Models, Fictions, and Make-Believe:
Fictional modeling and simulation in scientific practice

A Dissertation Presented

by

Adam M. Rosenfeld

to

The Graduate School
in Partial Fulfillment of the
Requirements
for the Degree of

Doctor of Philosophy

in

Philosophy

Stony Brook University

May 2012

Stony Brook University

The Graduate School

Adam M. Rosenfeld

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Dr. Don Ihde
Distinguished Professor, Department of Philosophy, Stony Brook University

Dr. Robert P. Crease
Professor, Department of Philosophy, Stony Brook University

Dr. Patrick Grim
Distinguished Teaching Professor, Department of Philosophy, Stony Brook University

Dr. Annamaria Carusi
Oxford University e-Research Cetre
and Norwegian University of Science and Technology

This dissertation is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

Abstract of the Dissertation

Models, Fictions, and Make-Believe:
Fictional modeling and simulation in scientific practice

by
Adam M. Rosenfeld

Doctor of Philosophy
in
Philosophy
Stony Brook University
2012

The notion that there are fictitional elements in science is not new, but historically the term has been used dismissively, as part of anti-realist positions about scientific theories. Recently, due to a growing philosophical focus on modeling, as well as the proliferation of computational modeling and simulation in scientific practice, the prospect of a progressivist account of the role of fictions in science (*viz.* one where fictions make a positive contribution to scientific progress) has been taken more seriously. This dissertation lays out a progressivist theory of fictional modeling that not only accommodates the representational character of scientific models, but their performative character as well. It distinguishes between a wide sense of scientific fiction, which I argue is an exaggeration of abstractions and idealizations that are arguably approximately true, and a narrow sense which is discontinuous with serious, non-fictional hypotheses and experimental practices. Even in this narrow sense, the inclusion of fictional elements in scientific models can solve pragmatic problems, facilitate accurate prediction, contribute to genuine discoveries, and perform explanatory functions; and they frequently do so not despite their fictional distortions of target phenomena, but precisely because of those distortions. Key to the satisfaction of these functions is the way that fictional models serve as props in acts of playful make-believe and are creatively re-interpreted in the course of virtual performances that can reveal novel features, structures, and possibilities for action in "serious," non-fictional contexts. This raises questions about how we might distinguish the intentional attitudes involved in make-belief from those involved in genuine belief, as well as normative questions about the extent to which effective scientific practice requires that we become fully absorbed in the make-believe of virtual scientific performances.

For my parents, who have taught me the meaning of unconditional love and support; and for Frances Bottenberg, without whom this dissertation would not have been possible.

Table of Contents

Preface

This dissertation is about fictions and the roles they play in scientific practice. Retrospectively, the seeds for it were sewn during a three year period when I was teaching high school science, and recognizing a recurring pedagogical pattern of articulating models, exploring where they break down, and then using these same, “broken-down” models to develop better ones. My focus was narrowed here at Stony Brook under the influence of Patrick Grim, who brought my attention to a variety of scientific computer simulations, and Don Ihde, who helped me to see and analyze these simulations, first, and foremost, as technologies.

The fundamental thesis of my analysis is this: Scientific models and simulations act as surrogates for certain target phenomena, and sometimes (but not always) these surrogates are inaccurate stand-ins in some important respects, such that those who use them would freely admit that the model/simulation represents its target as other than it actually is. I resist two facile responses to this observation.

The first is that the recognition of the fictionality of a model/simulation is grounds for dismissing it as a sound scientific tool. Fictional models *do*, in fact, work, and work in a variety of ways that contribute to scientific progress. It is not my intention to criticize the use of fictions in science, but rather to draw attention to them, and ask how it is that we are able to use constructions we know to be false to develop ideas that we believe to be true and practices that prove to be reliable.

The second facile response that I resist is that the functional utility of what I am calling fictions derives solely from key similarities between a target phenomenon and a model that is analogous to it – that they are alike in all of the relevant or essential ways, and the respects in which a model is “fictional” are not relevant to the way it functions as a model. I do not deny that models and simulations do exhibit similarities to their targets, but referring to these similarities in order to dismiss the relevance of dissimilarities is a mistake. Not only are such references to "relevant similarity" unhelpful in scenarios where the determination of just what the relevant features of a phenomenon are is precisely what is on the line, but they draw our attention away from the variety of ways that the dissimilarities between surrogate and target are significant and make positive contributions. We can do things with models and simulations that we can't do with their targets, and this is due to the ways in which they are dissimilar.

Thus models and simulations, particularly fictional ones, are rife with ambiguity. They are both similar and dissimilar to their targets in significant ways. Their interpretation is dependant on theories and an assumed representational relationship to the world, and yet they are also autonomous and taken on their own terms. And while this ambiguity offers the temptation to subsume their function under a similar representational ambiguity found in metaphors and analogies, this is only part of the story. Models and simulations do not simply represent (although we cannot completely disregard the fact that they are always models/simulations *of* some target); they are also concrete entities that are taken up and manipulated in practice. They are not simply extensions of or extrapolations from theory, nor are they simply pictures of the world; they are technologies. Moreover, they are a specific sort of technology that is ambiguously both representational and semi-autonomous — that is to say, they are "props" for scientific performances.

Insofar as we may think of models and simulations as props, the performances that fictional models and simulations are enlisted in have an essential component of playful make-believe to them; specifically, we make-believe that some target phenomenon *is* as its model portrays it. This attitude of make-belief is necessarily distinct from genuine belief. Nonetheless, it engenders a credulity that is difficult to disambiguate from genuine belief and this is what allows scientific fictions to have their distinctive scientific character: namely that they may give rise to serious, non-fictional discoveries. Rather than dismissing a fictional model outright as incapable of telling us anything interesting about the "real" world that we didn't already know, we stick with it, even when it makes challenging suggestions. We find novel ways of interpreting what the features of a model represent, negotiate novel ways of manipulating both model and target systems, and discover features of the target that were previously overlooked or thought to be irrelevant. This make-believe must be robust enough for the epistemological utility of a fictional model to play itself out, but not so robust that we are altogether incapable to recognizing where it falls short.

Overview of the dissertation

The path to this position unfolds over the course of six chapters. The first three chapters are loosely concerned with the representational character of scientific fictions, and the final three are loosely concerned with their practical and performative character.

Chapter 1 - In chapter 1, I set the background with an attempt to account for why the role of fictions in science is, for the most part, a relatively recent topic in philosophy of science. Hans Vaihinger's *Philosophy of 'As If,'* taken by many to be a crucial touchstone for this conversation, was published in 1911. Yet, it wasn't until Arthur Fine's revival of Vaihinger's theory in his 1991 essay "Fictionalism," that there seemed to be much appreciation of it. What changed in the intervening 80 years?

I argue that debates between scientific realism and anti-realism created a climate where a progressivist theory of fictions (i.e. one where fictional constructions are recognized as making a distinctive contribution to genuine epistemic progress) was unlikely to be explored. Between logical positivism's anti-realism, which cast theoretical concepts as either meaningless or in need of interpretation in testable observable terms, and social constructivism, which called into question the very possibility of objective epistemological progress in science, there was little room for the sort of discussion of fictions that I argue needs to be had. As long as philosophers of science were busy articulating and responding to these gambits in the realism/anti-realism debate, the recognition that many scientific models are both fictional, and instrumental to the development of non-fictional scientific theory, was unlikely.

Fortunately, the late 1970's and early 1980's saw the recognition of a stalemate in this debate, and ushered in a more moderate perspective. I take up two figures who made significant contributions that cleared the way for the progressivist approach to fictions in science I take.

Nancy Cartwright's "simulacrum account of explanation" challenges van Fraassen's assumption that theories may be explanatory while still being empirically adequate by making a distinction between fundamental theoretical laws and phenomenological laws. Phenomenological laws, Cartwright argues, can be empirically adequate, but they are messy and complex. Ultimately they are true, but at the expense of not being explanatorily satisfying. Fundamental theoretical laws, on the other hand, are explanatory, but at the expense of being true in anything other than imaginary conditions. In order to mediate between these two, we need models.

Ian Hacking's "intsrumental realism" shifts Cartwright's focus from laws, to entities, and perhaps more importantly shifts the general focus from a representational idiom, to a praxical interventionist one. Unlike the sort of instrumentalism that can easily be read as an anti-realism, one in which theoretical concepts are *merely* useful, but not literally true, Hacking argues for the necessity of a realist attitude when using theoretical entities as tools in experiments. His is a realism not simply predicated on reliable manipulation of objects as evidence that we are conceiving of them in an approximately true way, but one in which we must believe that things are real *in order* to use them the way we do in experimental inquiry.

Chapter 2 - In chapter 2, I continue laying out the groundwork for a progressivist theory of fictions by way of a discussion of scientific models and the ways they represent their targets. Models, like any other representation, have an ambiguous ontology. They can be real objects, or imaginary ones, and they can span a spectrum of relative abstractness or concreteness. Compared to theories, they are concrete. Compared to some particular real target system, they are abstract. Following Ronald Giere, I articulate a model of models as mediating between theories and real-world targets. Models model theories by instantiating a world where those theories can be interpreted as literally true, and can be thought of as constructed from law-like theoretical principles.

Their representational relationship to the world is a bit looser and more difficult to pin down. Models are more "iconic" than "symbolic" in that they do not represent merely arbitrarily or by intentional fiat, yet they are not necessarily faithful copies, even of the abstract formal relationships between features of the model/target. Giere refers to this relationship as "fit" in order to accommodate the looseness and elusiveness of the similarity between a model and its real-world target.

Within this notion of "fit" (or even in situations where the "fit" between models and targets is strained), there are many distinct ways that models can represent. They are frequently taken to "idealize" their targets. I examine two distinct types of idealization which, following Michael Weisberg, I refer to as minimalist idealizations (sometimes called abstractions, or negligibility assumptions) and Galilean idealizations. In particular, I consider the notion of "de-idealization" which, contra Weisberg, I don't think can adequately be used to distinguish between minimalist and Galilean idealizations, but is nonetheless a very useful concept. One notable aspect of models that do not represent faithfully is that some can be gracefully "de-idealized" (i.e. corrected for an arbitrary amount of precision as our needs demand) and some can not.

Among methods of representation that can not be gracefully "de-idealized" (or "de-fictionalized" in cases where distortions aren't really idealizations in any strict sense) are caricatures and analogies, which I take to be distinct from Weisberg's idealization scheme precisely in the way they resist both a straightforward account of the similarities and differences between model and target and how we may "de-fictionalize" by systematically correcting dissimilarities. Models that represent in this way draw our attention toward certain features of phenomena and away from others.

This drawing of attention to certain features and away from others depends as much on the distortions and dissimilarities between model and target as it does upon the similarities and indicates a non-neutral dimension to the way that models represent their targets. Following Catherine Elgin (who, in turn, is building off of Nelson Goodman), I call this 'exemplification,' and trace the complex structure at work therein. Key to exemplification is the representation of a real target *as* something else — and this, I maintain, is typical of all models that distort or otherwise represent without complete similarity. In this '*as*' structure— where model *x* represents target *z* as if it were *y*— *y* need not refer to any real, existing thing, even if *z* does.

The chapter concludes with a brief consideration of the technological side of models. Up to this point the focus has been on how models *represent* theories and their targets, but there is a praxical, *interventionist* dimension to them as well. Mary Morgan and Margaret Morrisson refer to this by drawing our attention to the manipulability of models and their role as investigative tools that provide epistemic access to their target systems. Models do not simply offer a fit between theory and the world, they must be fitted. This is an activity, as is the activity of de-idealization/de-fictionalization (whether it can be done gracefully or not). Models have an undeniable representational function, but as mentioned, this is only half the story. They do not simply sit there representing, they are taken up and put into motion. This indicates a need to keep our eye on the performative dimensions of modeling practices.

Chapter 3 - In chapter 3, I take up the question of what we mean by "fiction," what qualifies as a "scientific fiction," and what sort of approach to fictional models and simulations can accommodate both the representational and the performative character that is required.

Of primary concern is maintaining a sensitivity to a moderate scientific realist attitude, and avoiding a conception of fiction that fails to discriminate between models that are intended/taken to be approximately true, and those that are clearly not. Accordingly, I consider attempts to distinguish between a wide and a narrow conception of fiction. Hans Vaihinger offers such a distinction, claiming that the wide sense of the term may be applied to any proposition that is in conflict with reality in any way, and reserving the narrow sense for propositions that are self-contradictory. I reject this on the basis that genuinely contradictory models are difficult to find, and would be of dubious utility. Mauricio Suarez offers a better alternative, distinguishing between a wide sense that he calls "fictive," where fictive models represent existing targets falsely, and a narrow sense that he calls "fictional," where models represent objects that don't exist at all. Despite its superiority to Vaihinger's stance, this runs into trouble due to vagueness regarding how false a representation of an existing object must be before it is a representation of a non-existent one, as well as ambiguities regarding whole models that represent existing targets, but contain parts that represent non-existent objects.

As an alternative I suggest a distinction based on the notion of de-idealization/de-fictionalization discussed in the previous chapter. "Fiction" may be used in a wide sense to refer to models that are gracefully de-fictionalizable and can accommodate a continuous range of precision and accuracy as our needs demand. Models that abstract or idealize and provide "approximately true" depictions of their targets, ones that moderate realists may protest calling fictions at all, are fictions only in this wide sense. In the narrow sense, "fiction" refers to models that cannot be gracefully de-fictionalized. They may be approximately true but only in very specific contexts and for very specific purposes and they require significant alteration or re-interpretation in order to be applied outside those contexts.

I then address an ambivalence regarding "scientific fictions" that is easy to overlook in all of the above efforts to discriminate between wide and narrow senses of the term. While it seems useful to point out that fictions are false in some significant way, focusing on their falsity almost misses the point of a fiction. An essential characteristic of our attitudes toward fictions is that we don't care about their truth or falsity. I argue that the failure to emphasize this aspect is a symptom of thinking of scientific fictions as scientific first, and fictional second, where the reverse is more appropriate. By thinking of fictional models as fictional first, and scientific second, we would first take up models without regard to whether they represent their targets accurately (or even represent anything in the real-world at all), and only after that do we take up questions of whether and how they fit target phenomena.

