Science, Technology, & Human Values
37(1) 64-87

DOI: 10.1177/0162243910385417
http://sthv.sagepub.com
SAGE

Creating Convincing Simulations in Astrophysics

Mikaela Sundberg[1]

Abstract
Numerical simulations have come to be widely used in scientific work. Like experiments, simulations generate large quantities of numbers (output data) that require analysis and constant concern with uncertainty and error. How do simulationists convince themselves, and others, about the credibility of output? The present analysis reconstructs the perspectives related to performing numerical simulations, in general, and the situations in which simulationists deal with uncertain output, in particular. Starting from a distinction between idealized and realistic simulations, the paper presents the principal methods of evaluation in relation to these practices and how different audiences expect different methods. One major challenge in interpreting output data is to distinguish between "real" and "numerical" effects. Within the practice of idealized simulations, simulationists hold the underlying model accountable for results that manifest "real" effects, but because "numerical" and "real" effects cannot be distinguished on the basis of what they derive from, attempted causal explanations are rather justifications for their conclusions. At the same time, simulationists' explanations are part and parcel of their contradictory perspectives, according to which

[1] Department of Sociology, Stockholm University

Corresponding Author:
Mikaela Sundberg, Universitetsvägen 10B, Stockholm 10691, Sweden
Email: mikaela.sundberg@sociology.su.se

they believe in simulations largely due to the underlying model, while painfully recognizing everything they have to add to make computations doable on the basis of this model.

Keywords
astrophysics, numerical simulation, scientific practice, uncertainty, credibility

Computer simulations have become important components in the machineries of contemporary scientific knowledge production, particularly in the natural sciences where traditional experiments cannot be carried out (cf. Knorr Cetina 1999).[1] The temporal and spatial scales of phenomena such as galaxy formation, ocean currents, or climate change render them impossible to bring into and explore inside a laboratory, but as mathematical representations they can be dealt with using computers. Numerical simulations are based on transformations of mathematical models into algorithms, which are then translated into computer code (Winsberg 1999). In the physics-based sciences, computer simulations tend to reduce the full complexity of the phenomena under study to a small number of physical laws. It is the related equations that define the dynamics of the system, but ad hoc expressions and/or adjustments are often included for computational reasons.[2] The simulation is the execution of a program that performs enormous numbers of calculations.

Numerical simulation activities share features with traditional experimental work (see, e.g., Dowling 1999; Winsberg 2008), and simulationists, therefore, face some problems similar to those faced by experimentalists.[3] The particular similitude as well as point of departure for the present article is the fact that numerical simulations result in large quantities of numbers (output data) that require analysis and constant concern about uncertainty and error (see, e.g., Winsberg 2003; Lahsen 2005). Expectations—based on previous conceptions developed from the scientific literature and individual background and training—are important for understanding the reception of the (output) data (see, e.g., Collins 1985). It is nevertheless unclear what the correct outcome of an innovative "numerical experiment" or a traditional wet-lab experiment is and how to interpret this outcome (cf. Lynch 1985; Turkle 2009, 51).

In her analysis of the levels of uncertainty/certainty surrounding climate models in relation to various practitioners (developers, users) and audiences (empirical meteorologists, etc.), Lahsen (2005) shows how climate modelers tend to trust their own simulations. However, considering the fact that modelers seek to predict the future development of a complex system, it must certainly occur that the plausibility of results is sometimes difficult to determine. If simulationists do not automatically trust output data owing to a belief in the physical basis of the theoretical model (cf. Sundberg 2006; Lenhard 2007), it remains unknown *how* they actually proceed to establish that the outcome of their simulations is trustworthy.

On the basis of a case study of practices surrounding numerical simulations in astrophysics, the overall aim of the current article is to reconstruct the perspectives created in relation to performing numerical simulations, in general, and the situations in which simulationists have to deal with uncertain output, in particular. Astrophysicists are frequently unsure as to how to interpret the output of their simulations.[4] Using expressions like "all the time" or "many, many times," these scientists relate how their output is often unexpected and how difficult it is to determine whether it should be regarded reasonable. Data in some form, produced by measurements or calculations, are the basis for "sequences of practice" (cf. Amann and Knorr Cetina 1990), and the particular sequence of practice around which the current article revolves is how simulationists convince themselves, and others, about the credibility of the output. One core issue related to this is how simulationists speak of "effects" in relation to their data—just like experimentalists do (cf. Hacking 1983, 224ff). The manifestation of "real" or "physical" effects is preferable, whereas "numerical effects" correspond to what is referred to as artifacts in traditional experiments. Determining how to conceive of and distinguish between these two different types of "effects" is a complicated task, not only in terms of whether they look reasonable as such, but also in terms of what their causes are. In their discussion of what novel philosophical issues simulations raise, Frigg and Reiss (2008) suggest that one of the few things that are new compared to analytical models is that when things go wrong with a simulation model, it is unclear where one should put the blame. If things go well, it is unclear why.[5] This is because the computational process leading from abstract model to output is opaque (Humphreys 2008), and the ingredients of a model cannot be distinguished in its output (Boumans 1999). On the basis of these statements, the current paper develops the analysis of numerical simulations by arguing that "numerical" and "real" effects in fact cannot be distinguished on the basis of what they derive from. This conclusion is based on a discussion of how "real" and "numerical" effects are understood within two simulation practices that involve different preferences for how to evaluate results.

Performing Numerical Simulations: Practices and Perspectives

A perspective is an ordered view of one's world, an organized conception of what is taken for granted, plausible and possible, but also a coordinated set of actions (Becker et al. 1961, 34; see also Mead 1932). Perspectives develop in relation to problematic situations in which choices are required. If the situation occurs frequently, the perspective becomes an established way of dealing with the world (Becker et al. 1961, 34ff). Dealing with questionable output is a complicated part of the everyday lives of scientists who work with numerical simulations. As such, it is a situation that generates particular perspectives.

