized conceptual model (2017, p.165). The scientist draws the final inference that ties the results back to the target phenomenon by assuming that at any given step the methods involved *worked as intended*. Note that Beisbart recognizes the distinction between warrants supporting the content of simulations and warrants supporting our reliance on their results. His analysis highlights the empirical and historical considerations that ground our trust in computational methods.

However, as we have seen, the practices involved in simulation processes are such that they are not the product of transparent conveyers *even when they seem to work as intended.* The reasons why each of the steps can be said to be reliable and/or trustworthy are independent from the step before or after it, offering no single kind of warrant transmission that spans the entire process of computer simulation (Beebee, 2001; Davies, 2004; Pryor, 2012; Moretti and Piazza, 2013). As we will explain below, whatever reason we have for trusting the results of one stage in a simulation as we move to the next should generally not be ad hoc; should not be based solely on the factors that are unique to particular instances of a simulation practice itself but should be the product of established theoretical principles and engineering standards.

It is important to note that the methods and technologies used in practice are not simply deployed in an unchanged form across different applications and platforms. They are modified, often in non-trivial ways, to suit the task at hand. Throughout these modifications there are ways in which one can establish benchmarks that help to sanction the results of a simulation. These might involve, for example, repeated runs and internal comparison, comparison to outputs of other simulations, or most importantly, comparison to real-world data (Symons, 2008; Winsberg, 2010; Gramelsberger, 2011). Thus, when computer simulations can be trusted it is because of their adherence to theoretical principles, empirical evidence, or engineering best practices and not because of their output alone.

In addition to the importance of recognizing different kinds of epistemic practices in different stages of the construction of the simulation, it is also the case that even when the foundational components of the simulation are sound, the processes of combining them into a working application can fail in challenging ways. To put this point simply, we can say that the inference from warrants at the level of “parts” to warrants at the level of the “the whole” computer simulation is not sound. The reason that it is a mistake to regard an epistemic warrant as being successfully conveyed from relatively simple, “component-level” inferences to the behavior of the simulation as a whole, even in cases where the components of a software system are well understood, is because of the role of software engineering itself. Software, which is an essential part of any general purpose computing system, is a source of error insofar as it is created via human engineering and is built to serve human purposes. Human engineered software contains errors at the level of coding and vagueness at the level of specification (Tymoczko, 1979 p.74). These errors mean that normal notions of epistemic-warrant-transmission from elegant and well-understood foundations to the behavior of aggregates of these components simply do not hold (Winsberg, 2010; Kaminski, 2016; 2017). In other words, as Stéphanie Ruphy notes, "the computer simulation does not simply inherit" its epistemic credentials (2015 p.139) Rather, its reliability is rather the product of a range of diverse sanctioning processes.

In practice, our trust in computer simulation is fundamentally unlike the epistemic relation of the oracle believers to their mechanical oracles in our thought experiment. In themselves, computer simulations do not constitute good sources of testimony. Our trust in computer simulation takes place within a tradition of engineering and scientific practice. As we have seen, the actual process of building the simulation involves distinct stages, each of which has its own epistemic standards and warrants.

4. Computer Simulations are not Sources of Expert Testimony

According to Burge-style views of epistemic entitlement, when we are dealing with a given piece of testimony that we have *no plausible reason to doubt*, the default rational position is to accept it as truthful (Burge 1993). If, for example, I hear that weather models predict that a hurricane is likely to hit my city, it would rational for me to heed the meteorologist's warning without rigorously investigating the methods and evidence supporting the prediction. In this scenario, it would seem that I am trusting the model in the same way that I would trust an expert, or perhaps even in the same way that the oracle believer in our thought experiment trusted the boxes with the blinking lights. This, however, is not the right way to understand the non-scientist's attitude towards the prediction. First, we would argue that in the weather forecasting case I do not trust the models as experts. Instead, I trust the judgment of the human meteorologists with respect to the models in question.

Second, while it is reasonable for me to regard the meteorologist as a reliable source of testimony (Borge, 1993; Jenkins, 2007), it would not be reasonable for the meteorologist him- or herself to maintain this epistemic attitude towards the computer simulation. It is reasonable for us, in ordinary life, to trust the weather forecaster insofar as they have expert standing. Neither we, nor the weather forecaster should treat the simulation itself as an expert. Again, the simulation in itself is not self-validating. When things go well, *the community of experts and the tradition of using simulations in successful scientific practice grounds* the confidence of experts in the use of simulations.