A way of accommodating this reversal was already indicated by the "representation-*as*" structure discussed in the previous chapter, but another, thicker and more performative account of this can be found in Kendall Walton's account of mimetic play. I argue that despite the fact that perspectives from the performing and fine arts such as Walton's have gone underappreciated by philosophers of science who address the topic of the putative fictionality of scientific models, these approaches are particularly well suited to bridging the gap between the representational character of models, and their technological and performative character.

I close this chapter by reiterating that the account given here of fictional scientific models is one that can accommodate a moderate scientific realism. I do not claim that all models are fictions, at least not in the narrow sense of the term, only that some can strain even the loose sense of "fit" that folks like Giere argue for, and can rightly be called fictional. As such what I am putting forward is a theory of fictional scientific models, not a fictionalist theory of scientific models.

Chapter 4 - Having dealt with the representational character of models in chapters 2 & 3, the fourth chapter fleshes out the technological and performative character that has been argued to be their other, more overlooked aspect. Rather than simply extensions of theory, the manipulation of models in simulational practice has significant parallels to experimentation. Yet, due to their capacity for fictional representation, the virtuality of models and simulations introduces a significant epistemological complication that makes them distinct from experiments. But, apart from their virtuality, it is not immediately evident how these so-called "dry-lab experiments" are distinguishable from "wet-lab" experiments.

Experimental technologies represent phenomena, and frequently do so by introducing distortions. Yet, there is a difference between technologies that *make* representations, and technologies that *are* representations and function as surrogates for some target system. I distinguish experimental representations from simulational ones by virtue of a relatively direct and traceable causal relationship to their representanda for the former, and a complex, indirect, or even non-existent causal relationship for the latter. Simulational technologies do not intercede between an object of inquiry and an investigator so much as they are constructed by an investigator to stand in for the object of inquiry.

It is notable that despite the centrality of the role of imaginative make-believe in interpreting these surrogates, and the relatively abstract character of many simulations (especially computer simulations) any attempt to distinguish simulational practices from experimental ones by virtue of a supposed non-materiality of simulations is misguided. *All* simulations have an unavoidable materiality, and this plays a central role in the ways they are taken up and manipulated as props. Furthermore, attempts to soften this position by recourse to a allegedly “deep material similarity” for experiments that does not obtain in most simulations, fails to explain why the type of material involved should constitute a *relevant* similarity.

Against these sorts of attempts to distinguish between experiments and simulations, I offer a provisional distinction between simulational paradigms and experimental ones, based on the assumption that experiments typically begin with a target phenomenon and “cut” features away, while simulations begin from scratch and add features. Both practices converge on an ideal middle ground that is relevantly similar to the phenomenon under investigation, but also relatively more intelligible and/or easily manipulated. This indicates significant differences, not so much regarding what counts as a *successful* simulation or experiment, but rather regarding what counts as a *promising* simulation or experiment.

This schematic is easily complicated in intermediate sorts of constructions, which Joseph Rouse refers to as “laboratory fictions.” In particular, so-called “model organisms” and many other surrogates that are found ready-made more so than constructed are frequently understood as belonging to a "wet lab" paradigm. Additionally, a careful analysis to the ways that features are "added" and "subtracted" from systems reveals that such "additions" and "subtractions" are not so easy to tell apart. In concrete systems, we must frequently "add" in order to "subtract." While the general schematic of simulational and experimental practices may be a useful heuristic, it is more accurate to acknowledge that either practice rarely stands on its own, but is, instead, part of a complex set of practices that are sometimes simulational, sometimes experimental, and sometimes ambiguously between the two.

I conclude this chapter with a case study involving the simulation of a fluid dynamic system in 2-dimensions that illustrates these points, showing how simulational structure and experimental apparatus converge through a series of re-interpretations of the simulational model, and discoveries of novel ways of manipulating the experimental system. Through these re-interpretations, the simulation in question shifts from being a representation of the target system to a blueprint for novel ways of manipulating the experimental apparatus by revealing the salience of the ends of a cylindrical obstruction in the determination of wake dynamics.

Chapter 5 - Having introduced the notion of successful and promising models and simulations in chapter 4, chapter 5 investigates a variety of functions according to which promise or success could be measured. These functions are not intended to be an exhaustive account of the functions of models, but rather indicate the diversity of aims which they can serve, and perhaps more importantly for the focus of this dissertation, how the introduction of fictionalized elements can contribute to the satisfaction of these aims.

First I address pragmatic functions, which are perhaps the most frequently acknowledged motivation for studying a surrogate system rather than the target system directly. We opt for surrogates when access to the real thing is difficult, costly, or morally/politically problematic.

Notably, such pragmatic concerns are non-epistemic (or at least not directly epistemic in nature), and if models/simulations that satisfy such pragmatic ends are epistemologically acceptable, it is despite their fictional distortions, not because of them.

A less contentiously epistemic function can be found in prediction. Particularly in cases where the initial conditions and causal mechanism of the target system are not well understood, many simulational studies hone in on predictive success by way of parametric tweaks and kludges. The resulting simulations are capable of reasonably accurate predictions, but their structures and mechanisms become "black-boxed" and are not expected to offer any insight into the structure and mechanism of the target system.

Additionally, even in situations where inputs and mechanisms are fairly well understood, making predictions is, itself, a praxis, and rendering the complex phenomenological laws that are capable of making accurate predictions computationally tractable frequently demands the introduction of fictional distortions. These distortions, especially when combined with those introduced for non-epistemic pragmatic purposes, can quickly compound into a complex web of fictional elements where modelers must "get things wrong locally in order to get them right globally."

It is important to note that this introduces a slippery distinction between "models" whose mechanisms are so irrevocably "black-boxed" that they are reducible to the technological function of prediction-making-machines, thus ceasing to represent their targets in any strong sense, and those whose structure and mechanism are only temporarily "black-boxed" with the intention of eventually articulating a robust correspondence between model and target.

This latter case introduces a way that models also serve to facilitate discoveries of novel structures and previously overlooked salient features of target systems. I call this function "heuristic," drawing on the etymological roots of the term that emphasize discovery. This is, perhaps, the most complex and controversial function of fictional models and simulations, as it is often argued that these tools can only yield discoveries about their own fictional worlds, not about the actual world.

Drawing on the case study from the previous chapter, as well as an additional one involving the model-organism *C. elegans*, I illustrate how the complex hermeneutic task of interpreting and re-interpreting the relationship between a model and its target reveal surprising and overlooked features of both. In these cases, it is unlikely that the discoveries made would have been possible through the direct investigation of the target systems alone, illustrating the unique utility of the *dissimilarity* between the surrogate and target systems. The discoveries facilitated by fictional models rely on the eventual translation of what is discovered virtually into reliable manipulation of real systems. This does not discount the importance of the model in such discoveries, but rather underscores the way that discoveries are processes that unfold over time and are not localizable to a single moment.

Finally, I address the explanatory function of models. This function is clear enough for non-fictional and/or "successful" models where a robust correspondence between relevant structural features and causal mechanisms of the model and target have been established. It is less clear in the case of models with significant fictional distortions. In order to embrace an explanatory function for these models, we must accept a fairly controversial, Cartwright-ian conception of effective explanation as not necessarily (and perhaps even necessarily not) true. Examples of models and simulations that fulfill this purpose include those that illustrate how complex systems can emerge from relatively simple sets of rules.

Chapter 6 - The final chapter of this dissertation addresses the nature of make-believe and its relationship to belief. As such it marks a return to a variety of claims that have been made throughout the other chapters that are in need of elaboration. I argue that the attitude of make-believe that scientists bring to fictional models and simulations is remarkably similar to, but nonetheless distinct from, genuine belief.

This argument is made by way of Hans Georg Gadamer's account of play and the peculiar "seriousness" that authentic absorption in play requires. It is impossible to be fully engaged in a make-believe performance, and simultaneously be explicitly aware that it is *merely* make-believe, and not genuine belief. Structurally, this claim is similar to Moore's paradox, which highlights the apparent contradictoriness of the conjunction of an assertion of some state of affairs with a second assertion that one does not, in fact, believe the first assertion.

I entertain previously articulated ways out of/around Moore's paradox, but ultimately reject them as unsuitable for clarifying a distinction between make-believe and genuine belief. Instead, I opt for an account of what Francois Recanati calls "quasi-belief," which describes a fragile credulity hinging on a tension between the relatively well insulated internal coherence of a well constructed fiction, and its conflict with our network of previously held genuine beliefs. Insofar as authentically absorbed make-belief is not simply the act of a willing subject, but is instead characterized by a passive responsivity to the internal logic of a "game," acts of make-believe, including those involved in fictional modeling and simulation, have an essential robust fragility. Make-belief is distinct from genuine belief not in terms of our attitude when we are absorbed in it, but rather in how it breaks.

I conclude with comments regarding the normative consequences of these insights. The full function of scientific models and simulations, including those with fictional distortions, requires that scientists be adequately immersed in their virtuality that they are allowed to fully play themselves out. At the same time, they must be capable of recognizing where these models and simulations fall short in crucial ways in order to re-interpret them and/or avoid mistaking fictions for non-fictions. What this requires is a virtuous mean between the extremes of excessive credulity and excessive skepticism. A fuller exploration of what should count as virtuous scientific make-belief is required. Additionally, the education of young scientists to cultivate virtuous scientific make-belief, and the strategic design of models and simulations to facilitate the same is a necessary goal for the development of a science in which the role of modeling and simulation practices is becoming increasingly ubiquitous.

Chapter 1
Fictions and Anti-Realisms

Those who are particularly well versed in philosophy of science may note that the attribution of the term “fiction” to certain elements of scientific practice is not an unprecedented idea. “Fictionalism” goes back as far as positions laid out by Piere Duhem and Hans Vaihinger in the early 20th century. Duhem’s fictionalism can be seen largely as an extension of the spirit of Comte’s positivism which sought to liberate empirical science from metaphysical speculation. Duhem sought to address a perceived cultural schism in science based on stereotypes of a British method which emphasized the development of mechanical models to explain phenomena and a German logical empiricist method (somewhat ironically exemplified by Newton, an Englishman). While Duhem had significant criticism for both methods (Ariew 2011), he goes out of his way to launch a particularly strong critique of James Clerk Maxwell, whom he took have needlessly infected otherwise laudable advances in electrodynamics with physical models that could not be taken seriously as genuine scientific hypotheses (Ariew and Barker 1986). Hence, Duhem’s fictionalism was a dismissive sort of fictionalism, directed toward what he took as a misguided affinity for physical modeling, and which labeled such models as *mere* fictions, By identifying physical models as fictions, Duhem meant to say that we could and should refrain from building our empirical scientific theories around them.

I’m less interested in this sort of dismissive fictionalism than I am in a theory of fictions that asserts a significant epistemic role for scientific models while still acknowledging their fictional status. Let’s call this a “progressivist” theory of scientific fictions. Such a theory can be found in Hans Vaihinger's 1911 publication *Philosophie Des Als Ob* (*Philosophy of 'As If'*). Vaihinger's fictions, though they may be known to be false (and occasionally logically impossible) representations of reality, are nonetheless characterized as having potential practical value. By assuming that the world actually behaves 'as if' it was working according to fictional models, we may be able to accomplish much and even eventually discard our fictions in favor of non-fictional representations. Thus fictions may be tools, not only for practical activities, but for inquiry as well.

Vaihinger's theory of fictions was intended to be applicable to all domains of human knowledge but it seems (if only on the basis of the examples used by Vaihinger to illustrate his thesis) to be especially well suited to describing and explaining physical science, which by the early 20th century, with the rise of quantum and relativistic theories, was drifting into metaphysically strange territory. Many of the fundamental phenomena being studied by physics could be modeled with familiar mechanisms that failed to adequately capture what was going on, or they could be modeled with more adequate but bizarre mechanisms that defied coherent comprehension. Either way, it seemed as if the only ways to talk about the invisible mechanisms behind what was empirically observable was to adopt a language of "as if." Fictions were, seemingly, everywhere.

Yet, despite some enthusiastic initial reception both on the continent and in America, Vaihinger's theory of fictions was met with sharp criticism and a subsequent fall into relative obscurity within the philosophy of science. Especially noteworthy as a cause of Vaihinger's poor reception was the rise of logical positivism— a doctrine opposed in spirit to the idea of an indispensable role for fictional models in science—

which came to dominate much of early 20th century philosophy of science. Ironically, though the term "logical positivism" would come to be identified with the Vienna Circle's logical empiricism, it was Vaihinger who first used it and did so to describe his own position (Frank 1949, p43, Vaihinger 1911, p163).[1] While the similarities between the work of Vaihinger and that of figures such as Schlick, Mach, Carnap, and Neurath go beyond their self ascribed labels, "Vaihinger is seldom cited by the Vienna positivists as a precursor, and never as an ally" (Fine 2009, p20). So strong was the desire on the part of Vienna positivist to distinguish themselves from Vaihinger that Moritz Schlick would write that Vaihinger's claim that his *Philosophy of 'As If'* was a form of idealistic positivism was "but one of the contradictions from which [his] work suffers" (Schlick 1932, p85) and that Vienna positivism "is not a 'Theory of As If" (Schlick 1932, p 107).

I am not interested in the academic exercise of sifting through the similarities and differences between Vaihinger's "positivism" and that of the Vienna Circle. In fact, I wish to table any further discussion of Vaihinger's *Philosophy of 'As If'* until chapter 3 in favor of properly laying the groundwork for a more contemporary theory of the role of fictions in science. What I am more interested in first, is how it came to be that despite its being relatively ignored for most of the 20th century, there has been a significant revival of interest in fictions/fictionalism and in Vaihinger's contributions to thinking of fictions in science recently. Assuming that this reintroduction of fictionalism is finding a warmer reception now than it did at the beginning of the 20th century, what has changed in the intervening 80+ years?[2] Was Vaihinger simply ahead of his time, and the inhospitable context into which he released his *Philosophy of 'As if'* simply historically

[1] Vaihinger also uses the terms "critical positivism" and "idealistic positivism."