Perspectives do not only develop over time as an effect of situations but also in relation to a reference group (cf. Shibutani 1955). People look at the world around them as well as at themselves from the perspective of this reference group: "The socialized person . . . sets the same standards of conduct for himself as he sets for others, and he judges himself in the same terms . . . . [H]is perspective always takes into account the expectations of others" (Shibutani 1955, 564, cf. Mead 1934). Applied to the topic of the current paper, there are two significant implications of this quote. First, it implies that socialized astrophysicists use the same standards to convince themselves about *their own* numerical simulations as they would use in judging simulations produced by *others*. However, previous research shows that uncertainty is unevenly distributed across the audiences of scientific results (cf. Collins 1985; Mackenzie 1990; Lahsen 2005) and that scientists recognize and anticipate the expectations of peers (cf. Knorr Cetina 1981; Sundberg 2006). This leads to the proposition that simulationists rely on different (rhetorical) resources depending on their audience, that is depending on whether they seek to convince themselves, other simulationists, or astronomers in general (cf. Kennefick 2000, 26). The aim of my analysis is to investigate what methods simulationists use to convince themselves and others, and what expectations they respond to.[6] Whether they actually convince anyone or create certainty is a different question. Second, although the article does analyze numerical simulation practices, this behavior is intertwined with how astrophysicists think simulations *should* be practiced, that is their rules or standards of conduct. Scholarly discussions of validation in the literature and textbooks on simulation and modeling provide insights into these standards exclusively but to enable reconstruction of the simulationists' perspectives, qualitative analysis of interviews and observations is required (cf. Becker et al. 1961). As a consequence of this type of material, the current paper does not explicitly distinguish between how numerical simulations are practiced and in what ways this may differ from how simulationists think they should be practiced (i.e., what is stated in handbooks on simulations, etc.).

Material and Analysis

For this study, eleven astrophysicists were interviewed, selected because all of them were, or had been, working with numerical simulations.[7] These interview accounts have been interpreted as descriptions of social practice or as part and parcel of social practices (as perspectives linked to practices; Gubrium and Holstein 1997).[8] Observations gave the opportunity to listen in on presentations as well as both formal and informal discussions.[9] I attended three days of a symposium, including twenty presentations, and during three simulation code user meetings, I attended twenty-two "science" presentations, and several "code" discussions that were less strictly organized. I also participated in lunches, coffee breaks, and so on, during these different gatherings. Additional occasions of observation include eight department colloquia, two dissertation defenses, and one interdisciplinary doctoral student course on multiscale modeling and simulation (six days). To follow-up interesting issues that arose in the interviews, presentations, or discussions, I have had personal e-mail communication with a number of the astrophysicists.

Observing interaction between the astrophysicists serves to verify what the interviews have revealed regarding how simulationists work. It is also an important source of data as concerns establishing the collective character of perspectives. For example, discussions after presentations and questions from the audience reveal expectations and perspectives as well as give insights into what simulationists—and astronomers more generally—disagree about and discuss. All gatherings I observed were intended for a scientific audience. A few simulationists from closely related disciplines attended the code user meetings. On all other occasions except for the doctoral course, the audience consisted of astronomers, mostly simulationists, but always at least a few observers.[10] The doctoral course was the only fieldwork that took place in an education context rather than a research context.[11]

My analysis focuses on astrophysicists' practices of performing simulations and their methods of evaluating output, rather than on the details of their work activities.[12] It is based, first, on the simulationists' descriptions of their work, both in general terms and regarding concrete occasions when they have struggled with the output. Second, it is based on my observations of presentations and discussions, where I draw upon excerpts from my field notes describing occasions when credibility of output has been debated. Owing to my interest in forms of knowledge production rather than the content of knowledge, I have not followed or analyzed how different simulationists work with any particular research problem. The informants I have interviewed and the presenters I have listened to work on a variety of research topics, ranging from objects such as planets and disks, stars (including the Sun) and galaxies, to the large-scale structure of the universe (cosmology), and they investigate various processes and forces, such as turbulence, convection, magnetic fields, and radiation transfer.

Idealized and Realistic Simulations

Previous analysis has shown that there are different approaches to performing numerical simulations in the natural sciences (see, e.g., Kennefick 2000; Sundberg 2009; Shackley 2001). Throughout this article, I distinguish between idealized and realistic simulations (cf. Sundberg 2005, 138ff.). These two dominant perspectives on the use of simulation are in one sense actor-defined, as the simulationists recognize and acknowledge the differences themselves, but my data-driven analysis looks at the content of these perspectives, especially with regard to uncertain output and evaluation methods (cf. Becker et al. 1961, 37). For pedagogical reasons, I will present a brief introduction to these perspectives below.

The aim of an idealized simulation is to understand the relative importance of different physical processes for a particular phenomenon or its underlying mechanisms. The simulation excludes descriptions of other real-world processes that are not considered of major importance to the particular problem by "switching" off specific functions (such as the representation of dust coagulation, particles, etc.) in a code or using a code with fewer initial process descriptions. The simulationists' argument for this practice is that additional process descriptions interfere with interpretation and understanding of what happens during the simulation.

Realistic simulations include more detailed process descriptions that interact in the calculation as a means to better reflect the variability and complexity of real-world processes. The models are often three-dimensional rather than one- or two-dimensional (in space). Compared to idealized simulations, more empirical knowledge is required to set up a realistic simulation. For example, if the calculation is to start from more detailed initial conditions, one has to set a larger number of parameter values (cf. Kennefick 2000, 9). One astrophysicist working on stellar atmospheres explained that the reason for "dressing up the model" (whatever model it may be) with "all these details" is to enable comparison with observations.