A scientist may be justified in relying on the computer simulation's pronouncements once it is integrated into her practice or into the practice of other experts (Audi, 1997). As explained above, this is a long and gradual process. It is never the case that a scientist would give as a reason to trust a simulation the claim that the simulation is an expert. Treating our trust in simulations by analogy with our trust in human experts, would be to miss the actual warrants that ground the trust granted by human experts to simulations. When non-experts take the testimony of experts seriously, they are not granting that same level of credence to the tools used by the experts they trust. For example, they might not even know of the existence of those tools. In our case, most consumers of the weather forecast simply trust the experts to interpret computer simulations and other tools for them.

Justification to trust testimonial evidence comes from observing a "general conformity between facts and reports [by which] with the aid of memory and reason, we inductively infer that certain speakers are reliable sources of knowledge." (Lackey, 1999 p.474). It is the conformity between facts and reports that does the heavy epistemic lifting when we trust experts. This may be the case for those humans that we consider experts, but it is not the case for computer simulation. For simulations, the reasons that experts would count them as reliable sources of knowledge is their relationship to theoretical principles and engineering best practices in addition to their predictive successes. For example, a reason to trust a given computer simulation of a system is not merely because it is able to predict a given state of the system, but rather *because it does so in the appropriate way*- in conformity with the laws of physics, in conformity with engineering practices that are amenable to error assessment, etc. (Resnik, 1997; Symons and

Alvarado, 2016) The task for experts in the evaluation of computer simulations is fundamentally different from the consumption of computer simulations by non-experts. Non-experts depend on experts to have certified those simulations as worthy of attention. The testimony that non-experts are trusting is the testimony of expert communities, not directly the output of their simulations.

5. Scientific Inquiry and Everyday Epistemic Practices

We have argued that Burge-style approaches to epistemic entitlement are inappropriate in the context of computer simulations when it comes to scientific or other high-stakes applications. In areas of life that involve great risk or expense, appropriate epistemic diligence will be called for. However, we contend that science should be, and often is also organized around stringent epistemic norms for the evaluation of computer simulations. In fact, it might even be the case that since time constraints are less pressing in much of science than they would be in, for example, military or policy contexts, *scientists have the luxury of being even more rigorous and demanding with respect to the epistemic standards governing their devices.*

Clearly there are aspects of scientific practice that are just like the epistemic practices that we would find in ordinary life. Scientists are people too, and in this sense, Burge's reflections on the Acceptance Principle will have an important role. We are not denying the fact that in order for scientists to engage in research, they need to trust one another, they cannot maintain an attitude of radical skepticism, and they need to make reasonable trade-offs between time spent in various epistemically relevant tasks.

However, there are a few distinctions that ought to be drawn in this discussion. First, philosophical treatment of computer simulations should distinguish between the kinds of entitlement being appealed to in distinct contexts. Is the entitlement in question, for example, an entitlement to trust, or an entitlement to not doubt. [12] Second, there are

[12] It is unclear, for example, the extent to which Barberousse and Vorms assume a negative (entitlement to not doubt) or positive (entitlement to trust) role for entitlement in their arguments. However, this particular paragraph focuses on the completely distinct categories between kinds of entitlements (a kind of warrant that gives someone *a reason* to believe (x)) beyond their epistemic variety.