[2] Taking the 1993 publication of Arthur Fine's essay "Fictionalism" to mark a renewed attention to Vaihinger's work.

unfortunate and accidentally mismatched to the temperament of his thesis? Has science itself changed in the intervening time in some important way that makes a theory of fictions more appealing or plausible today than it was in 1911? There is some truth to both of these answers. Science has, indeed, changed in important (especially technological) ways, especially with respect to an increasing prevalence of computational simulations. But additionally, there have been important developments in the philosophy of science, developments which concern sorting out which aspects of science are game for being regarded as fictions, and which are not. This is what I will focus on in this chapter.

What follows is an attempt to sketch out the broad contours of these changes in the philosophy of science, specifically concerning attitudes toward scientific realism/ anti-realism, between Vaihinger's publication of *The Philosophy of 'As If'* and its relatively belated revival. This is an ambitious undertaking for a single chapter, perhaps too ambitious. It should not be taken to be a thorough account of 20th century philosophy of science (even if limited to scientific realism/anti-realism) or of the various positions that I touch on (logical positivism/empiricism, abductivist realism, and constructivism). There are already scores of far more rigorous treatments of these topics. What this chapter does offer is a sense of how the shifting intellectual climate throughout this debate made it difficult for a discussion of the contributions of fictions to scientific progress to take place..

Naïve Realism and the Burden of Proof

The question of which (if any) aspects of scientific inquiry and practice can be regarded as "fictional" is the compliment of a more familiar question in philosophy of science, namely: which (if any) aspects of scientific inquiry and practice can be regarded as "real?" The position that science by and large provides us with knowledge of how the things that it studies "really" are can be loosely designated "scientific realism." It is important to begin a discussion of scientific realism and arguments against it with a careful consideration of where the burden of proof lies. We may acknowledge that science is not infallible, that it occasionally, or even frequently gets things wrong. This is evident enough to any even casual science student who may realize, perhaps with some frustration, that the history of science is littered with the carcasses of once well respected theories which have since been abandoned for other theories. But this acknowledgement of fallibility is not, on its own, the same as a rejection of scientific realism. For the pattern of abandoning theories seems to be one which replaces flawed theories with better ones, producing a succession of more and more *approximately true* theories, each more accurate than the previous one, ultimately converging (however distantly) on the truth—that is to say on a description of the world the way it really is. Moreover, the world and the truth about it are the way that they are independently of what we think (and perhaps even whether we think about them at all).

Such is the commonsense understanding of science and its progress; that any reasonably current science textbook talks, for the most part, about things that really do exist and says approximately true things about them, with future editions being, in all likelihood, even more accurate. Let us call this "naïve scientific realism." As a commonsense position, this hardly qualifies as a philosophical position, which is not to say that philosophical positions necessarily conflict with commonsense ones, but rather that it has not been "earned," so to speak, via careful consideration against critique from a contrary position.

For the naïve scientific realist, it is not evident that a reasonably coherent contrary position even exists. Richard Boyd puts this well with a reference to more mundane sorts of commonsense realism like that concerning tropical fish. Tropical fish realism would be "the doctrine that there really are tropical fish; that the little books you buy about them at pet stores tend to get it approximately right about their appearance, behavior, food and temperature requirements, etc.; and that the fish have these properties largely independently of our theories about them." (Boyd 2002) Absent any reasons why we ought not be tropical fish realists, it is completely rational for us to be so.

The same goes for any commonsense realism, which is to say it is unreasonable to demand proof that commonsense beliefs are true if there is no reason to doubt them. Essentially, this amounts to a principle that designates naïve belief as the starting point of philosophical thought rather than radical skepticism. Though it has not always been scrupulously adhered to in the history of philosophy, it is returned to again and again, from Socrates entreating his interlocutors to begin by saying what they think, to C.S. Peirce's critique of Descartes' radical methodological doubt (Peirce 1878), to Russell who writes, "Skepticism, while logically impeccable, is psychologically impossible, and there is an element of frivolous insincerity in any philosophy which pretends to accept it" (Russell 1948, 9). This principle does not render commonsense beliefs immune from criticism, but does require that the burden of proof be on those who would critique them.

For scientific realism, if indeed naïve scientific realism is the default pre-philosophical position, the burden of proof is on the anti-realist.

The Empiricist Challenge

So why shouldn't we be satisfied with naïve scientific realism? There are, as mentioned earlier, significant anti-realist challenges, which, along with realist rebuttals constitute what is arguably *the* major debate in 20th century philosophy of science. The first of these comes from empiricism and is taken up by the logical positivists who sought to distance themselves from Vaihinger's fictionalism. Logical positivism (or logical empiricism, I will use the two labels interchangeably[3]) is characterized by a turn to linguistic analysis in order to develop a rigorously empirical account of science, one which clearly separates scientific thinking from metaphysical thinking and thus claims that science can, and should, free itself from any and all lingering traces of metaphysics. The cornerstone of empiricism, in particular what is usually referred to as *knowledge* empiricism, is the fairly plausible claim that all knowledge of the natural world is derived from sensory experience, and the connection of this claim to a positivisti rejection of metaphisics finds its roots in the empiricism of David Hume.

Hume makes a distinction between relations of ideas and matters of fact. We can reason analytically about necessary relationships between ideas, but matters of fact are known empirically, and any ampliative, synthetic reasoning about matters of fact must be done inductively. Attributing the necessity of a causal relation between two empirically

[3] While I realize that the conflation of logical positivism and logical empiricism is not altogether uncontentious, discriminating between these two terms and exploring the fine differences between the positions they represent is beyond the scope of this investigation. Recall that I am concerned, here, with the general climate of discourse in the philosophy of science throughout the 20th century. Some fine distinctions will necessarily be glossed over.

experienced events, therefore, amounts to an illicit confusion of relations of ideas and matters of fact. At best, we can identify an historically consistent conjunction of empirical experiences, one which we may inductively infer will likely continue in the future, but never causation. For if causation is anything other than a reliable conjunction of two empirically observed phenomena, it involves the necessity of the relationship between cause and effect, and such necessary relationships are only properly to be found in analytic relations of ideas, not in what can be known about matters of fact.

Without the ability to invoke a principle of causation, all manner of unobservable entities and "secret powers" of observable things are off limits. I may be justified in believing that I will be nourished when I eat bread if being nourished has consistently followed eating bread in the past, but I am not justified in attributing the "secret power" of nourishment as an essential property of the bread. To do so is to grant the bread the status of *cause* for the effect of being nourished on the basis of empirical experience, and empirical experience alone won't support this. Likewise, any inference of some unobserved state of affairs that is responsible for, that is to say *causes* our empirical experiences, is attempt at the illicit alchemy of turning observations of fact into relations of necessity between ideas, of turning physics into metaphysics. (Hume 1977, 1-37)

The consequences of this run deep, implicating not only unobservable theoretical entities such as electrons or gravitational fields, but also the legitimacy of metaphysical and speculative philosophy in general. Hence, we have what is known as "Hume's fork," that the only acceptable methods of reasoning are the logical analysis of relations of ideas, or the empirical study of matters of fact, most provocatively expressed; "If we take in our hand any volume; of divinity or school metaphysics, for instance; let us ask, *Does* *it contain any abstract reasoning concerning quantity or number?* No. *Does it contain any experimental reasoning concerning matter of fact and existence?* No. Commit it then to the flames: for it can contain nothing by sophistry and illusion" (Hume 1977, 114).

The impact of Hume's skeptical challenge was somewhat softened by a response from Immanuel Kant, who argued that in addition to analytic, *a priori* relations of ideas and synthetic *a posteriori* matters of empirical fact, there are "synthetic *a priori*" judgments which make a metaphysical foundation for natural science possible. But early 20th century developments in physics, particularly those in the emerging quantum and relativity theories such as particle/wave complementarity and non-euclidean space-time, undermined the legitimacy of the classical intuitions about objects, space, and time which Kant claimed to have proven as synthetic *a priori*, and led those in the Vienna circle to reject synthetic *a priori* judgments as a source of knowledge. Similarly, advances by Russell and Wittgenstein in the articulation of the logical grounding of mathematics undermined Kant's insistence that mathematics was a synthetic, rather than analytic enterprise. Consequently, logical positivism/empiricism can be viewed as a return to the spirit of Hume's "fork."

The potential consequences run even deeper, implicating not only unobservable theoretical entities such as electrons or gravitational fields, but also the substantial external objects responsible for our empirical experiences. In short, Hume's "skeptical problem" threatens even "tropical fish realism" in a way that isn't clearly rescued by his "skeptical solution" (Hume 1977, 25-37). We may rationally believe that certain sense impressions will be accompanied by others, but not that any particular external object is causing them. In short, Hume offers no clear way to avoid slipping from empiricism to radical phenomenalism, the belief that all that exists, or at least all that can be known, is reducible to the mental experiences of percievers.

A logical positivist, however, is not necessarily (though they may be) committed to such a radical empiricism as that which leads to more implausible phenomenalist conclusions. She may be perfectly content to subscribe to a realism for ordinary observable things provided that they can be translatable into a language whose terms only refer to actual or potential empirical experience.[4] As Rudolph Carnap puts it "One of the fundamental theses of *positivism* may perhaps be formulated in this way: every term of the whole language L of science is reducible to what we may call sense-data terms or perception terms" (Carnap 1936, 67). Precisely what is meant by "reducible" here is a matter of debate throughout the evolution of the positivist project. It ranges from analytic "definition" and immanent, complete "translatability" (Carnap 1928) to "verifiability" and modifications on what counts as "verifiable." But the core sentiment is always that the meaning of any statement made about the world lies in its observable consequences. A.J. Ayer makes this quite clear, writing, "For I require of an empirical hypothesis, not indeed that it should be conclusively verifiable, but that some possible sense-experience should be relevant to the determination of its truth or falsehood. If a putative proposition fails to satisfy this principle, and is not a tautology, then I hold that it is metaphysical, and that, being metaphysical, it is neither true nor false but literally senseless" (Ayer 1936, 41). Since anything beyond analytic definitions and statements that could be

[4] The distinction at work here between the positivist brand of scientific anti-realism and the radical phenomenalism I am prepared to reject out of hand is a subtle but important one. Essentially it hinges on the difference between saying that everything is *reducible to* mental events (as the radical phenomenalist holds), and requiring that a statement be translatable into empirical language in order to have any meaningful epistemic content. The former is a metaphysical position, the latter an epistemological one which places normative limits on what science talk meaningfully about.

empirically tested by observation is meaningless and has no place in empirical science as far as the logical positivist is concerned, all that is left for *philosophy* of science (if not philosophy in general— as the early Wittgenstein declared) on this view, is a second-order logico-linguistic analysis of scientific theories through which the lingering traces of metaphysical superstition could be either translated into empirical language, or purged from the discourse.

Thus, an empiricist challenge to scientific realism that stems from logical positivism is not a wholesale rejection of the truth of scientific knowledge, but involves a rejection of the treatment of unobservable theoretical entities as real. This rejection can be derived from a very concise argument, generally referred to as the "underdetermination argument" (Boyd 1983). A given theory about unobservable entities may be posited by a scientific realist to be supported by empirical evidence; specifically evidence in the form of confirmed observable predictions derived from said theory. The empiricist anti-realist would respond by asserting that there exists at least one possible competing theory that is significantly at odds with the first regarding its claims about the putative unobservable entities, but which is empirically equivalent, that is to say it makes the same observable predictions. There is no means for choosing the first theory over the second (or any number of empirically equivalent theories) since the available empirical evidence underdetermines which of the competing theories is correct. Absent any compelling empirical evidence for the unobservable entities posited by a theory over those posited by some competing empirically equivalent theory, there is no *meaningful* distinction between them. Theoretical concepts and theoretical entities do not denote anything meaningful outside of the observable conditions they correspond to.

Quoting Carnap again,

> I wish to emphasize here that this talk about the admission of this or that kind of entity as values of variables in L_T [*a theoretical language, as opposed to L_O, and observational language*] is only a way of speaking intended to make the use of L_T […] more easily understandable. Therefore the explanations just given must not be understood as implying that those who accept and use a language of the kind here described are thereby committed to certain "ontological" doctrines in the traditional metaphysical sense. The usual ontological questions about the "*reality*" […] of numbers, classes, space-time points, bodies, minds, etc., are pseudo questions without cognitive content.

and later

> […] We may, in view of our partition of the total language L into the two parts L_O and L_T, distinguish between the use of "real" in connection with L_O, and that in connection with L_T. We assume that L_O has only one kind of variable, and that the values of these variables are possible observable events. In this context, the question of reality can be raised only with respect to possible events. The statement that a specified possible observable event […] is real means the same as the statement that the sentence of L_O which describes this event is true, and therefore means just the same as this sentence itself. (Carnap 1956, 44-45)

Thus, if there appear to be significant distinctions between two empirically equivalent theories— say, the putative existence of some unobservable entity— these apparent distinctions are not meaningful. Statements about the existence of such entities can't ever be said to be true or false, as they do not actually refer to the unobservable entities they appear to, and are instead figurative expressions which ought to be understood in empirical terms.

It is relatively easy to see how this argument, if sound, would spell trouble for any "progressivist" theory of fictions. The ostensible appeal of theoretical statements about unobservable entities, that observable phenomena are understandable as being 'as if' they were caused by unobservable states of affairs are apparently without any epistemic warrant. This is all well and good for the fictionalist, who doesn't claim that such epistemic warrant exists (on the contrary, such "as if" claims are intended to be *fictional* statements), but the logical empiricst's objection goes beyond denying the epistemic warrant of theoretical claims about unobservables. They maintain that such statements are either meaningless or refer figuratively to empirically observable conditions, and a fiction is neither meaningless nor something that must be figuratively interpreted. Furthermore, the logical positivist/empiricist maintains that science can, and indeed ought to do without such claims, that they amount to infecting empirical science with metaphysics and give rise to theoretical debates which are nothing more than pseudo-problems that cannot be resolved. According to the logical positivist, a theory of fictions that asserts a positive or necessary role for 'as if' statements in scientific inquiry is misguided.