For the simulationists, it appears as if the distinction between idealistic and realistic simulations (as empirical constructs) is related to what is included in the simulation code, whereas the current analysis shows how the crucial difference between idealistic and realistic simulations *as perspectives* (analytical constructs) concerns what the simulationists expect from the calculation. For example, it is not obvious what constitutes the processes of "major importance" that should be taken into account. There are conventions and commonly shared ideas regarding what is necessary to include—and what is not—when investigating specific research problems, and these conventions change over time, in part as an effect of debates among scientists. The question of taking magnetic fields into account in simulations is a case in point. Observations and interviews indicate that this issue has been debated in astrophysics for a couple of decades or so. The representation of magnetic fields in simulation codes is typically included through the so-called magnetohydrodynamics (MHD) equations. Some simulation codes include these equations (as an option), whereas others do not. The significance of taking one or the other approach was a question that kept popping up in dialogues after presentations. Choosing what process descriptions to include is also scale dependent. Although the representation of phenomena of different scales in a simulation creates computational problems, the computation is sometimes not considered physically meaningful if fine-scale information is ignored. It is the research problems a simulation code is designed to be applied to that determine the level of detail and sophistication of process descriptions and approximations. For example, the vast spatial scale cosmological simulations have to account for, combined with limited computational power, constrain them to coarser resolution and simple approximations on smaller-scale processes. According to one astrophysicist working on radiation transport on cosmological scales, this causes simulationists who try to simulate processes of the Sun, referred to as "solar people," to consider some of the approximations made in cosmology as "crap."

The selection and representation of processes is primarily an aspect of constructing simulation codes. Thus, the inner construction of simulation codes will only be a topic for the current analysis to the extent that it is raised as an issue in relation to interpretation of the simulations' outcomes. Importantly, there are a number of publicly available codes that astrophysicists may download and use for their research. This means that astrophysicists "run" either simulation codes they have developed themselves *or* codes developed by others. Different patterns of developer/user roles are not discussed here (cf. Sundberg 2010), nor is how this affects levels of uncertainty (cf. Lahsen 2005). Nevertheless, distinguishing among types of simulationists performing idealized and realistic simulations, not only among simulationists and observers, adds another layer to the analysis of how simulationists convince various audiences. However, this analysis does not extend to audiences external to the discipline. This is potentially of more significance for understanding the internal dynamics in fields such as climate modeling, in that climate modelers have to face expectations not only from their scientific colleagues, but, to a much larger extent than astrophysicists, also from the public and the politicians.

Exploring Output through Replications

One simulationist I interviewed performed simulations of the solar dynamo, which is the physical process that generates the Sun's magnetic field, as well as of magnetic fields in astrophysical environments more generally. He told me about common problems that occur when working with simulation codes and about how some output was easily recognized as "rubbish because you see all these kind of like spikes and everything. That's the first thing you know, if you get this spiky kind of things, then you are doing something wrong." For this simulationist, there is no question as to whether an output profile that exhibits "all these kind of spikes" is problematic. For astrophysicists in general, rapid oscillations in parameter values—manifested as "spikes" in diagrams and plots—are signs that a simulation generates unrealistic instabilities. Implausible values such as, for example negative energy, are also clear signs of "error." Simulationists are preoccupied with debugging and finding errors in codes (see Kennefick 2000), but error management is analytically distinguishable as a sequence of practices that starts when simulationists distrust output data and begin looking for errors (bugs) in the simulation code. This differs from when they have doubts, but still have hopes, and are trying to determine whether the data are credible. As this commonly occurs, it is a relevant subject for the analysis of perspectives.

There are many occasions on which it is uncertain whether one's output is wrong or reasonable. Like in traditional wet-lab experiments, innovative numerical experiments do not have clear answers (cf. Lynch 1985). Astrophysicists need to develop an appreciation of the types of errors likely to emerge under different circumstances and of how to diagnose problems (cf. Collins 1985). For example, the position in the computational domain where a given feature (parameter values representing, e.g., turbulence, convection, etc.) occurs may indicate whether the feature is plausible. One doctoral student talked about one of his simulations of planet formation in accretion disks and mentioned that he was not sure what to make out of it: "It's a weird thing that this vortex here shows up in the boundary of the buffer zone right. I think that it's a hint that it is something numerical." This account mentions that the feature (a vortex) is likely to be "numerical." The prevalence of "numerical" effects is difficult to determine, and simulationists have to learn what a "hint" of a "numerical" effect looks like. To create more certainty, they use different test methods, several of which aim at testing the stability of output, or in other words, at replicating output with slightly different settings of parameter values, boundary conditions, and so on (cf. Collins 1985).[13]

One of the methods that simulationists use is to change the resolution of the calculation by increasing the distance between grid points or the number of particles. If a feature changes or disappears with increased resolution, it is taken as an indication that it is a "numerical" effect. On a few occasions, I observed speakers presenting the results of simulations with different resolution as a means to support their claim that these simulations exhibit some structural pattern or particular effect. If such findings are not presented, someone in the audience will typically ask the speaker whether a feature is resolution-dependent, despite the fact that it is often considered too time-consuming to reproduce simulations at different resolutions because they are such computer-resource-intense activities.[14]

Another procedure to increase the details of a simulation and thereby test the stability of the output is to increase the number of spatial dimensions taken into account in the simulation. Simulation codes often exist in different versions, including one-, two- and three-dimensional versions. When speakers present one- or two-dimensional simulations during presentations, it often happens that someone in the audience will ask what the result looked like when, or would look like if, the simulation was run in, for example, "3d" instead of "2d."

Both of the two methods described above imply a shift from idealized to more realistic simulations, in the sense that the world is ultimately three-dimensional in space and phenomena of many scales interact. Including more of this complexity entails more realism and less idealism. Yet all these methods are based on small adjustments to test the stability of the simulation, and do not imply moving from more to less idealized settings.

A third way to check whether "numerical" effects are present is to change some input parameter value and then run the computer simulation again. One astrophysicist emphasized solidity in relation to this practice: "If a feature of a simulation is not robust to the choice of free parameters, it is evidence that this feature is artificial, not really the way nature behaves." Importantly, this account suggests that lack of robustness is "evidence" that output is *not* credible, rather than evidence that it is.