epistemic and non-epistemic entitlements that can be invoked as justifications to trust something. For instance, an epistemic entitlement might be playing a role in our decision to trust a medical device given our belief that it may be beneficial to a patient. A non-epistemic reason to believe might involve our commitments to our affiliation group. There might be good social reasons to believe what a particular person says in cases where the matter under consideration is trivial or involves no harms or significant expense. If the cost of error is low, then arguably it is reasonable to simply accept the other person's position, in cases where the social (or other) benefits are important. An individual might accept or reject some scientific hypothesis for social reasons. This can be regarded as reasonable insofar as that person's individual commitments have little practical import. Her views for example on climate change, evolution, or other socially controversial matters have little real significance for either the practice of science or for policy decisions based on the scientific hypothesis in question. Insofar as their position on these matters causes no significant harm, it is at least understandable why consumers of science would choose to practice bad epistemic hygiene with respect to the science in favor of strengthening their position within an affiliation group. They strengthen their place within their social network by assenting to the consensus position of that group, perhaps even assenting to an exaggerated version of that consensus position. In those cases one may indeed have reasons to believe/trust, but they are not strictly speaking *epistemic reasons* (See for example Dretske, 2000 p.594). By contrast, according to Dretske, for example, if what we are appealing to in accepting some proposition as true is to be regarded as an epistemic reason then the only grounds are its truth or possible truth (p.593-4).

The truth or possible truth of the values produced by an instrument deployed in the aid of science can only be a product of a reliability assessment, which is in turn distinct from personal justification (Williams, 2000). Epistemic entitlement for computer simulations, should not be reduced to, or confused with the effort to provide what Dretske calls "a pragmatic vindication of the practices we use to fix beliefs" in place of efforts to "validat[e] the beliefs themselves" (2000, 598). Merely showing that a "practical purpose or need is served by accepting certain practice does not address the problem of our *epistemic* right to accept the individual beliefs that occur in this practice." (ibid)

Scientists are held to a higher epistemic standard than ordinary epistemic agents (Nola and Sankey, 2014). This is one of the reasons why they can serve the role of providing expert testimony. While no individual scientist is an epistemic saint, the scientific enterprise collectively seeks to ensure truth-aptness and reliability independently of the shortcomings of each individual practitioner (2014). In relation to modeling and simulation, this has practical import. Humphreys, for example thinks that practitioners should always possess working knowledge of a given instrument and he argues that the background theory of its principles should always be at hand to the practitioner. This is not only because some instruments, for example, an MRI scanner, require this knowledge in order to be used effectively (Humphreys, 2004 p. 38), but also because, the theory of how the instrument operates is "routinely needed to protect the user against skeptical objections resulting from ways in which the instrument can produce misleading outputs." (p.38)

Computer simulations are often used in contexts where highly demanding scientific standards are in place and where even ordinary well-grounded scientific practice falls below the threshold for acceptability (Ruphy, 2011; Morrison, 2015). The famous five sigma standard governing evidence at CERN in the discovery of the Higgs Boson (where computer simulation played a central role) is an extreme, but important instance of the kinds of high epistemic standards governing scientific research.[13]

Scientific standards would be regarded as unreasonably strict if we were to adopt the perspective of everyday epistemic practices such as normal decision-making, trust in perception, and credence with respect to the testimony of others.

6. Conclusion

We have argued that the epistemology of computer simulation is a highly complex domain where assumptions concerning the transmission of warrants and the characterization of epistemic entitlements must be understood in relation to established

[13] The five-sigma standard corresponds to a p-value, or probability, of $3x10^{-7}$, or about 1 in 3.5 million. In this case it is the probability that *if* the particle does *not* exist, the CERN team would find what they observed. It is extremely unlikely that they could have generated the data by accident. How unlikely? 1 is 3.5 million unlikely!

scientific practices. The actual epistemic situation is fundamentally dissimilar to the situation depicted in our thought experiment. Neither pragmatism, nor an analogy with ordinary epistemic practices will prove useful as guides to the epistemology of computer simulations.

Instead, we have argued that trust in simulations should be grounded in empirical evidence, good engineering practice, and established theoretical principles. Without these constraints, computer simulation risks becoming little more than unmoored speculation

Bibliography

Adler, Jonathan, "Epistemological Problems of Testimony", *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2015/entries/testimony-episprob/>.

Audi, Robert. "The place of testimony in the fabric of knowledge and justification." *American Philosophical Quarterly* 34, no. 4 (1997): 405-422.

Barberousse, Anouk, Sara Franceschelli, and Cyrille Imbert. "Computer simulations as experiments." *Synthese* 169.3 (2009): 557-574.

Barberousse, Anouk, and Marion Vorms. "About the warrants of computer-based empirical knowledge." *Synthese* 191.15 (2014): 3595-3620.