Theory-Ladenness of Observation

The preceding section gives an account of an anti-realist movement within the early 20th century that undermines naive realism and is inhospitable to the notion of a crucial role for fictional models in science. However, this climate saw a significant shift in the beginning of the latter half of the 20th century following arguments that scientific methods of observation were inherently and deeply "theory laden" (see, e.g. Hanson 1958; Sheffler 1967). If, indeed, empirical observations require certain theoretical assumptions in order to be made and organized meaningfully, then a significant blow is dealt to a key assumptions of logical positivism, namely that there is something special about direct sensory perception that privileges it over instrumentally mediated "observation" of otherwise unobservable entities; a distinction that was coming under attack just as the influence of logical positivism was beginning to wane. (see e.g. Feigl 1956, Maxwell 1962)

The positivist's account relies on a distinction between observable entities and unobservable ones, but even a casual analysis reveals that this distinction lacks clarity. Is empirical observation limited to "naked" observation of phenomena? Scientists frequently "observe" otherwise unobservable objects via instrumental mediation, as when a single photon is detected by a photo-multiplier tube, or *alpha* or *beta* particles are observed in a cloud chamber. The empiricist anti-realist may protest that such objects are not *observed* so much as *inferred* from observations. To say that a photon is observed begs the question of the acceptance of a theory that posits the existence of photons, which theories are, they would argue, underdetermined by direct empirical evidence. For the same reason, atoms are not observed, say, *via* electron microscopy. But how are we to differentiate these instrumentally mediated "observations" from more mundane ones? Are the moons of Jupiter theoretical objects simply because we need a telescope to see them? Surely the scientist whose myopia requires that she wear eyeglasses still makes visual observations, and we would not discount everything she sees as unwarranted theoretical inferences about objects. So what meaningful difference separates her from the scientist who "observes" through an optical microscope, an electron microscope, or a mass spectrometer?

To be sure, there is a sophistical flavor to this argument. The inability to identify precisely where the shift from observable to unobservable occurs does not necessarily indicate that there is no distinction between the two, and provided there are clear cases of an unobservable entity, say for instance a gravitational field or a quark, the distinction may be said to hold. At best, we may say that the class of observable objects is slightly larger than a hard-core empiricist would suppose, but not that all empirical effects of putative theoretical entities counts as empirical observation of those entities. However, the ambiguity cuts the other way as well, and if it can be said that all empirical observation (even "naked" observation) rests upon theoretical presuppositions, then this crucial distinction for the positivist's anti-realism gets shaky.

The vast majority of data production and collection in modern science is considerably a more complex and active affair than the passive connotations of the term "observation" suggest. If we consider that this data is produced in experiments, it becomes clear that experimental design and execution determine the sort of data that is available for observation. If we design an experiment in a slightly different way, we will get very different data, and experimental design is, more often than not, a thoroughly theoretical affair containing assumptions not only about the phenomenon explicitly under investigation, but the inner workings of the battery of instruments used in a modern laboratory. To even begin to construct these laboratories, let alone get involved in the process of discriminating between which laboratories are producing reliable data and which are conducting flawed experiments, requires recourse to theoretical tools, including theories about what isn't being directly observed and phenomena which are, in principle, unobservable. Even the explicit theoretical assumptions involved in the activity of experimental science are so densely tangled up with observation as to make the prospect of giving a empirical account with a merely figurative understanding of theoretical concepts rather far-fetched. Add to this the acknowledgement that there are myriad tacit theoretical commitments as well, and the task of reconstructing science exclusively in operational and observational terms begins to look impossible.

Furthermore, even more mundane acts of perception appear to be “theory laden” as well. R.H. Steuwer (1985) offers a telling account of James Chadwick’s investigation of conflicting observation reports of scintillations (thought to signify the emission of charged particles from certain elements under radioactive bombardment) between the Rutherford lab in Cambridge where he was working and the Petterson lab in Vienna. Members of the Petterson lab continued to “observe” scintillations even when Chadwick was manipulating the experimental apparatus so as to make it impossible for emitted particles, even if they did exist, to hit the scintillation screen. T.S. Kuhn (1962) recounts similar results in perceptual psychology by Bruner and Postman that showed subjects to be relatively incapable of recognizing anomalous playing cards, reporting to have observed, for instance a red four of hearts when presented with a black four of hearts. Such accounts suggest a deep sense in which what is expected conditions what is observed, or as N. R. Hanson (1958) put it, that “seeing is a ‘theory laden’ undertaking.”

It is possible to dismiss these sorts of accounts as identifying a regrettable psychological succeptibility to confirmation bias which can, with effort and careful attention, be combated. Failing any one person’s ability to control their own confirmation bias, this failure can still be recognized when pointed out. However, there is an additional dimension to the theory-ladenness of observation that is more difficult to dismiss. Observation reports do not convey the entirety of observable phenomena in an experiment, but rather are selective. This selectivity comes in to play not only in the decisions made in experimental design about what to measure and record, but also in post-experimental interpretation of data. Data does not speak for itself, but requires that certain bits of data be identified and grouped as significant while others are discarded as insignificant. Any intelligible observation requires that the observer have some idea of what they are observing in order for an observation report to be meaningful. This requires that all observations be conditioned by meaningful theoretical commitments that are external and prior to empirical experience.

Constructivist Anti-Realism

That observation is deeply and unavoidably theory laden is fairly widely accepted by now, and considered by many (along with similar arguments regarding semantic holism from W.V.O. Quine (1951) that dismantle the analytic/synthetic distinction) to have put the nail in the coffin for the version of logical positivism described above. By 1967 John Passamore would announce in his entry for “Logical Positivism” in the Encyclopedia of Philosophy that “Logical positivism [...] is dead, or as dead as a philosophical movement ever becomes” (Passamore 1967). However, this does not leave scientific realism as the victor-by-default in this debate. The appeal to theory-ladenness opens the door to a much stronger and more diffuse sort of anti-realism known as ‘constructivism.’ If observations are conditioned by theoretical commitments, the constructivist says, then it is circular to use empirical data as evidence for a theory; what counts as “evidence” is a flexible enough category that a theory could conceivably accommodate any empirical data, even if this “accommodation” is to dismiss inconvenient data as an ignorable anomaly. How, then, can the fact that science appears to progress, that old theories are abandoned in favor of new ones, be explained?

As difficult as it is to give an adequate generic account of logical positivism/empiricism, constructivist anti-realism is even harder to pin down. A fairly representative figure, however, can be found in Thomas Kuhn. In his *The Structure of Scientific Revolutions* (1962), Kuhn articulated an historical account of major theoretical shifts in terms that render scientific theory selection (and by extension all of the theory laden practices of science) a primarily social phenomenon. To do this, Kuhn introduced the simultaneously powerful and problematic concept of a "paradigm." Though there is significant debate amongst Kuhn scholars about the proper extension of the term, it is clear that Kuhn intended it to mean more than just theoretical commitments. Additionally, it includes traditions of what constitutes acceptable methodology, standards of evidence, the meanings of terms, and even standards of rationality. It is, Kuhn says, impossible to do science without a paradigm. Furthermore, so deep is the influence of one's paradigm in the construction and interpretation of evidence, that it is incommensurable with other distinct paradigms, meaning that there is no common ground upon which meaningful discourse between paradigms can occur.

While the majority of scientific practice can be described a "normal" "puzzle-solving" enterprise in which a single dominant paradigm constrains how experimental evidence is deployed to settle disputes, normal science can arrive at periods of "crisis" in which disputes are the result of competing incommensurable paradigms. When a new paradigm succeeds at replacing the previous one, a scientific revolution has occurred. An important and controversial consequence of this account of scientific revolutions is that, given the incommensurability of the paradigms involved, there is no rational way to adjudicate between them. The process of a scientific community's movement from one paradigm to another is a social affair, not one based on appeals to evidence or rational argument, and a revolution typically concludes when the members of the old guard retire, die, or simply splinter off into a new discipline.

The consequences for Kuhn's account are disputed, but none quite so much as the notion that a social constructivist account of science is essentially a sort of epistemological relativism. Kuhn apologists, and even Kuhn himself consistently maintain that he never advocated for a realtivist or irrationalist depiction of science (Hoyningen-Huene 1993, Kuhn 2000[1983]), but others (e.g. Sankey 1993) maintain that Kuhn's attempts later in his career to distance himself from relativism and irrationalism involve "major transformation(s)" of his initial (and most influential) account given in *The Structure of Scientific Revolutions*. I'll leave such issues of Kuhn scholarship alone and focus on the received "Kuhnian" caricature that has influenced the constructivist anti-realist position and characterized a post-positivist intellectual climate that was ultimately inhospitable for a progressivist theory of scientific fictions.

Constuctivism, especially social constructivism — the position that scientific "truth" is not discovered, but rather an invented product of convention— is, if not essentially a form of relativism, problematically indistinct from it. If we are to take early Kuhn (ca. *The Structure of Scientific Revolutions*) at face value, particularly concerning the deep and ubiquitous influence of paradigms and their incommensurability with one at other, then the paradigm shift that occurs in a scientific revolution cannot be said to constitute progress without begging the question of which paradigm is more correct. From the point of view available within any paradigm, a conflicting paradigm is quite literally incomprehensible. The illusion of progress, if there is one, is nothing other than our tendency to reconstruct the history of science from the perspective of our own paradigm, in light of which it would be no small wonder that previous, pre-revolution theories appear inferior to post-revolution ones. Proponents of a Ptolemaic cosmological system are as certain of its superiority over the Copernican model as Copernicans are of the opposite. Priestly would have found "oxygen" an incomprehensible concept for his study of dephlogisticated air, just as Lavoisier could not understand how to coherently use the concept of "phlogiston" in his work with oxygen. If theories can only be adequately understood from within their own paradigm, then there is no way to compare pre- and post-revolutionary science, effectively, no way to say that today's theories are any better than another era's.

This relativist reading of Kuhn's constructivism is, as mentioned, not the only available reading, but it isn't all that far fetched, and the mere fact that Kuhn was pressed to deny it[5] speaks volumes. That the theory ladenness of observation renders the logical positivist project of translating scientific discourse into a language composed exclusively of analytic relations of ideas and empirical observations implausible, and that the work "paradigms" do in guiding research is deeply holistic and indispensible seems hard to deny. That this commits philosophers of science to epistemological relativism and to the position that pre-revolutionary and post-revolutionary paradigms are incommensurable seems considerably more extreme and even implausible.

[5] In a postscript to the 1969 edition of *SSR,* Kuhn addresses the accusation of relativism and the issue of progress saying "Later scientific theories are better than earlier ones for solving puzzles *in the often quite different environments to which they are applied*. That is not a relativist's position, and it displays the sense in which I am a convinced believer in scientific progress." (p206, emphasis mine) Without the italicized qualification, I would be inclined to believe his claim that his position is not relativistic. But with it, it seems fairly clear that current scientific theories are superior to previous ones only *relative* to the context of their environment.

Nonetheless, it presents a significant challenge for anyone who wishes to retain even a whiff of scientific realism. Between the Kuhnian constructivist position, and those like Feyerabend who were advocating epistemological “anarchism” (1975, the seeds of which grow out of arguments concerning theory ladenness of observation in Feyerabend 1962), a “post-modern” philosophy of science was emerging out of the ashes of logical positivism. Additionally, a significant presence of historical and sociological approaches in philosophy of science was also on the move, and the trendy (and perhaps even justifiable) methodological relativism with which these disciplines approached their subject matter was becoming a general epistemological principle in sociologically inflected philosophy of science (see, for example, Collins 1981). To be sure, positivist philosophers had been guilty of convenient reconstructions of the history of science that reinforced their idea of an empirical science shedding the superstition and mysticism of metaphysics and its physical models of unseen causes. But we should be equally wary of an image of science produced by social sciences, especially when that image appears to commit the same crime of being influenced by its own methodology, one ironically intended to avoid ethnocentric distortion.

That this depiction of science is inhospitable to scientific realism (or at least a form of realism that is not trivial) is fairly obvious; that it is inhospitable to a theory of epistemically productive fictions is slightly less so. While logical positivism was dismissive of fiction to the point of banishing it from meaningful scientific discourse, constructivism’s openness to “irrationality” may permit talk of fictions but seems to preclude their putative productivity by calling into question the possibility of any genuine scientific progress. Furthermore, we may wonder what would set ficitons apart from "non-fictional" elements of a scientific "paradigm." That fictions could contribute to what merely *appears to be* "progress" would be no more surprising than the contributions of earnest theoretical hypotheses under the constructivist model.

Getting Past the Debate

To be fair, the preceding account of 20th century philosophy of science is a caricature, one structured around the most extreme interpretations of the most extreme positions. However, my account of logical positivism and constructivism doesn't stray far from the received accounts, and is meant to highlight the most significant challenges each poses to scientific realism. The purpose of the preceding historical reconstruction of the debate between realism and anti-realism is to show that for most of the 20th century, the prospect of recognizing, let alone developing an account of how fictions can be a productive element of scientific progress was unlikely. The early part of the debate is dominated by logical positivism and realist responses to it. Logical positivism is content to identify many fictions in scientific theories, but dismisses them and seeks to eliminate them in favor of an scientific language whose meaning is determined empirically. The abductivist response from the realist camp is compelling for those who already have realist leanings, but is question begging from the empricist perspective. Considerable ground can be made against logical positivism by pointing out the theory ladenness of observation, but this ground comes at the expense of conceding ground to constructivist anti-realism, which came to dominate the conversation in philosophy of science for much of the latter half of the century. And we have just indicated the way that constructivism is inhospitable to a theory of epistemically productive scientific fictions.