Speakers presenting simulations as well as commentaries on them sometimes mention that even if features in simulation output remain despite changes such as those mentioned above, they are still not necessarily "physical" or "real" effects. Because both speakers and audiences express this view, I take the comment as a caveat, signifying awareness, rather than as devastating critique. Why? First, because it is unlikely that speakers wish to inject serious doubt about the simulations they present themselves. Second, and more importantly, is it unlikely that speakers will present results they do not believe in themselves. Summarizing thus far, all of the above methods test persistence primarily to identify possible sources of *failure* (cf. Roth 2009; Lynch 1985). One astrophysicist made the general remark that "[t]he pitfall is to interpret something as real without ruling out numerical explanations," in other words, to claim that something is a real effect before challenging the stability of output with different measures that would indicate that numerical issues play a role. Drawing upon the distinction between data, which is *recognized* in the laboratory, and the evidence *published* in papers clarifies this point further (Amann and Knorr Cetina 1990). Stability is sufficient for the producer of a simulation to recognize and trust its results, but it is insufficient as evidence to convince others. What I refer to as replicative methods, therefore, generate a form of negative knowledge (cf. Knorr Cetina 1999), rather than providing positive support that output is believable. The following sections accentuate the distinction between idealized and realistic simulations by presenting two such support methods and examining how different practitioners and expectations are related to them.

Connecting to the Underlying Model

One way of trying to understand and possibly support the feasibility of output is to look into the simulation code. This takes place by exploring and unfolding the code, in particular its underlying equations (cf. Knorr Cetina 1999, 71f; 197f.). This is typical of the practice of idealized simulations, which sometimes rely on more traditional, analytical modeling. Comparing parts of the approximate solution of a numerical model with a solution of a similar analytical model is considered a solid way of checking that the output is reliable (see also Kennefick 2000, 25). However, only "simple" problems can be formulated with solvable analytical models, and many mathematical models are too complex to allow for analytic evaluation. This is, of course, part of the reason why numerical rather than analytical methods are used to approach more complex problem formulations in the first place. It may seem self-evident that there is a link between the equations the calculations of the simulation code are based on and the outcome of these calculations, but because simulation codes are generally constructed to deal with equations that are analytically intractable, it is much harder to determine exactly how equations and results relate to each other than a layperson might think. One way to try to get around the obdurate character of simulation output is to strip away from the mathematical model everything that underlies the simulation code except for what is taken as its most essential mechanism. Then this part is investigated using analytical methods. One astrophysicist working on star winds told me about what she did to determine how to interpret unexpected output: "You do something you call toy models so you take the equations [in the simulation code] and reduce them ... throw away almost everything except the part that you suspect causes the effect [in the simulation], so to speak. And ... and then you use a sheet of paper so to speak to see what it can be." Tackling the problem with analytical methods (use a sheet of paper) is an approach to uncertain outcomes, the aim of which is to reach an understanding of what happens inside the simulation code when it calculates. Based on the result of this, simulationists hope to determine whether and in what sense features in output can be considered to derive from the underlying model.

What is the underlying cause of "real" effects? One astrophysicist who simulates interactions between gas and radiation in various astrophysical environments repeatedly referred to "real" effects in different simulations. When I asked what he considered the underlying causes of these effects, he replied:

> It is in the equations. Hopefully you'll be able to *argue* why this actually happen in a more descriptive way or may even use the equations to show what this is, what you can expect to happen. That's ... that's the typical thing, you try to, if you see something that is unexpected, but you can show from simpler arguments, using the equations, that this actually can happen, this is the strongest *evidence* that this is a physical effect and not a numerical effect.

This quote suggests that the establishment of a link between the equations and the result is regarded one of the best ways to establish that unexpected results are credible. This simulationist holds it up as the most convincing way to "argue" that "something unexpected" is "physical,, rather than "numerical." When this support is lacking, it may be asked for. The occasion described below provides an example of when simulationists are expected to explain their results in relation to the underlying equations (see also Kennefick 2000, 16ff). During a thesis defense in astrophysics, one of the committee members, a simulationist himself, doubted the conclusion that the generation of vortices in a particular simulation was "real" rather than "numerical." The committee member questioned the doctoral candidate by going through the equations of the mathematical model in a stepwise manner to pin down exactly which term in one of the equations the unexpected feature derived from. The doctoral candidate was unable to provide an answer to this. I interpreted the questioning as an exercise aimed at convincing the committee member that the candidate claim was legitimate, but the fact that the candidate was unable to argue for what the effect derived from indicates that *he* believed in the result of the simulation *anyway*. In the context of a discussion of experimentalists' skepticism toward meteorological simulations, it was claimed that the way results are discussed and questioned depends on whether or the dissenter is from the same group of practitioners, for example, whether she or he is a simulationist or an observer (Sundberg 2006, 61). Thus, reference to the "physics of the model" may perhaps end a face-to-face dialogue with a critical observer but certainly does not silence critical fellow simulationists, who do not hesitate to dig deeper into the content of the underlying model. However, if idealized and realistic simulations have different aims, simulationists of different kinds do not ask for the same type of support. In the last example, the committee member was someone who performs idealized simulations. This is illustrated by what he said about his own research on solar dynamos in an earlier interview.

> We are trying to move away from these boxes ... which help us to learn the equivalence between the equations, simply the mathematic equations and what is seen in simulations. But you actually should be moving closer to real solar geometry, and we are working on this and we are trying to turn spherical geometry into the code.

This quote shows that investigating the equivalence between the underlying equation and what is seen in the output of the numerical simulations is not *only* an exercise to convince others or a matter of defense. It is part and parcel of idealized simulations as a practice and perspective. Second, this astrophysicist refers to a numerical domain that has the shape of a square box, whereas the real Sun is a spherical object. A simplified domain (a box or a so-called Cartesian domain) is sufficient to investigate the connection between equations and output. Shifting to spherical geometry shifts the process toward trying to achieve greater realism (in this example, in the field of solar dynamo simulations). The account shows how simulationists have the ambition to develop ("move away") from idealized toward more realistic simulations and the time gradient that characterizes the relation between the two perspectives.