Beebee, Helen. "Transfer of warrant, begging the question and semantic externalism." *The Philosophical Quarterly* 51.204 (2001): 356-374.

Beisbart, Claus. "Advancing Knowledge Through Computer Simulations? A Socratic Exercise." In *The Science and Art of Simulation I*, pp. 153-174. Springer, Cham, 2017.

Bonjour, Laurence. “A priori Knowledge” in Bernecker, Sven, and Duncan Pritchard, eds. *The Routledge companion to epistemology*. Routledge, (2011): 283-293

F. Boschetti, E. Fulton, R. Bradbury, J. Symons “What is a model, why people don’t trust them and why they should” in *Negotiating Our Future: Living Scenarios for Australia to 2050.* Australian Academy of Science (2012) 107-118

Borge, Steffen. "The word of others." *Journal of Applied Logic* 1, no. 1-2 (2003): 107-118.

Burge, Tyler. "Content preservation." *The Philosophical Review* 102.4 (1993): 457-488.

Casullo, Albert. "What is entitlement?." *Acta Analytica* 22, no. 4 (2007): 267-279.

Davies, Martin. "II—Martin Davies: Epistemic Entitlement, Warrant Transmission and Easy Knowledge." *Aristotelian Society Supplementary Volume*. Vol. 78. No. 1. The Oxford University Press, 2004.

Dretske, Fred. "Entitlement: Epistemic rights without epistemic duties?." *Philosophy and Phenomenological Research* 60, no. 3 (2000): 591-606.

Dummett, Michael. "Testimony and memory." In *Knowing from words*, pp. 251-272. Springer Netherlands, 1994.

Floridi, Luciano, Nir Fresco, and Giuseppe Primiero. "On malfunctioning software." *Synthese* 192.4 (2015): 1199-1220.

Frigg, Roman, and Julian Reiss. "The philosophy of simulation: hot new issues or same old stew?." *Synthese* 169.3 (2009): 593-613.

Gramelsberger, Gabriele. "Generation of evidence in simulation runs: Interlinking with models for predicting weather and climate change." *Simulation & Gaming* 42, no. 2 (2011): 212-224.

Horner, Jack, and John Symons. "Reply to Angius and Primiero on software intensive science." *Philosophy & Technology* 27, no. 3 (2014): 491-494.

Humphreys, Paul. *Extending ourselves: Computational science, empiricism, and scientific method*. Oxford University Press, 2004.

Humphreys, Paul. "The philosophical novelty of computer simulation methods." *Synthese* 169, no. 3 (2009): 615-626.

Jenkins, Carrie S. "Entitlement and rationality." *Synthese* 157, no. 1 (2007): 25-45.

Keller, Evelyn Fox. *Models, simulation, and "computer experiments"*. na, 2003.

Lackey, Jennifer. "Testimonial knowledge and transmission." *The Philosophical Quarterly* 49, no. 197 (1999): 471-490.

Lackey, Jennifer. "Testimonial Knowledge" in Bernecker, Sven, and Duncan Pritchard, eds. *The Routledge companion to epistemology*. Routledge, (2011): 316-325.

McEvoy, Mark. "The epistemological status of computer-assisted proofs." *Philosophia Mathematica* 16, no. 3 (2008): 374-387.

McKinsey, Michael. "Transmission of warrant and closure of apriority." (2003). In Nuccetelli, Susana ed. *New essays on semantic externalism and self-knowledge*. MIT Press, 2003.

Moretti, Luca, and Tommaso Piazza. "When warrant transmits and when it doesn't: Towards a general framework." *Synthese* 190.13 (2013): 2481-2503.

Morgan, Mary S., and Margaret Morrison. *Models as mediators: Perspectives on natural and social science*. Vol. 52. Cambridge University Press, 1999.

Morrison, Margaret. "Models, measurement and computer simulation: the changing face of experimentation." *Philosophical Studies* 143.1 (2009): 33-57.

Morrison, Margaret. *Reconstructing Reality*. (2015)

Naylor, Thomas H., J. L. Balintfy, Donald S. Burdick, and Kong Chu. "Computer Simulation Techniques, John Wiley & Sons." *Inc., New York, New York* (1966).

Nola, Robert, and Howard Sankey. *Theories of scientific method: an introduction*. Routledge, 2014.