By the late 1970's and early 1980's, many philosophers of science were beginning to come to the conclusion that the debate between realism and anti-realism was an unadjudicable stalemate, at least on the terms in which it had been argued for the past 60-or-so years. Particularly in the 1980's, a wide variety of approaches emerged which were undoubtedly informed by, but which blurred the lines between the camps which had come to define contemporary philosophy of science. In general, post-realism/anti-realism philosophy of science is characterized by a comparatively moderate spirit, distinct from the excesses of full-blown scientific anti-realisms and metaphysical realism and resembling a more sophisticated return to the basic position of naive realism (without the naivete) (*c.f.* Fine 1984a & 1984b; Giere 2006). These approaches are too diverse to articulate a unified "post-realism/anti-realism" philosophy of science, but in the remainder of this chapter I would like to identify two important insights that inform the approach to a contemporary theory of fictions in science that I will take in the remainder of this dissertation.

Cartwright's Simulacrum Account of Explanation

Nancy Cartwright's 1983 collection of essays, *How the Laws of Physics Lie* offers a different way of re-imagining anti-realism. Cartwright makes some notable departures from empiricism, perhaps the most obvious of which is Cartwright's willingness to believe in unobservable entities.[6] "I am not concerned exclusively with what can be

[6] I did not previously emphasize one of the most significant aspects of the received view of van Fraassen's constructive empiricism, that scientists should be agnostic about the existence of unobservable theoretical entities, for two reasons. Firstly, I am not inclined to agree with this point, so I do not intend to carry it forward in the subsequent analysis of this dissertation. Secondly, it's not entirely clear that this is actually a necessary consequence of van Fraassen's position. It is true that belief or disbelief in theoretical entities goes beyond the scope of the aim of science for the constructive empiricist. Accordingly, a scientist need

observed. I believe in theoretical entities and in causal processes as well. […] All sorts of unobservable things are at work in the world, and even if we want to predict only observable outcomes, we will still have to look to their unobservable causes to get the right answers." (Cartwright 1983, p 160)

How then, can Cartwright claim to be an anti-realist? She does this by through recasting how it is that we ought to think of scientific laws, making a distinction between "fundamental" theoretical laws and "phenomenological" laws. The fundamental laws, Cartwright tells us, are not true. They're not even approximately true, or true most of the time. Phenomenological laws, on the other hand, are true (or at least empirically adequate), and so Cartwright is a "realist" about these. To understand what Cartwright means by "fundamental" and "phenomenological" laws, and why she claims to be a realist about the latter, but an anti-realist about the former, it is important to note the bold position she takes on the relationship between explanation and truth.

Contemporary empiricists, such as Bas van Fraassen, seem to hold that the pragmatic virtue of explanation in theories can be exercised in harmony with the ultimate goal of empirical adequacy. Even though he acknowledges that explanations assert some sort of causal mechanism (van Fraassen 1980, pp124-126) and belief in the truth of those causes goes beyond the scope of science, a good explanatory theory can still make true predictions about observables. Most causal accounts of explanation maintain that a causal explanation must be true (in some strong sense) in order to be a good one, but Cartwright disagrees. "Scientific theories must tell us both what is true in nature, and how we are to explain it" but unfortunately these two functions are quite distinct, and in

not take any metaphysical position in order to gauge the empirical adequacy of a theory, and may remain in an agnostic state of non-belief. However, this does not preclude a scientist's believing in theoretical entities for other, non-scientific metaphysical reasons.

fact opposed to each other (Cartwright 1983, p 44). Fundamental laws offer effective explanations; they group diverse phenomena together as instances of similar causal structures and represent these phenomena in a mind-sized way so that their salient features can be identified and rendered mathematically tractable. But such laws come with *ceteris paribus* qualifications, that is to say, assumptions that other causal structures at work in real phenomena can be omitted, and these *ceteris paribus* conditions hardly ever obtain in real phenomena. These laws can, in principle, be adjusted to be more empirically adequate to real phenomena, and if this is accomplished they would be good phenomenological laws. But making a law more true makes it incredibly complex. Oftentimes, these more “realistic” mathematical laws are so complex that they cannot be analytically solved for any particular case. Sometimes they are so complex that they can’t even be written down. The closer to the truth they get, the worse they function as explanations. There is a seemingly inevitable tradeoff between explanation and truth, such that we may say that “the truth doesn’t explain much” (Cartwright 1983, pp 44-73).

Nonetheless, Cartwright points out that science has historically shown a tolerance of, if not a preference for fundamental laws, and for good practical reasons. Accurate phenomenological laws are difficult to come by, impossibly complex, and it would be the mark of a poor science to necessitate a new phenomenological law for every particular system one wishes to study. Furthermore, where empiricists tend to subsume the wide variety of scientific goals under the effort to make accurate predictions of observable conditions, Cartwright sees a multiplicity of potentially incompatible goals that depend on the particular problems that motivate our inquiry.

> We may wish to calculate a particular quantity with great accuracy, or to establish its precise functional relationship to another. We may wish

> instead to replicate a broader range of behaviour, but with less accuracy. One important thing we sometimes want to do is to lay out the causal processes which bring the phenomena about. (Cartwright 1983, 152)

How can we maintain a reasonable continuity when coordinating such diverse goals? Cartwright turns to models as the key to what she calls a "simulacrum account of explanation." "To explain a phenomenon is to find a model that fits it into the basic framework of the theory and that thus allows us to derive analogues for the messy and complicated phenomenological laws which are true of it." (Cartwright 1983, 152) More often than not, a model that can do this sort of work must possess properties that are not possessed by the actual system they are modeling. Sometimes they must possess properties that it would be impossible for the real system to possess. That this is compatible with a fictionalist account should be clear enough, even if Cartwright didn't come right out and explicitly say that "A model is a work of fiction." (Cartwright 1983, 153).

Thus for Cartwright, models mediate between the fictions of fundamental laws and the non-fictions of accurate phenomenological accounts. "The route from theory to reality is from theory to model, and then from model to phenomenological law" (Cartwright 1983, 4). Her account of just how this mediation works is less than straightforward, and it's fairly clear that she doesn't think this happens in only one way. In the best of cases, our explanatory models give an "as-if" description of the phenomenon that pick out its relevant features. "If the right kinds of descriptions are given to the phenomena under study, the theory will tell us what mathematical description to use and the principles that make this link are as necessary and exceptionless in the theory as the internal principles themselves. But the 'right kind of description' for assigning an equation is seldom, if ever, a 'true description' of the phenomenon studied; and there are few formal principles for getting from 'true descriptions' to the kind of description that entails an equation." (Cartwright 1983, 132-133). This is the thrust of Cartwright's anti-realism. The mediation is not a fluid translation, and our scientific explanations tend not to be true. Thus the activity of scientific theorizing is always somewhat flawed and incomplete. The spirit of this account guides the next two chapters of this dissertation, where the question of fictional representation in scientific models is explored.

Hacking's Instrumental Realism

Cartwright offers a reconceptualization of the role of scientific laws that forefronts the role of models in the mediation between fictional fundamental laws and realistic phenomenological ones and opens up a space for a discussion of a progressivist theory of scientific fictions. But an excessive focus on scientific laws and representation only tells half of the story about scientific modeling. An adequate discussion of the performative dimension of models, the way they are taken up and manipulated is obscured by such a focus. Thus, a second insight I wish to take from what I take to effectively be the inconclusive conclusion of the realism/anti-realism debate, comes from Ian Hacking. While a great deal of the discussion of scientific realism/antirealism is focused on questions of the nature and role of theoretical formulations, Hacking aims to help us recover the realism that is so intuitive for the practicing scientist. His (Hacking's) position is frequently referred to as "instrumental realism," although this title might be a bit misleading. "Instrumentalism" can be understood as an anti-realist position, essentially claiming that something, say a theoretical entity or principle, is *merely* useful, or instrumental, but not necessarily real or true. Hacking's instrumental realism is not merely a blanket denial of the instrumentalist's claim that instrumental utility fails to provide warrant for belief in the literal truth of theories, for he has a very specific type of instrumentality in mind; the instrumentality at work in experimentation. Those entities that are mobilized as instruments of investigation in the laboratory can be, Hacking argues, considered real. Hacking's instrumental realism is frequently represented by the slogan, "If you can spray them, they're real" (Hacking 1983, p 23). Thus, like Van Fraassen's constructive empiricism," Hacking's "instrumental realism" seeks to achieve a novel synthesis of sorts.

Understanding this slogan demands that we understand it as the culmination of Hacking's critique of philosophy of science's focus on scientific representation at the expense of the practical manipulations and interventions of the laboratory. Indeed, the vast majority of the preceding account is primarily, if not exclusively, concerned with an understanding of theories and theoretical entities as representations of what is going on in observable phenomena. The battle between realism and anti-realism thus ends up being a struggle to determine whether or to what extent these representations could ever be considered to be accurate depictions of reality. In *Representing and Intervening* (1983), Hacking concludes that if we insist on the representationalist bias exhibited in the terms this debate has been conducted, then it is necessarily inconclusive, or, at the very least, concludes with a picture of science that is unrecognizable to those who practice it. "If there were a sharp distinction between theory and observation, then perhaps we could count on what is observed as real, while theories, which merely represent, are ideal. But when philosophers begin to teach that all observation is loaded with theory, we seem completely locked into representation, and hence into some version of idealism." (Hacking 1983, p 130)

Where, if anywhere, is realism to be found? Merely turning to experimental practice does not, all by itself, reveal the answer. We have already seen that the reliable production of "observations" of theoretical entities begs the question if taken as evidence of the reality of those entities. These entities being "observed," the ones under investigation, are still hypothetical entities. It may be said that our experimental designs produce reliable regularities by making a variety of theoretical assumptions and by manipulating unobservable, so-called "theoretical" entities. This is closer to the realism Hacking is after, but still not quite there. To say that we may infer the truth of these theoretical assumptions and reality of unobservable entities from the successful regularities produced by our experiments "gets the time order wrong," for we do not experiment first, and then believe in the theoretical entities used to create the desired laboratory phenomena. Rather, when we use an unobservable entity to influence another hypothetical entity, we are already committed to the belief that the "instrumentalized" entity is real. "We are completely convinced of the reality of electrons when we regularly set out to build— and often enough succeed in building— new kinds of devices that use various well understood causal properties of electrons to interfere in other more hypothetical parts of nature." (Hacking 1982, p153)

It must be said that there are aspects of Hacking's argument that are suspicious. Despite his protests that his instrumental realism is not essentially an abductive inference drawn from our ability to successfully manipulate certain unobservable entities, nor is it description of an unavoidable psychological tendency of scientists to believe in the objects they are ostensibly working with (one which is, as van Fraassen would say, a pragmatic virtue for experimental practice, though not warrant for serious metaphysical belief), it is difficult to see how else the argument for his position could work. This isn't to say that his position isn't attractive, only that the details of the supporting argument are opaque and his credulity regarding entities that can be "sprayed" is arbitrary for those of us who have already become skeptical about theoretical entities.

However, Hacking does seem to be pointing out something important about how far our skepticism might be reasonably extended. It may have once been reasonable to be skeptical about electrons, or neutrinos, when these were hypothetical entities being studied by scientists. However, once we began using these particles to study other objects, having a realistic attitude about them becomes as natural as being a realist about tropical fish. No miraculous advance occurred in our theoretical representations of these particles, but an important practical shift occurred. This practical shift requires a shift in attitude that makes skepticism silly if not impossible. The implications of Hacking's refocusing our attention to experimental practice are somewhat mysterious, but significant.

I am cautious about accepting the doctrine of instrumental realism whole-hog, but the assertion that we ought not get lost in the world of representations at the expense of neglecting how things are tools not only for thinking, but for *doing*, is one that I will try to follow. First and foremost, we should keep this in mind in the following chapters as we turn our attention to models which figures like Cartwright indicate as integral to the relationship between potentially fictive representations and their real targets. Hacking seems to think that models belong squarely in the domain in representation, and that while experimenters must be realists about the entities that they "spray," modelers are generally anti-realists about their models. I'm not so sure that the distinction between modeling and experimenting can be drawn so cleanly, especially when we begin to look at simulations. Hacking is far too quick to relegate models to the domain of representation and identify them with theory construction. This isn't altogether wrong, since models are, in fact representations. But they are representations (or, to use Morgan and Morrisson's term, 'representatives') that we do intervene on and manipulate. The big difference is that these are virtual interventions on virtual systems. Intuitively, we want to refrain from calling the entities that we virtually "spray" real, but Hacking does not give us any tools to distinguish between virtual interventions and real ones while also appreciating similarities between the two. Thus Hacking is an ambiguous ally for me, whose insights about experimental practice are useful tools for addressing models and simulations, but become complicated in this context. I will return to this in the later chapters of this dissertation, when it has become clear that simulations are fictitious performances, which involve interventions and reliable manipulation of models, and which are difficult to distinguish from the manipulations employed in the artificial confines of the laboratory. Such performances use models as props, whose quasi-reality demands that we quasi-believe in them.

Chapter 2
Models and Representation

In the preceding chapter I have sketched out the broad contours of a debate between scientific realism and scientific antirealism, arguing that this debate created a climate that was inhospitable to the development of a progressivist conception of scientific fictions. Despite the fact that the general debate stalls out and cannot be resolved without question-begging assumptions, there appear to be certain significant elements of scientific inquiry which refer to entities that don't exist (or at least don't exist in the way they are described), and the recognition of these elements does not necessarily commit one to a broader anti-realist position. Most notable of this class of scientific objects, and where a key overlap between realist and ati-realist intuitions lie, are scientific models. Scientific models are exemplars of the sort of representations that scientific realists and anti-realists alike would freely admit play a significant role in scientific theory construction and practice, but do not necessarily aim at truth in any strict sense.

Ontology

What sorts of things do we mean when we talk about scientific models? By way of examples, the extension of this term seems impossibly diverse for the purposes of discussing the ontological status of models in general. We may consider the Bohr model of the atom, Schrödinger's equation as a mathematical model of the probability distribution for an electron in a hydrogen atom, a model of the double helical structure of DNA (in either its abstract or physical instantiation), a balsa-wood scale model of a steel bridge, a physical scale model of a dynamic target like an ecosystem, or so-called "model-organisms" like *Drosophila melanogaster*. It is my intention to make an argument that many models qualify as fictional distortions of their targets, but exactly what this means and why anyone should believe it won't be clear until the conclusion of the next chapter. For the meantime, let us see why the question "what sorts of things are models, and what sort of relationship do they have to their targets?" is a difficult one.