Convincing and Conventional Use of Observations

Framing means to consider objects or pieces of information in light of other such components that serve to check, control, extend, or compensate the former (Knorr Cetina 1999, 72 ff). In relation to simulations, observations constitute one such piece of information. This is in line with the conventional view that measurements are capable of guiding scientists in determining what constitutes reasonable results (Knorr Cetina 1999, 52). It is a common view in numerical simulations (particle physics, geophysics) that the realism of a model is tested through comparison with observations (see, e.g., Merz 2006, 168f; Oreskes, Shrader-Freshette, and Belitz 1994). Astrophysicists speak of observations as principally synonyms with "truth," "reality," or "Nature," but scarcity of observations commonly justifies the need for numerical simulations in the first place, and observations cannot always be used as guidance.

It is mostly through telescopes that detect electromagnetic radiation that astronomers are provided with observations. However, temporal and spatial scales create problems for observational techniques. Some phenomena are extremely rapid, and processes such as supernovae are too fast to capture with current observational methods. On galactic scales, nothing happens on human timescales, and there are basically no observations available at all from "the Dark Ages" —when the Universe was between 380,000 years and 400 million years old—because there were no stars. This creates different opportunities for research fields within astrophysics. For example, simulationists interested in solar physics may have access to a great deal more useful observations than do simulationists working on cosmological problems.[15]

Simulationists performing realistic simulations compare output with observations both when trying to understand and evaluate output for themselves, as well as when presenting output to others at seminars or conferences. We could expect there to be different views on what constitutes agreement, especially among simulationists and observers (cf. Sundberg 2006), but these interpretative practices are not explored here. Instead I discuss comparison to observations to highlight the principal role this plays in realistic simulations, the expectation that idealistic simulations will adhere to the standard of comparison to observations, and finally, the distinction between "numerical" and "real" effects in relation to realistic simulations and problematic ingredients in simulation codes.

One primary point is that, for realistic simulations, comparison of simulations and observations is "the moment of truth," as one astrophysicist working on models of stellar atmospheres puts it. He repeated this later on in the interview and explained:

> The moment of truth is when you are standing there and compare with observations. Okay, but this feeling can be of different strength. I think I belong to (...) the type of researcher who feels strongly: [you have to] "prove!" With your models. Whereas some others are more fascinated by the models and less of this confrontation with reality.

This quote is illustrative of the difference between idealized and realistic simulations. The view of "proof" as related to the comparison of output with observations differs from the view of evidence presented earlier, where connections between output and the underlying model are essential. Importantly, comparison seems to be something this astrophysicist has a strong "feeling" about in relation to convincing *himself*. He does not mention it in relation to convincing others, but only so as to *differentiate* himself from others. Others "are more fascinated by the models"—presumably those who perform idealized simulations and look into the equations underlying simulation codes. For this latter type of simulationist, enabling output to be placed next to observations is not a primary goal. Part of the audience to whom simulations are presented may nevertheless expect comparisons, regardless of whether this is the simulationists' intention. For example, during one PhD defense, the doctoral candidate presented what he referred to as a "theoretical study," using simulations to investigate, among other things, mass loss from stars and stellar winds. The discussant—an observer—repeatedly asked for observations that would strengthen the credibility of the simulations from his perspective.

A Swedish dissertation in the natural sciences consists of a collection of, preferably, published articles, but sometimes also submitted or about-to-be-submitted manuscripts. Publication constitutes a context for exposing simulation results that is different from the context of presentations. For the current analysis, the interesting aspect of this part of the research process is that simulationists' accounts of how they relate to observations in publications illustrate the convention of addressing observations rhetorically—even when there is no intention to compare to them in practice. One astrophysicist had conducted highly "idealized simulations" of magnetic fields and explained why he had included a sentence on the importance of making "realistic simulations" in the article presenting this work.[16]

> If you would publish in a mathematical journal or a theoretical physics journal, then you would perhaps not write anything about observations. But because this is *Astronomy and Astrophysics*, if you don't have a little bit of connection to observations, then you may be afraid that they will think this is not astronomy.[17]

As socialized members of astronomy, simulationists (have to) care about what the audience (editors, reviewers, and readers, in short "they") of *Astronomy and Astrophysics* might think. This implies that dropping a line about "observations" is simply a conventional exercise for simulationists running idealized simulations, and that they do not need observations to be convinced. As it were, articles presenting results from realistic simulations even have titles such as "Models meet observations," illustrating how the connection plays an important role. More generally, scientific journals function as common communication channels and the choice of journal is, therefore, a choice of which expectations—which reference group—one acknowledges. In Mead's terminology, it is a question of which generalized other's role is to be taken (cf. Shibutani 1955, 568). If the audience were non-astronomers, for example readers of a "theoretical physics journal," the link to observations would not be as necessary.

Returning to realistic simulations, what does the related perspective say about "real" and "numerical" effects in relation to expectations for results? One astrophysicist conducting research on planet formation said he could not know for sure whether what he saw in the output of his simulations was "numerical or real," because he had no guiding observations. This implies that it is how output corresponds to observations, rather than how output can be connected to the underlying model that determines whether one might have a justified belief in the visible effect. If the aim of realistic simulations is to achieve agreement with observations, it is not surprising that observations are required to settle the issue of what is good or bad. Interestingly, however, this view signifies a completely different understanding of "real" effect compared to what was discussed above. The difference is that the notion of "real" effect does not seem to rely on what the effect *derives* from, but only relate to whether the effect is plausible in relation to what could be expected from observations.