Oreskes, Naomi. "The scientific consensus on climate change." *Science* 306, no. 5702 (2004): 1686-1686.

Oreskes, N., Shrader-Frechette, K. and Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, *263*(5147), pp.641-646.

Pryor, James. "When warrant transmits." *Mind, meaning, and knowledge: Themes from the philosophy of Crispin Wright* (2012): 269-303.

Quine, Willard V. "The roots of reference." (1974).

Resnil, Michael. *Mathematics as a Science of Patterns* (1997) New York, Oxford University Press.

Ross, James F. "Testimonial evidence." In *Analysis and Metaphysics*, pp. 35-55. Springer Netherlands, 1975.

Ruphy, Stéphanie. "Limits to modeling: balancing ambition and outcome in astrophysics and cosmology." *Simulation & Gaming* 42, no. 2 (2011): 177-194.

Ruphy, Stéphanie. "Computer Simulations: A New Mode of Scientific Inquiry?." In *The Role of Technology in Science: Philosophical Perspectives*, pp. 131-148. Springer Netherlands, 2015.

Schickore, Jutta. "Scientific Instrument." In *Encyclopedia of Systems Biology*, pp. 1897-1901. Springer New York, 2013.

Symons, John. "Computational models of emergent properties." *Minds and Machines* 18, no. 4 (2008): 475-491.

Symons, John and Jack Horner. "Software intensive science." *Philosophy & Technology* 27, no. 3 (2014): 461-477.

Symons, John and Jack Horner "*On Some Limits to Model-Based Proof of Software Correctness*." In Powers, T. (ed.) *Philosophy and Computing: Essays in Epistemology, Philosophy of Mind, Logic, and Ethics.* Springer pp. 2017

Symons, John, and Ramón Alvarado. "Can we trust Big Data? Applying philosophy of science to software." *Big Data & Society* 3, no. 2 (2016): 2053951716664747.

Symons, John, and Fabio Boschetti. "How computational models predict the behavior of complex systems." *Foundations of Science* 18, no. 4 (2013): 809-821.

Norton, Stephen, and Frederick Suppe. "Why atmospheric modeling is good science." *Changing the atmosphere: Expert knowledge and environmental governance* (2001): 67-105.

Taddeo, Mariarosaria, and Luciano Floridi. "The case for e-trust." *Ethics and Information Technology* 13, no. 1 (2011): 1-3.

Tymoczko, Thomas. “The Four-Color Problem and its Philosophical Significance.” *Journal of Philosophy* vol. 76 (p57-82)

Vallor, S. (2017). AI and the Automation of Wisdom. In: Thomas Powers (Ed.) *Philosophy and Computing: Essays in Epistemology, Philosophy of Mind, Logic, and Ethics*. Philosophical Studies Series, vol 128. (pp. 161-178). Springer

Ware, Willis H. *The history and development of the electronic computer project at the Institute for Advanced Study*. Rand Corporation, 1953.

Williams, Michael. "Dretske on epistemic entitlement." *Philosophy and Phenomenological Research* 60, no. 3 (2000): 607-612.

Wright, Crispin, and Martin Davies. “On Epistemic Entitlement.” *Proceedings of the Aristotelian Society, Supplementary Volumes*, vol. 78, 2004, pp. 167–245. www.jstor.org/stable/4106950.

Weisberg, Michael. *Simulation and similarity: Using models to understand the world*. Oxford University Press, 2012.

Wiener, Norbert. *The human use of human beings: Cybernetics and society*. No. 320. Da capo press, 1954.

Wiener, Norbert. "Cybernétique et société (1950)." *Paris: Union Générale d'Editions* (1971).

Wikipedia contributors, "Patch (computing)," *Wikipedia, The Free Encyclopedia,* https://en.wikipedia.org/w/index.php?title=Patch_(computing)&oldid=784124212 (accessed June 23, 2017).

Winsberg, Eric. "Simulated experiments: Methodology for a virtual world." *Philosophy of science* 70.1 (2003): 105-125.

Winsberg, Eric. *Science in the age of computer simulation*. University of Chicago Press, 2010.

Winsberg, Eric, (2015) "Computer Simulations in Science", *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2015/entries/simulations-science/>.