The first part of this question, "what sorts of things are models?" addresses the ontology of models, and appears to have a dual meaning. We might take it as suggesting a domain of possible responses determined by the opposition between the real and unreal, or alternatively between the abstract and the concrete. On both counts, models seem to defy categorization, running the full spectrum between real and unreal, abstract and concrete.

With respect to reality and unreality, let us examine the use of a pendulum to model another harmonic oscillator like a weight bobbing on a spring. We might do this in a variety of ways. We may actually construct a pendulum, in which case the model itself is uncontroversially real. We may simply imagine a pendulum, which is bit harder to classify as real, since it is imaginary; but real pendula do exist (i.e. they are not like unicorns, for example) so we would have to qualify any assessment of real or unreal imagined objects, or introduce a range of intermediaries between the two. Additionally, our imaginary pendulum could involve any number of distortions that cause it to differ from any particular actual pendula. For instance, it may involve a point-mass at the end of a massless string, which oscillates without any energy loss due to friction and swings through a perfectly uniform gravitational field. This pendulum could not possibly exist in reality. It is quite clearly as unreal as a unicorn (perhaps even more-so if we judge certain features of it to be impossible in principle). All of the above examples are models in what appears to be the same sense, and yet they span the range of possibility between real and unreal.

Some indication of how models should be classified along the concrete/abstract dimensions of our ontological question should be already evident from the above. Those with materialist inclinations may, in fact, collapse this analysis with the previous one. However, while material concreteness does seem to be a plausible sufficient condition for being real, it is not altogether clear that abstract models are unreal. Nor is it clear that materiality is all we mean by concreteness. We might just as well understand the abstract/concrete distinction to refer to a distinction between universals and particulars.

Regardless of how we understand the relationship between real/unreal and abstract/concrete, it should be evident that models can't be pinned down as generally abstract or concrete. Our constructed pendulum is concrete in the material sense, while the imaginary ones are abstract (with the idealized pendulum perhaps more abstract). All are abstract representations of any harmonically oscillating system in general, that is to say that they refer to a large class of systems, but this doesn't settle the matter for models in general. A scale model of a bridge or ecosystem is both materially concrete and may also refer us to a single particular target, and both model and target have particular values for lengths and stress bearing capacities among other features. Mathematical models appear to be on the abstract end of the spectrum, in all senses of the term, but can be more concrete (in the particularist sense) when their variables are set to reflect some specific set of conditions.

All of this serves to show that at first blush, inquiring into the ontology of models tells us very little about just what models are. I will return to this question again at the end of this chapter, at which point we may be able to make a bit more headway on it, but for the time being let us allow it to stand as the first of many puzzles that models offer. Many of the puzzling aspects of an inquiry into the ontology of models stem from the fact that models are representations. The trouble we have pinning down the ontological character of a model is the same as the trouble we have pinning down the ontological character of any symbol or picture. Symbols and pictures can possess varying degrees of concreteness or abstractness (in all the senses mentioned above) and can denote existing particular objects, classes of objects, or even represent non-existing things.

Models of Theories

It is not immediately obvious, however, what sorts of things models are intended to represent. There are two likely candidates which may determine one's approach to the representational character of scientific models. Scientific models can be thought of as representing scientific theories, or as representing the world. The relationship between models and theories is not nearly as complex or puzzling as that between models and the world, but it does require some care in distinguishing a model from a theory.

A full discussion of exactly what a scientific theory is, and what makes for a good scientific theory, is beyond the scope of this project. Nonetheless, some degree of specificity regarding what we mean by a scientific theory is necessary for our purposes. In casual conversation, especially amongst practicing scientists, terms like 'model,' 'theory,' and 'hypothesis' are used interchangeably (as in "it's just a model at this stage," suggesting that the so-called model is a tentative, or hypothetical theoretical position). (Morgan and Morrisson 1999, 7) Such conflations are, however, problematic if for no other reason than their imprecision and sloppiness. Ernest Nagel writes, “The only point that can be affirmed with confidence is that a model for a theory is not the theory itself.” (Nagel 1961, 116) While I agree with the spirit of Nagel’s comment, I think we can say a bit more about models and theories with a fair degree of confidence.

For the purpose of differentiating a model from a theory, we can get away with the following stipulative definition of a theory: theories are sets of *formal statements* expressing general assertions intended to be about the world (or a portion of the world) (Morgan and Morrisson 1999, 2). The important feature of theories that distinguishes them from models is that they are formal statements, which is to say the have a necessarily linguistic structure. Models, by contrast, as we have seen from the above examples are not necessarily linguistic, and are even frequently regarded as necessarily non-linguistic. Consequently, theories, when measured against the world they make claims about, may be true or false, depending on whether or not the claims they make pan out. This gives us a decisive (though not necessarily sufficient) tool for determining which theories are adequate and which need work. Models, since they may represent non-liguistically, do not easily admit of binary classifications such as “true” and “false,” however. As we will see, something a bit more complicated and intentionally conditioned is at work when determining the extent to which a model satisfactorily represents the world.

One way of thinking of scientific theories in this way is to identify theories with sets of law-like axiomatic statements along with the logical consequences of those axioms. This is generally referred to as the "syntactic" view of scientific theories. (Morgan and Morrisson 1999; Frigg and Hartman 2006; Lutz 2010) This view is common in logical empiricist accounts of scientific theories and has fallen out of favor as logical empricism has fallen out of favor amongst philosophers of science.[7] On such an understanding, scientific theories would be defined by their logical form and are interpreted by assigning meaning to the terms of the formal axioms (which meanings are, for the logical empiricist, restricted to potentially observable referents).

Opponents of the syntactic view have offered what is usually called the "semantic" view as an alternative where models are used to fix the meaning of theoretical terms. A model of a theory is regarded as providing an interpretation of the theory by instantiating conditions in which the formal statements of a theory are true. Proponents of the semantic view (e.g. Suppes 1967; Suppe 1989; Giere 1999; van Fraassen 2000) characterize a scientific theory according to the groups of models which satisfactorily provide these conditions rather than the formal structure of the theory, where this grouping may only be characterizable as a "family resemblance" (Giere 1999, 97-117). It is noteworthy that this does not preclude the formalization of a theory into logical propositions, but it does mean that theories are interpreted through their models rather than empirical correspondence rules that translate theoretical claims directly into claims about the world (Morgan and Morrisson 1999, 3).

Following Cartwright (a discussion of whose position can be found in the previous chapter), we know that a vast majority of the formal laws that would comprise

[7] The extent to which a syntactic approach to theories necessarily requires one to adopt other problematic aspects of logical empiricism is unclear. For a fuller discussion of the rise and fall of the syntactic view, see Suppe 1977 and for a discussion of the syntactic view which attempts to disentangle it from the problems associated with logical empiricism, see Lutz 2010.

theories are not even approximately true of the real world without significant qualifications tacked on, qualifications which do not obtain in any actual phenomena. However, in a world where those qualifications do obtain, say where the Earth and moon are the only two bodies that exist and they only interact via gravitational forces, or where a pendular weight swings frictionlessly on a massless string in a uniform gravitational field, these universal laws are true. Such worlds are models; more specifically, they are models of theories. Hence, we may describe a model of a theory as an accurate model of said theory if it instantiates a system where the claims made by that theory are true, and if these these theoretical claims are not true of *actual* systems in the world, then there are potentially fictional elements in the theory's models.

By adopting these stipulative definitions of 'theory' and 'model of a theory,' it is easy to see how the relationship between a theory and a model is less complicated than that between a model and the world. The formal linguistic statements which comprise a scientific theory act, as Ronald Giere states, "as rules devised by humans to be used in building models to represent specific aspects of the natural world" (Giere 1999, 94-95).[8] Models of theories are constructed according to theoretical rules/principles in order for them to be accurate models of theories, but these rules do not exhaust all the features that may be included in a model. Theories underdetermine their models, which is to say that a model of a theory may contain features not prescribed by theoretical laws and still be a

[8] Here, Giere makes a subtle distinction between "laws" and what he calls "principles," where it is principles that are used to instruct model construction." The distinction is not a strong one, however, and seems to serve to draw our attention to whether the formal statements in question aim directly at the world or are used to construct a model. I gloss over this distinction because it seems clear in Giere's account that a scientific principle is nothing other than a scientific law understood through a model-theoretic account of scientific practice. One significant additional benefit of refering to theoretical propositions as "principles," rather than "laws" or even, *pace* the above quote, "rules," is that it avoids the suggestion that models scrupulously adhere to and follow these "laws." Instead, it seems more accurate to loosen the role of theories in model construction a bit, indicating that they guide, rather than definitively determine models.

satisfactory model of that theory (provided that the model and theory do not conflict). These extra features may facilitate the model's representation of real systems or satisfy some other modeling virtue, such as ease of construction, interpretation, or manipulation.

While the actual construction of models of theories is no easy feat, identifying when a model of a theory is successful (if we restrict "success" to the accurate modeling of a theory) is relatively straightforward; *viz*. when conditions under which the theory is true are instantiated. Modelers may compromise on this for one reason or another (explanatory power, greater similarity to actual systems, or even due to limitations of the material the model is constructed from —be it balsa wood, software, or the modeler's own imagination), but such trade-offs clearly result in the models being less satisfactory models *of theories*.

While a successful model of a theory need only instantiate conditions where the theory's statements are literally true, some further questions are implied by the above remarks. I suggested that a model may sacrifice successful representation of a theory in exchange for successes on other fronts. We may, following Ronald Giere (1999) distinguish between conceptions of models that focus on representing theories and conceptions that focus on representing actual real-world systems. This latter representational relationship is a bit more complicated than the former.

Models of the World

It should be clear that a model is not merely a *symbolic* representation. Here, a *symbol* is to be understood in distinction to an *icon* (or "likeness"), roughly the way that C.S. Peirce distinguishes the two. Symbols denote their denotata by fiat, but a model purports to represent its target in a less arbitrary way. It is more like an *icon* in that its ability to represent depends on characteristics belonging to the model itself rather than solely on the intentions of those who construct and interpret models (Peirce 1894). This appears true of the way that models model theories, given that models are non-linguisitic representations that instantiate the conditions under which a theory is true. The iconicity of models is equally revelatory in the way they represent the world.

An iconic understanding of models of real-world objects/systems seems straightforward enough for concrete, scale models like balsa wood bridges. However, many iconic representations aren't completely free of a certain arbitrary, symbolic dimension. This dimension seems to encroach as we shift to more abstract models, like mathematical representations of physical systems. Despite the fact that the equation '$x(\mathrm{t}) = \mathrm{A} \sin(2\pi ft+\phi)$' offers a mathematical representation of a simple harmonic oscillator, none of the parts of the equation bear any resemblance to the parts of an oscillating spring or pendulum system, nor does the equation as a whole bear any resemblance to any such system as a whole. The representational function of the *parts* of the equation to the *parts* of a real system depend entirely on the intentions of those who use this equation and are arbitrarily fixed.

Nonetheless, we should stop short of saying that the representational relationship between a mathematical equation and the system it models is an arbitrary one. Features of a simple harmonic oscillator such as amplitude, frequency, and the position of a moving mass at a given time are arbitrarily denoted by symbols 'A,' 'f,' and '$x(t)$,' but once these and other symbols are assigned, a relationship between them must obtain which adequately represents the relationship between the corresponding features of the system being modeled. It is, therefore, the formal relationships between the parts of the model which bear a non-arbitrary relationship to the formal relationships between the parts of the target system despite the lack of any similarity between the outward appearances of the model and target.

Idealization

Many (van Fraassen 1980; Suppes 2002; Da Costa and French 2003) have followed this train of thought to the conclusion that models, if they are to adequately represent their targets, must be *isomorphic* to their targets. Isomorphism is to be taken literally, here, as an *identity of form* (or formal relationships between parts) between model and target. This strikes me as too strong a relationship to account for all models, but we would do well to examine it closely. A standard of isomorphism for adequacy in models is particularly appealing if we focus on mathematical models, if only because mathematical models capture the formal structure of systems in a way that abstracts them from the material substance of those systems. Most mathematical models, including the one for simple harmonic oscillators mentioned above, employ observable measurable quantities as variables, and those which invoke quantities that are not measurable by direct observation (for instance, the number of moles of a gas in the Ideal Gas Law) can frequently be translated such that they are measurable by observation (by, say, making number of moles a function of mass, which can, in turn, be interpreted as a function of weight, which, for it's part can be an observable measurement made on a scale).

But we saw earlier (in Chapter 1), via Cartwright's analysis of theoretical laws, that such mathematical models of simple harmonic oscillators or ideal gasses do not accurately represent any real systems when interpreted literally.[9] Theoretical laws, represented in mathematical models, "lie," and therefore do not maintain an isomorphic relationship to their targets in the world. Through the employment of *ceteris paribus* qualifications and negligibility assumptions, these models portray *idealized* versions of their respective targets. These idealizations may be *similar* to said targets, but they can hardly be said to be isomorphic.[10] Furthermore, following Cartwright's account, it is precisely this idealization that, while it sacrifices the "truth" of the model, enables the model to have some explanatory power.