It is well known that simulation output may correspond well with data (or generate reasonable results more generally) despite there being erroneous formulations inside the simulation code.[18] I will now briefly discuss bugs and so-called artificial terms. Bugs such as misspellings, redundant characters, wrong numbers, and so on, are common, but they do not always seem to influence the outcome in disadvantageous ways (and vice versa). According to one astrophysicist who claims to perform realistic simulations on stellar atmospheres, "the big problem" with bugs is that it is only when results are considered evidently poor that simulationists put energy into debugging their codes: "If you manage to come close to what you think is reasonable, then you often stop the efforts to debug," he said (cf. Kennefick 2000, 26). In addition, one of the reasons for a simulationist to use a code developed by someone else is exactly to avoid debugging, because it is expected that others (developers) will take care of it (cf. Sundberg 2010). Many simulation codes also contain deliberately fictitious "artificial" terms. These are added to the original equations (and therefore included in the code) to keep simulations on plausible tracks (Winsberg 2001, 2006, 2008). One astrophysicist made an analogy between bank transactions and simulations to illustrate the effect of including a term representing artificial dissipation: Generating reasonable transaction values requires that a bank leak money like water pores from an open tap, in spite of the knowledge that this is definitely not the case in real life. Do simulationists who apply these terms actually trust them on the basis of their background knowledge, as philosopher Winsberg (2008) claims? Or is it rather that there are few other alternatives *if* simulationists want to generate what they consider plausible output, and especially if plausibility requires comparison to observations? The representation of shock waves is a classic problem in astrophysical simulations and provides an illustrative example of how perspectives on these types of "artificial" terms differ among those performing idealized and those performing realistic simulations (cf. Sundberg 2009). During a simulation code user meeting, participants discussed how to handle the issue of shock waves when performing simulations with the code. With a critical tone of voice, one of the founders of the simulation code asked the audience if anyone was using "artificial viscosity." One of the other developers answered that he was and he defended it by saying that it was needed to solve "real physical problems," as opposed to solving problems under "very idealized conditions," which is what the developer considers the founder to be doing in his simulations. After some other users had reported on their inclusion of such terms too, the founder of the simulation code concluded the discussion by recommending that users be cautious with "unphysical things" in the code. These terms are certainly used by simulationists for pragmatic reasons, but they still cause principle debates.

Concluding Remark

To understand modern science, we need to direct our analytical gaze toward numerical simulations. Reconstructing the simulationists' own perspectives is a way to approach this aim using sociological methods. On the basis of astrophysicists' accounts of how they handle the everyday situation of dealing with unexpected output, the current article describes the activities involved in this work and their relation to practice-based perspectives. Although the generality of these perspectives remains to be further explored, the analysis nevertheless contributes to our understanding of simulation-based knowledge production by revealing *how* simulationists in astronomy draw upon internal (the model) as well as external (observations) resources to generate support that will convince them and others about the credibility of the output.

We can also summarize the principal patterns of the three types of resources for creating certainty, three types of scientists, and their expectations with regard to support. Replicative tests are important for both types of simulation practices, but particularly simulationists performing idealized simulations seem to trust their output if it is stabile throughout various conditions. To convince others performing idealized simulations, they connect to equations. To convince the rest of their reference group, they relate rhetorically to observations but do not necessarily compare to observations. Simulationists performing realistic simulations evaluate their own *and* others' simulations by comparing to observations. Observers are likely to expect simulations to be compared to observations, no matter what type of simulation has been performed. Although realistic simulations are always potentially comparable to observations, a tentative conclusion is that idealized simulation in fact *means* it is impossible to compare. Claiming to perform such simulations is, therefore, also an excuse for avoiding comparisons.

It should now be obvious to the reader that one key aspect of interpreting simulation output is for astrophysicists to distinguish between "real" and "numerical" effects. What are simulationists referring to when they use such expressions? It implicitly appears as if estimations, uncertain parameter values, coarse resolutions, time stepping procedures, and so on, are considered the underlying causes of untrustworthy results that exhibit "numerical" effects. Yet these messy ingredients are certainly not only distorting output. Numerical simulations would be impossible to pursue without the ingredients added for purely computational reasons, including the controversial "artificial terms" (see, e.g., Edwards 2001; Lenhard 2007; Winsberg 2006). There are several tricks through which simulationists achieve what they consider reasonable results, and these can be achieved for the "wrong" reasons.[19] In addition, simulation codes with bugs generate acceptable output as well. Thus, simulationists (performing idealized simulations) point at certain terms to argue for why a feature found in the output is "real," *at the same time* as they, or at least their fellow realistic simulationists, deliberately include "artificial" terms (which, in a sense, generate "numerical" effects by definition) to avoid output features they believe are false. At least within the practice of idealized simulations, simulationists hold the underlying model *accountable* for results when they consider them worth believing in—when they present "real" effects (cf. Winsberg 2006, 7). Somewhat analogous to Fine's (2007, 125) description of how meteorologists create forecasts intuitively and then use theory to justify their forecasts in a post hoc fashion, astrophysicists' attempted causal explanations can be seen as *justifications* for their conclusions, or as rhetorical devices aimed at bolstering credibility. Notably, these simulationists' explanations are *also* part and parcel of their contradictory perspectives, according to which they believe in simulations largely due to the underlying model, while at the same time painfully recognizing everything they have to add to make computations doable on the basis of this model.

Frigg and Reiss (2008) compare simulations to experiments and argue that the relationship between what happens in a "numerical laboratory" and in the target system is equivalent to the relation between what happens in a wet-laboratory experiment and in reality, but we can twist this epistemological question into a sociological framework and ask ourselves whether simulationists' *perspectives* on what is "real" might differ from, for example, those of wet-laboratory experimentalists. This opens the door to discussions of intriguing questions. What position on epistemological core issues, such as realism/antirealism and causality, is expressed in the *practices* of simulationists (cf. Hacking 1983)? If taking for granted that work in the natural sciences is conducted along a realist vein is inadequate, we should look further into what epistemological standpoints are actually being performed among these relatively new types of scientists.

Authors' Note

Previous versions of this article have been presented at the Swiss STS Meeting ScienceFutures, Zürich 2008 and the biennial conference for the Society for Philosophers of Scientific Practice, Minneapolis 2009.

Acknowledgments

I am extremely grateful to the astrophysicists I have interviewed, the organizers of the different types of astrophysics meetings I have attended, and those astrophysicists I have talked to during these occasions. I also wish to thank the anonymous reviewers for comments that improved the paper. For recommendations and comments on earlier drafts, I thank Ericka Johnson, Petra Jonvallen, Francis Lee, Per-Anders Lindén, Monica Nordvik, Árni Sverrisson, Lars Udhen, and Staffan Wennerholm.