This appears to be in keeping with the iconic nature of models mentioned earlier. The vast majority of typical, stylized icons which come to mind usually employ some sort of distortion in order to convey information clearly and efficiently (e.g. gender markers for bathroom doors, a road sign for merging lanes, maps, etc). In simply noting

---

[9] The issue of literal interpretation is a potential sticking point when discussing the ways that models represent. While fundamental theories (according to Cartwright) and many models do not literally represent their targets in a true way, it is certainly possible to say that they do so *figuratively*. Nelson Goodman articulates the way in which figurative representation, such as metaphor is contrasted with fiction. The contrast is subtle, since the same thing may be understood to represent either fictionally or metaphorically. "Don Quixote', taken literally, applies to no one, but taken figuratively, applies to many of us – for example, to me in my tilts with the windmills of current linguistics" (Goodman 1978, 103). "Fiction, then, whether written or painted or acted, applies truly neither to nothing nor to diaphanous possible worlds but, albeit metaphorically, to actual worlds" (Goodman 1978, 104). I believe the gist of it is that any prospective fiction can be interpreted to represent non-fictionally but figuratively. Fiction and metaphor are compatible in that a fiction can be metaphorically interpreted, but what is implied here is that interpreting a representation as figuratively applying to some real thing is to not interpret it literally, and therefore to not interpret it as a fiction. Thus, as frameworks for analysis, fiction and metaphor are not compatible, as one asks us to interpret a model literally and the other asks us to interpret it figuratively. This is not to suggest that scientific models are not metaphorical, but a metaphorical framing gives a degree of latitude in the interpretation of models that obscures a fictional one. Perhaps the choice between frameworks is a matter of temperament, and for my own part, I have found accounts from thinkers who insist upon a literal interpretation of scientific theories (e. g. van Fraassen and Cartwright) to be particularly compelling. As this project is concerned with understanding fictional models, the prospect of figurative representation/ interpretation is bracketed. I could have just as easily used the opposite framework, but that would have yielded a very different project.

[10] This insistence on similarity over isomorphism as the key relationship between models and their targets (asserted in Giere 2004 and Teller 2001) makes a two-fold break from isomorphic accounts, a break from both the *iso-* and the *morphe-* components of the term. Similarity requires neither perfect correspondence, nor is it restricted to the formal aspects of models and targets. Instead, it is strikingly non-specific regarding both the degree and respect of the correspondence between two objects that are said to be "similar."

that many models give idealizaed representations of their targets, however, the nature and degree of the distortions involved in idealization remain uncharacterized. What, exactly, do we mean when we say that models represent their targets by idealizing them? Do we even mean one thing?

Discussion of idealization in philosophy of science literature generally makes a distinction between two sorts of idealization, one which omits features which are present in the target phenomenon, and one which adds distorting features to the target phenomenon. I follow Michael Weisberg by calling these *minimalist* idealizations and *Galilean* idealizations respectively (Weisberg 2007).[11]

Minimalist idealizations, sometimes referred to as "abstractions" (Cartwright 1989), "negligibility assumptions" (Musgrave 1981), or "method of isolation" (Mäki 1994) seek to portray a simplified version of a phenomenon. Cartwright's discussion of modeling the behavior of massive bodies by way of Newton's universal law of gravitation offers a nice example of this species of idealization. When we represent the spatial relationship between, say the earth and the moon, by way of the attractive force due to gravity between two masses at a distance, we omit several features of the real system such as electrostatic forces as well as gravitational forces between the earth or moon and other celestial bodies. These features, present in the target system, are missing from the model. This simplification is frequently justified by way of an appeal to the claim that what is left out has a very small effect on the real system, and this effect is

[11] "Galilean idealization" is a term Weisberg takes from Ernan McMullin's extensive work on Galileo's method (see McMullin, 1985). Minimalist idealization is discussed extensively by Cartwright (1983) but under the term "abstraction." While I have no particular attachment to either term, I have already used the word 'abstraction' in a less precise sense earlier in this chapter. I use "minimalist idealization" solely to avoid confusion. It should be noted that Weisberg also identifies what he calls a thrid form of idealization, which he refers to as "multiple models idealizations." I ignore this "third kind" here because it concerns the relationships between several models, rather than the sorts of distortions involved in a single given model. However, there is a breif discussion of the use of multiple models in chapter 5.

negligible for the purposes of our model. We also, incidentally, omit other features of the real system, like the color and chemical composition of the earth and moon, and these may be left out because they play no causal roles whatsoever with respect to the aspects of the system we are interested in. What is revealed in this sort of idealization is that simplicity is a virtue for some models, and we are willing to sacrifice the accuracy of our model, even in cases where the omitted features do play a causal role in the target phenomenon, provided that the ensuing inaccuracies are relatively small.

Galilean idealizations are often distinguished from minimalist ones in that they model by adding distorting features to the target system, rather than omitting features. The name for this style of idealization comes from McMullin's analysis of Galileo Galilei's experimental methodology (see, for instance, McMullin 1985). Galileo's own writing on the behavior of falling bodies is instructive here:

> We are trying to investigate what would happen to moveables very diverse in weight, in a medium quite devoid of resistance, so that the whole difference of speed existing between these moveables would have to be referred to inequality of weight alone. Since we lack such a space, let us (instead) observe what happens in the thinnest and least resistant media, comparing this with what happens in others less thin and more resistant.[12]

The idealizations being discussed here are complex, but we can see that something different than simply ignoring a feature of the target system is at play. There is reference to a frictionless medium that does not exist, and approximations of that frictionless medium by way of constructing experiments in successively "thin" media. This is not simply the omission of the medium, or the causal role of friction, but rather an experimental distortion of the medium by way of portraying it as different than it actually is in ordinary cases. Such distortions are more clearly distinct from minimalist

[12] Quoted in Weisberg 2007

idealizations in other cases. For instance, in the case of the model of the earth and moon discussed earlier, we may also treat the earth and moon as perfectly spherical. To do so is to distort our target system in a way that differs from the way we omitted electrostatic forces, or gravitational forces between the earth/moon and other heavenly bodies. Distorted features are not left out, they are represented as different than they actually are.

Whereas minimalist idealizations are motivated by the value of simplicity for models (a value that is frequently justified by reference to explanatory virtues – *c.f.* Strevins 2004; Cartwright 1989; Batterman 2002; Hatmann 1998), Galilean idealizations are motivated by a practical value for models, one of tractability. Analysis that may have been difficult or impossible for a more faithful representation of a phenomenon are easier in a distorted one. Continuing with the example of a model of the relationship between the Earth and moon, calculating gravitational forces between point masses or perfect spheres is far simpler than for irregular spheroids. As was the case for minimalist idealizations, these distortions, while they have a relatively big payoff with respect to tractability, are supposed to bring with them a relatively small loss in accuracy, one which, if the idealized model is an adequate one, is negligible for the purposes it is being used.

It is worth pointing out that minimalist and Galilean idealizations are not mutually exclusive. In fact, as is probably evident from the above examples, the distinction between the two is a vague one and there are many cases where determining whether an idealized model is minimalist or Galilean appears to be a fairly arbitrary and inconsequential matter. What, we might ask, is the difference between a model that distorts a real system by instantiating “the thinnest and least resistant” medium through which a moving body travels to approach frictionless motion, and one that simply omits the medium under the pretense that its effect is negligible for our purposes? While not every Galilean idealization qualifies as a minimalist idealization (e.g. treating the earth and moon as perfect spheres rather than irregular oblate spheroids omits nothing), it would seem that most, if not all minimalist idealizations could count as limit cases for extreme distortions in Galilean idealizations. The electrostatic forces between the Earth and moon can be omitted if we distort our model to portray each as neutrally charged, and the gravitational effects of other celestial bodies can be omitted if we distort our model to portray them as infinitely distant from the bodies under investigation. Accordingly minimalist idealizations can be regarded as a special cases of Galilean idealization; specifically cases where the feature in question has been distorted to the point of negligibility.

Weisberg concedes this point, but maintains that an additional distinction can be drawn regarding whether and how these different styles of idealization become de-idealized. Minimalist idealizations, he claims, given that they are motivated by a desire for simplicity (in turn motivated by explanatory virtues) tend not to be de-idealized. Galilean idealizations, on the other hand, are often a first run at modeling that can be adjusted as experimental and analytic tools develop. In these adjustments, the distortions that were once introduced for the sake of tractability can be, and frequently are corrected. "Galilean idealization takes place with the expectation of future de-idealization and more accurate representation" while minimalist idealization does not (Weisberg 2007, 4).

As distinctions between representational styles go, this is a fairly weak one. It focuses more on future intentions than on the relationships between model and target as they currently stand.[13] We might be inclined to draw a stronger distinction than this by attending to *how* these models are de-idealized. Despite Weisberg's claim that minimalist models tend not to be de-idealized, they certainly can be. In addition to the appeal that the omission of features of the target system from the model makes little difference, some might also claim that any of the omitted features (provided that they are known) can easily be added back to the model in order to make the model as accurate as we need it to be. In other words, the idealized model can be de-idealized as needed if all that the idealization entails is the leaving out of features which can be accounted for and re-introduced. But de-idealizing a minimalist idealization is a necessarily discrete process involving the addition of features omitted in the idealized model; either a feature is in the model or it isn't. De-idealizing the distortions in Galilean idealizations is a potentially continuous process by which a distorted feature of the model is made to be more like the corresponding feature of the target system; the degree of similarity between model and target with respect to this feature can be arbitrarily set at any point in the continuum of distortion, as practical needs dictate.

However, if we were correct in saying that all minimalist idealizations can be regarded as special cases of Galilean idealizations, then we could also conceivably de-idealize a minimalist model by re-introducing omitted features in a distorted form, and gradually un-distorting them to make our model more like its target. A putative strong distinction between minimalist and Galilean idealizations seems to hinge on whether or not there exist any cases of omission that cannot be re-interpreted as limit cases for extreme distortion. In the absence of any clear cases that come to mind, I'm prepared to

[13] This is not to suggest that the intentions of modelers, particularly whether they ever intend to correct the distortions in their models, is not a significant factor in understanding how models work. This idea will be a recurring theme in later chapters of this dissertation.

accept that no such cases exist. Beyond this, distinguishing between minimizing a feature of a model by omitting it, or by distorting it to the point where its effect is infinitesimal gets us into the sort of haggling over the distinction between negligible presence and outright absence that I would just as soon avoid, and which has no practical consequence for this discussion.

Furthermore, the respective values that motivate each type of idealization hardly exclude one another, as simplicity is not merely an explanatory virtue, but has a tendency to promote tractability as well (consider how much more difficult assessing the gravitational forces at work in a three body system is in comparison to a two body system). Therefore, I expect that the preceding discussion presents a convincing enough argument that "minimalist" idealization need not be seen as anything more than a special case of "Galilean" idealization, insofar as the potentially fictional distortions present in each can be "de-fictionalized" as our needs for precision and accuracy demand.

Caricatures and Analogies

Nonetheless, there does seem to be something worth pursuing in this question of whether and how the distortions present in models are removed, whether they can be easily de-idealized, or perhaps more broadly, "de-fictionalized."[14] While a strong distinction on this front appears to be elusive in the case of kinds of idealization, there are other ways that models represent that are far more difficult to "de-fictionalize." The inaccuracies of models may not be of much concern if one can always make them more accurate. This hinges on our ability to clearly identify which features have been distorted

[14] "Fiction" is being used in an intentionally loose way here. This concept will be tightened up in the next chapter.

or omitted and how we would need to adjust our model in order to make it a more faithful and accurate representation of its target. This is not always so easily achieved. Moreover, the motivation behind idealization, discussed above as driven by simplicity and analytic tractability seems not to apply to all models and, more consequently, misses the essence of modeling practice in general.

Caricature models and analogue models are instructive on this point. Both seem to be structurally reducible to the sorts of idealizations discussed above (insofar as they involve omissions and distortions) but aim at slightly different purposes. Caricatures are usually taken to be a combination of minimalist and Galilean idealizations in which certain features are omitted and others are distorted into extreme cases (Gibbard and Varian 1978, Reiss 2006). While we may entertain arguments for regarding idealizations as approximately true (provided that they can be de-idealized as needed), it is less plausible to make such arguments about caricatures, given the complex and extreme nature of their distortions. Perhaps more puzzling, if we take non-scientific caricatures as our basis, caricatures seem to have a tendency to represent their targets surprisingly well *despite* the extreme nature of their distortions. Caricature artists depict easily recognizable images of people by over-emphasizing features (like Barack Obama's ears, Jack Nicholson's eyebrows, Jimmy Durante's nose, or Jay Leno's chin). Similarly, scientific caricatures over-emphasize certain features of a target system, often to the point of excluding confounding variables.

Consider Ackerlof's (1970) model of the car market which overemphasizes the role of asymmetric information between sellers and buyers in order to investigate the well-known price differential between new and used cars, or Schelling's (1971) segregation model that overemphasizes a mild preference for not being the minority in an overwhelming majority to investigate the phenomenon of geographic segregation in housing markets. In both cases, easily recognizable effects (depreciation of the value of a new car and geographic segregation) are represented as emerging from by easily recognizable conditions (asymmetric knowledge between sellers and buyers and mild preference for not being in the minority).

While these caricatures can be explained solely in terms of omitted and distorted features in an idealized model, de-idealization/de-fictionalization seems to be a bit more out of reach than in our previous examples. For one thing, the caricature is a complex combination of omissions and distortions, one which may be somewhat difficult to untangle. As in the case of artistic caricatures, an incomplete step-wise de-fictionalization has the effect of making the representation in question *less* recognizable before it becomes *more* recognizable.[15] Secondly, in cases where the structure of the target system is relatively poorly understood (which cases are, in fact, constitutive of the vast majority of cases in ongoing research programs), even identifying the respects in which the model differs from its target becomes a difficult task.

Gibbard and Varian offer an explanation for this aspect of caricature models, which would otherwise appear to be a shortcoming if models are supposed to use relatively negligible distortions and omissions for the sake of simplicity and analytic tractability.

> The difference between applying a model as an approximation and applying it as a caricature lies in the intentions of the investigator: a caricature involves deliberate distortion to illuminate an aspect of […] life. If the uses of deliberate distortion are ignored, and the job of applied models is taken to be no more than accurate approximation under

[15] The phenomenon known as the “uncanny valley” is a good illustration of this.

> constraints of simplicity and tractability, many of the caricatures economic theorists construct will seem unsuited for their job. (Gibbard and Varian 1978)

Analogue models offer a slightly different route to the same sort of conclusion. Analogue models substitute one relatively concrete system for another, as in the hydraulic model of a circuit (Esposito 1969), the billiard ball model of a gas (Hesse 1966), the liquid drop model of the nucleus (Frigg and Hartman 2006), the MIT bag model of quark confinement (Hartman 1999, 334-340), or an oscillating spring and mass system as a model of another oscillating system like a pendulum. In such analogies, by substituting one concrete system for another the models in question do not merely omit or distort features of their targets, they also add wholly new features.