Declaration of Conflicting Interests

The author(s) declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: This research has received generous financial support from the Swedish Science Council under contract 2006-1296 and contract 2007-1627.

Notes

1. I use numerical simulations and computer simulations interchangeably. Simulation is used for the sake of brevity.
2. If the differential equations are nonlinear, they cannot be solved because it is impossible to write down closed-form equations that represent a unique solution. This means that the equations have to be approached using numerical methods instead. In the *Eulerian* approach, the continuous equations are made discrete and transformed into step-by-step difference equations for which solutions can be approximated, the computational domain is made discrete into finite cells and the equations are solved on a grid. The Eulerian approach focuses on *specific*

locations in space, whereas the basic idea of *Lagrangian* schemes is to *follow particles*—representing, for example, gas or water—as they move. In an Eulerian code, space is divided into a large number of points and the numerical solutions for the equations are calculated for each grid point. The smaller the distance between the grid points and the time step, the higher the resolution of the simulation code. In Lagrangian schemes, higher resolution corresponds to a larger number of particles. The resolution is high in high-density regions but low in low-density regions.

3. Although much of the philosophical literature on simulations addresses the theory–experiment split, it is a questionable starting point for sociological analysis. It is, therefore, not addressed here.
4. Strictly speaking, *astronomy* is the science of measuring the positions and characteristics of astronomical objects, and *astrophysics* is the application of physics to understand astronomy. Currently, the two terms are interchangeable, because all astronomers use physics. I use *astronomy* to refer to the discipline as a whole and *simulationist* and *observer* when there is a need to distinguish between the practitioners. More generally, I refer to astrophysicist as a synonym for simulationist.
5. Analytical modeling or methods refers to the writing of formulas and pen and paper calculations. The solution of an analytical model is an equation in the form of an algebraic expression.
6. Important contributions to philosophy of science discussions on the credentials of simulations include, for example, Frigg and Reiss (2008); Winsberg (1999); Oreskes, Shrader-Freshette, and Belitz (1994).
7. Of these interviewees, there were four doctoral students, one post-doctoral student, four research scientists, and two professors. Three of them worked at a research department in Denmark, the others at different astronomy departments in Sweden. Five interviews were conducted in English, the rest in Swedish and excerpts from the latter have been translated to English. All interviewees were granted anonymity.
8. Because perspectives develop during socialization, the status of doctoral students' accounts deserves a note. Despite the relatively few years of experience with numerical simulations that (most) doctoral students have, socialization of the doctoral students into the practice of simulation modeling involves adopting the perspectives of the researchers they are surrounded by. Some of their accounts can, therefore, be analyzed as expressions of similar, evolving perspectives as well.
9. Most of these gatherings were at the astronomy departments at Stockholm University or Uppsala University or at the Nordic Institute of Theoretical Physics (Nordita) in Stockholm. Nordita, Leiden Observatory in the Netherlands, and

Max Planck Institute for Astronomy, Heidelberg in Germany hosted the three simulation code user meetings.

10. Because of the focus on simulationists' perspectives, observers are only considered to the extent that they interact with the simulationists, who have to take their expectations into account (cf. Shibutani 1955). Three of the interviewees had also worked with observations.
11. In Sweden, thesis defenses are public. In the natural sciences, the doctoral candidate summarizes the thesis and she or he is then questioned by an external discussant, often a scientist from abroad, invited on the basis of her or his expertise in the particular field. There is also an examination committee of three to five research scientists from Sweden. Because this case study of astrophysics is part of a larger comparative study, observations of dissertation defenses in meteorology and oceanography have also taken place. Adding to this my own experiences as a sociologist attending defenses in sociology leads to the conclusion that this Swedish tradition provides excellent opportunities to learn about what is typically debated in different fields, including how one is supposed to provide support for claims.
12. Numerical modeling is a collective enterprise, but the simulationists' everyday work involves sitting alone, looking at a computer screen. There is no correspondence to the laboratory shop floor and hence, it is difficult to follow interactively accomplished interpretations in everyday work. For a discussion on what types of material on modeling work are suitable for sociological analysis, see Sundberg (2005, 221f.).
13. It is said to be good practice to apply a different simulation code to the same problem and try to reproduce the results. Very few simulationists have the skills or consider it worthwhile to set up and apply a different code for that purpose. In intercomparison projects, simulationists with similar research interests join to formulate a simulation problem, apply their different codes to it, and then compare their results. For analysis of these coordinated efforts, which put simulationists in a different situation compared to the situation under study in the current article, see Sundberg (forthcoming).
14. One general remark is that many problems and limitations of simulations are blamed on coarse resolution—both in time and in space. This, in turn, is blamed on the shortage of computer resources, which implicitly places the responsibility for improved simulations in the hands of computer scientists, and beyond the control of simulationists. At the same time, coarser resolution leads to demand for more (detailed) descriptions of phenomena that affect the smaller-scale processes. This potentially creates problems of a different kind, for example, how to improve models of processes from a theoretical viewpoint, and is looked upon from this perspective, supposedly directing the spotlight of responsibility back on the simulationists.

15. The Sun is not only our closest star, but it directly affects life on Earth. There are applied interests tied to solar physics, which is likely to generate investments in observational capacities. This is not further discussed in the article.
16. The astrophysicist used the terms "realistic" and "idealistic" in the article.
17. *A & A* is one of the important journals in astronomy and astrophysics.
18. When observations (or something else) are considered correct, one can try to force the simulation code to generate a similar output, for example by adjusting parameter values. This is known as tuning, or calibration. It is a general concern in numerical simulations as soon as there is a "right" answer, but for the reasons given, this is rarely the case. Tuning is a prominent issue in meteorological research, largely due to the close relationship with weather forecasting, where tuning is accepted as a necessary means to produce better agreement with observations (Sundberg 2009). This is rarely discussed in astrophysics.
19. Compare philosophical discussions of the fact that more than one model configuration may produce the same output (nonuniqueness). For discussions of the relation between causes and effects in simulations from the standpoint of philosophy of science, see Küppers, Lenhard, and Shinn (2006, 11ff.); Winsberg (1999, 2001).

References

Amann, K., and K. Knorr Cetina. 1990. The fixation of (visual) evidence. In *Representation in scientific practice*, ed. M. Lynch and S. Woolgar, 86-121. London: MIT Press.

Becker, H. S., B. Geer, E. C. Hughes, and A. Strauss. 1961. *Boys in White: Student culture in medical school.* Chicago: Chicago University Press.

Boumans, M. 1999. Built-in-justification. In *Models as mediators: Perspectives on natural and social science*, ed. M. S. Morgan and M. Morrison, 66-96. Cambridge, UK: Cambridge University Press.

Collins, H. M. 1985. *Changing order. Replication and induction in scientific practice.* London: SAGE.

Dowling, D. 1999. Experimenting on theories. *Science in Context* 12:261-73.

Edwards, P. N. 2001. Representing the global atmosphere: Computer models, data, and knowledge about climate change. In *Changing the atmosphere: Expert knowledge and environmental governance*, ed. C. A. Miller and P. N. Edwards, 31-66. Cambridge, MA: The MIT Press.

Fine, G. A. 2007. *Authors of the storm. Meteorologists and the culture of prediction.* Chicago: University of Chicago Press.

Frigg, R., and J. Reiss. 2008. The philosophy of simulation: Hot new issues or same old stew? *Synthese* 169:593-613.

Gubrium, J. F., and J. A. Holstein. 1997. *The new language of qualitative method*. Oxford: Oxford University Press.

Hacking, I. 1983. *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge, UK: Cambridge University Press.

Humphreys, P. 2008. The philosophical novelty of computer simulation methods. *Synthese* 169:615-26.

Kennefick, D. 2000. Star crushing: Theoretical practice and the theoreticians' regress. *Social Studies of Science* 30:5-40.

Knorr Cetina, K. 1981. *The manufacture of knowledge. An essay on the constructivist and contextual model of science*. New York: Pergamon.

Knorr Cetina, K. 1999. *Epistemic cultures. How the sciences make knowledge*. Cambridge: Harvard University Press.

Küppers, G., J. Lenhard, and T. Shinn. 2006. Computer simulation: Practice, epistemology, and social dynamics. In *Simulation: Predictive constructions of reality. Sociology of the sciences yearbook*, ed J. Lenhard, G. Küppers, and T. Shinn, 125-38. New York: Springer.

Lahsen, M. 2005. Seductive simulations? Uncertainty distribution around climate models. *Social Studies of Science* 35:895-922.

Lenhard, J. 2007. Computer simulation: The cooperation between experimenting and modeling. *Philosophy of Science* 74:176-94.

Lynch, M. 1985. *Art and artifact in laboratory practice*. London: Routledge and Kegan Paul.

Mackenzie, D. 1990. *Inventing accuracy. A historical sociology of nuclear missile guidance*. Cambridge: MIT Press.

Mead, G. H. 1932. The objective reality of perspectives. In *On social psychology: Selected papers*, ed. G. H. Mead, 1964. Chicago: University of Chicago Press.

Mead, G. H. 1934. *Mind, self and society*. Chicago: Chicago University Press.

Merz, M. 2006. Locating the dry lab on the lab map. In *Simulation: Predictive constructions of reality. Sociology of the sciences yearbook*, ed. J. Lenhard, G. Küppers, and T. Shinn, 155-72. New York: Springer.

Oreskes, N., K. Shrader-Freshette, and K. Belitz. 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 263:641-6.

Roth, W. M. 2009. Radical uncertainty in scientific discovery work. *Science, Technology, and Human Values* 32:313-36.

Shackley, S. 2001. Epistemic lifestyles in climate change modelling. In *Changing the atmosphere: Expert knowledge and environmental governance*, ed. C. A. Miller and P. N. Edwards, 107-34. Cambridge, MA: The MIT Press.

Shibutani, T. 1955. Reference groups as perspectives. *American Journal of Sociology* 60:562-9.

Sundberg, M. 2005. Making meteorology. Social relations and scientific practice. PhD diss., Stockholm University, Stockholm, Sweden.

Sundberg, M. 2006. Credulous modellers and suspicious experimentalists? Comparison of model output and data in meteorological simulation modelling. *Science Studies* 19: 52-68.

Sundberg, M. 2009. The everyday world of simulation modelling: The development of parameterizations in meteorology. *Science, Technology, and Human Values* 32:162-81.

Sundberg, M. 2010. Organizing simulation code collectives. *Science Studies* 23:37-57.

Sundberg, M. 2011. The dynamics of coordinated comparisons: How simulationists in astrophysics, oceanography and meteorology create standards for results. *Social Studies of Science*.

Turkle, S. 2009. *Simulation and its discontent*. Cambridge: MIT Press.

Winsberg, E. 1999. Sanctioning models: The epistemology of models. *Science in Context* 12:275-92.

Winsberg, E. 2001. Simulations, models and theories: Complex physical systems and their representations. *Philosophy of Science* 68:442-454.

Winsberg, E. 2003. Simulated experiments: Methodology for a virtual world. *Philosophy of Science* 70:105-25.

Winsberg, E. 2006. Models of success versus the success of models: Reliability without truth. *Synthese* 152:1-19.

Winsberg, E. 2008. A tale of two methods. *Synthese* 169:575-92.

Bio

Mikaela Sundberg is an Associate professor at the Department of Sociology, Stockholm University. Presently, she conducts comparative analysis of numerical simulation practices in astrophysics, oceanography, and meteorology, presented in publications such as e.g. Organizing Simulation Code Collectives (Science Studies, 2010) and Cultures of Simulations vs. Cultures of Calculations? The Development of Simulation Practices in Meteorology and Astrophysics (Studies in History and Philosophy of Modern Physics, 2010).