At first blush, such analogies are merely concrete instantiations of the sort of formal reductions discussed above in terms of isomorphic models. While it was shown that isomorphism is too strong a relation to use for *all* models, it may be argued that analogue models are similar to their targets by virtue of a two-step formal reduction which fits our analysis of minimalist models. Taking the oscillating spring and mass system as an analogue of a pendulum, one might say that both the model and target system can be represented by the same mathematical representation, which representation is nothing more than a formal model, i.e. a minimalist idealization of both systems.

However, to take this as an indication that analogue models are nothing more than complicated concrete versions of essentially minimalist idealizations is to miss a significant point. It assumes that the only active features of an analogy are the similarities between analogues. In fact, the differences play a role as well, specifically by drawing our attention to certain salient features of the model and target. When one makes an analogy, she implicitly states that features not shared by the analogous systems are not important. Mary Hesse (1966) addresses this in terms of a distinction between positive and negative analogies (where positive analogies refer to the ways in which analogous systems are the same and negative analogies refer to the ways in which the systems are different). Hesse suggests that negative aspects of an analogy can be discarded as insignificant, but this is a bit heavy handed. Negative analogies are significant in analogue models because they make a claim about the sorts of features that can be omitted or distorted in an idealization. They direct our attention to certain features and away from others.

Furthermore, analogue models are susceptible to previously mentioned problems regarding target systems that are poorly understood; in any active research program, we may be uncertain which features constitute positive analogies and which ones constitute negative analogies. Hesse addresses this via another distinction, identifying what she misleading terms “neutral” analogies. Neutral analogies are, in fact, either positive or negative but concern features for which the similarity or difference between the model and target are unknown and need to be explored.[16] Furthermore, while the question of whether or not a feature of an analogue model is actually similar or dissimilar to its target is relatively straightforward, dissimilarities can take many forms. In a neutral analogy it is unclear whether features in a model are problematically dissimilar to those of the target, justifiably distorted representations of features in the target, are formally similar, or don’t actually correspond to any features of the target system at all. Answering this sort of question depends not only on how much we know about the systems taking part in

[16] In this way, analogies direct our investigations, laying out non-arbitrary trajectories for extending models to novel contexts.

an analogy, but are also dependent on the context of the model and the intentions of the modeler.

Exemplification

At this point, two issues which have hitherto been running parallel to one another in this discussion, namely the way models are *non-arbitrary* representations of real-world targets, and the motivation to use of models that are inaccurate representations of those targets, merge. . We noted that models are neither purely arbitrary symbols of nor are they identical to their targets. It was determined that isomorphism, while tempting for many mathematical models, was too strong a relation to describe all models. Instead, a weakened relation of similarity, in which models are somewhat similar and somewhat dissimilar to their targets obtains. But similarity is an incredibly vague relation, and it may be noted that *any* two things are trivially similar in *some* respects and dissimilar in others. Mauricio Suárez (2003) (taking significant cues from Goodman 1968, 4) makes a somewhat stronger argument against similarity as *the* standard of scientific representation (of which models are a species), noting that similarity is a symmetrical relation (if A is similar to B, then B is similar to A) and also a reflexive relation (A is always similar to itself). These relationships do not seem to characterize models particularly well. If a model provides some sort of representation of some real-world target, it hardly seems appropriate to say that the real-world represents the model as well, or in other words that the target system models the model.[17] Furthermore a model is not really a model of itself.

[17] To do so would be to confuse the two distinct ways that models "model" theories and/or the world. The world may "model" a model the way a model "models" a theory, but not the way that a model "models" the world. This is just exploiting a terminological ambiguity. It does not address Suarez's central point, that the modeling relationships are asymmetrical, while similarity is symmetrical.

So the fact that A is similar to B cannot be a sufficient condition for saying that A models B,[18] and while some degree of similarity may be a necessary quality, without some sort of constraining specification that can act as a standard for the degree or respect of similarity between a model and its target, this necessary condition is fairly useless.

Suarez calls for a necessary intentional component to scientific representation which would capture the essential directionality of representations like scientific models. Yet, if we were to take human intention as both necessary and sufficient, we would be back at symbolic representation, where a model would represent their targets merely by our designating them as models of their target. What we are looking for in our attempt to capture the essence of the representational structure of models is something that offers a specification of the type or degree of similarity which is less strict than isomorphism, and which can account for the asymmetric directionality of the relationship between model and target. The previous discussion of caricatures and analogies provides a clue to this structure we are looking for. Both are intended to draw attention to some particular aspect of the target system in a way that depends as much on the distortions and dissimilarities between the model and target as it does on the similarities. Catherine Elgin (2009) refers to this relationship as "exemplification." Elgin builds off of Nelson Goodman's (1968) notion of "representing *as*" to elaborate the complex structure of exemplificatory representations. "Representing as" is a peculiar sort of representation, as Goodman illustrates in the example of representating "Churchill as a baby." Taken as a representation of Winston Churchill *when he was a baby* this is a straightforward denotative representation, referring to a real person as some specific time. But when

[18] The same critique applies for isomorphism as well (Suarez 2003).

taken as a representation of an *adult* Churchill *as if he were a baby*, something stranger and more complicated is going on.

Elgin notes a distinction between *p-representations* and *representations of p*, where the latter is a special case of the former, one in which 'p' really exists. For example, some of the tapestries hanging in the "Cloisters" branch of the Metropolitan Museum of Art are unicorn-representations insofar as they depict a unicorn in captivity, being hunted, or dead and being brought to the castle, but these are not representations *of* a particular unicorn, or unicorns in general, since unicorns do not exist.

Representations that exemplify are representations *as*; which is to say that some object, *x*, is a representation *of* its target *y*, and represents *y as z*. In this scheme, "*x* is a *z*-representation that *as such* denotes *y*." Elgin goes on to explain this "as such":

> It is because *x* is a *z*-representation that *x* denotes *y* as it does. *x* does not merely denote *y* and happen to be a *z*-representation. Rather in being a *z*-representation, *x* exemplifies certain properties and imputes those properties or related ones to *y*. The properties exemplified in the *z*-representation serve as a bridge that connects *x* to *y*. This enables *x* to provide an orientation to its target that affords epistemic access to the properties in question. (Elgin 2009, p85)

What is notable about this structure is that when something, *e.g.* a model, exemplifies some aspect of its target by representing the target *as* being a certain way, the target does not need to really be the way it is represented. In fact, the desire to exemplify may require this to be the case if the real target system is such that confounding features cannot be isolated without a significant distortion. Even when a model is a real, concrete system (e.g. billiard balls), its function of representing its target *as if* the target were the way the model portrays it (e.g. a gas whose molecules behave like billiard balls) may represent something that does not exist. It is worth emphasizing, here, that "*z*" in this structure is neither the model nor the target, but some third thing (neither billiard balls, nor a gas, but gas-as-billiard-balls/billiard-balls-as-gas).

Elgin goes on to say, still following Goodman, that exemplification cannot operate by stipulation alone (though there is a significant and necessary sense in which a modeler must denote a particular target in order for their model to represents it *as if* that target were a certain way). Intentional designation may be a necessary component of exemplification, but it is not sufficient. In order to exemplify some property, a model must *instantiate* that property. If what is being exemplified is a phenomenal property of the target system, then this phenomenal property must be in the model. If it is some formal relationship between properties, then this formal relationship must obtain in the model as well. Hence, the way in which a model represents is not arbitrary; it depends on some very specific and non-trivial similarity between the model and target. Yet, as mentioned previously, this similarity is not specifiable in either respect or degree as an objective criterion for all models. It depends upon the intentions of the modeler both for the denotation of the target and the selection of the feature that is being exemplified. Additionally, the intentionality of the modeler is not the only one to be concerned with. Since the function of models as tools for exemplification is to "afford epistemic access to" a particular feature of the target system, it needs to be able to draw an interpreter's attention to this feature.[19] This is accomplished through the strategic use of distortions which introduce dissimilarities between the model and target.

[19] This is a fairly complex and significantly socially constructed aspect of the communicative function of models. Though I do not address it further here, this is a substantial topic discussed in section II of this dissertation. One important thing to keep in mind about this, however, is that when models become more than representations, instead becoming the object of scientific inquiry and practice, exemplification in turn becomes less a process of the modeler intentionally highlighting some feature of the target system, and more one of the model *revealing* these features to an interpreter.

This offers a great deal in terms of characterizing the representational structure of models in general. At the risk of excessive reiteration, models exemplify some feature of their targets in a way that is objectively non-arbitrary but nonetheless dependent upon the intentions of the modeler. The degree and respect of the similarity between the model and its target are not definable independent of the modeler's intentions. Furthermore, certain dissimilarities between the model and target are not only to be tolerated, but are essential to the art of exemplification. These dissimilarities assert the features of the model and target that are taken to not be relevant. All of the representational styles discussed heretofore (i.e. isomorphism, idealization, caricature, and analogy) can be recast through this lens, and considered means to this broader goal of models in general— to facilitate epistemic access to a target system by exemplifying certain features of it.

Exemplification reveals not only the representational structure of models, but the aim that motivates us to use representations that are unfaithful to their targets as well. Previously, we saw other candidates like simplicity and analytic tractability, but these too, like the various representational structures discussed, can be subsumed under the broader goal of rendering a feature of some real phenomenon epistemically accessible by exemplifying it.

A word on intentionality (mentioned earlier as a necessary aspect of representation) is warranted here. Following a tradition that extends back at least to Edmund Husserl (*c. f.*, Ideas I, sections 87-98) (and possibly all the way back to Kant) we may note that all perception is a *perceiving as* which involves an interaction between perceiver and perceived. In fact, the distinction between perceiver and perceived is one that can only be made conceptually – every act of perception is, in fact, an interaction that is co-constituted by both. We might emphasize the subject-pole of this interaction by stressing perception as always interpretive, always a *taking as*, and this is what I take to be the main thrust of stressing the intentional dimension of perception, representation in general, and specific sorts of representation in particular. There is no perception or representation that is not *of* something (even some fictional thing) and does not require that we pick out, or intend an object as taken in some way.

But it is as big a mistake to exclusively dwell on the activity of the subject in this interaction as it is to suppose that perception is passively suffered by the subject. Objects must afford being taken as a certain way, and various styles of representation place different levels of constraint on the ways in which the representans can be taken. The loosest of these is styles is symbolic representation, but even here some constraint, typically in the form of a convention amongst symbol users, dictates what a symbol represents. A symbol may be designated by fiat, but without at least a tacit ratification of this designation by the community, a symbol cannot be said to represent its representandum. Even in the case of private representation, say a mark in the margins of a text, a symbol must be used in a relatively consistent fashion in order for it to represent.

What is often called "iconic" representation has more rigorous constraints on the sorts of representations that may be taken in a certain way. It must bear some non-arbitrary similarity to its representandum. Accordingly, what an icon may represent depends, in a sense, less on convention (or at least on something other than convention) though it is not altogether free of conventional aspects. To talk of a non-arbitrary similarity between representandum and representans introduces what some might take to be an unfortunate way of speaking about things, one that assumes features inherently belonging to objects, and therefore an objective basis for evaluating whether or not similarity between two things obtains. But we may excuse this as a naïve, or commonsense description of a more complicated state of affairs where the features in question are a product of an interaction between a subject and some object, and where this interaction is conditioned by a complex background of habits, conventions, and perceptual capacities.

A key characteristic of models is that they exemplify some key feature(s) of their targets. I take exemplification to be a subset of iconic representation, one that has a particularly rigorous constraint relative to iconicity in general. Iconic representation, generally understood, need only be a recognizable likeness of its representandum, one that does not require designation by fiat. For example, the icon frequently used for a women's restroom is a convention, but one which is relatively easy to decode even for subjects who are not all that familiar with the convention. Yet the only criterion of relevance for the similarity of features involved for this icon is recognizability. In order for something to exemplify, to be a exemplary representative of some target, it must not only bear a non-arbitrary similarity to what is being exemplified, but must instantiate features whose relevance goes beyond mere recognizability. For example, Nelson Goodman refers to a collection of five patterned fabric swatches (Goodman 1978, 133-137), all of which are recognizable as representatives of the same fabric, but one of which best exemplifies the pattern in a way that is not only recognizable, but relevant to the practical needs of someone who is, say, looking to upholster their couch with this fabric.

Again, it should be pointed out that increased constraints on what may adequately serve as an instance of any particular style of representation do not ever get clear of convention (*c.f.* Goodman 1978, 116-120). Nor can they, no matter how rigorous, ever negate the indispensable role of some subject who takes the representation as such-and-such. If we talk of "relevant features" being similar as if that relevance is objective, and those features inhere in objects themselves, this is just a common way of speaking that is a short-hand for a more complex account of the way that subjects and objects interact and co-constitute one another. We may, however, gesture toward the prospect of naturalizing some of these questions of relevance by referring them to historically entrenched habits of interaction that facilitate human flourishing given the sorts of bodily capacities we possess.

As mentioned before, models exemplify, but I would argue that they exemplify in a particular way that allows us to further specify the sort of representation involved in modeling. Given the persuasiveness of claims that models are tools for inquiry that reveal relevant features of their targets through their manipulation and use, I find it helpful to think of models as props. A prop may, like a symbol, be designated by fiat, but it must perform adequately when taken up and used as if it were what it stands in for. It must instantiate the features of its representandum that are relevant to its use. Thus, all that can be said about what is needed for an adequate exemplification can also be said about what is needed for an adequate prop, though we may emphasize that the features instantiated are relevant with respect to being taken up in a performance.

Some latitude is warranted if we expect a model to reveal something new and unanticipated about its target. If, as Goodman says, representation is a way of "making worlds," inquiry demands that we create worlds that enable us to see things in unanticipated and productive ways, ways that are obscured in the worlds we currently inhabit. Thus we cannot specify, in advance, what sorts of degrees or respects of similarity there must be between a model and its target. The relevant features that are revealed through the exploration of a model may only be apparent retrospectively.

Furthermore, given arguments that will be articulated in chapter six about the nature of make-believe, we may make further specifications about what is needed in a *fictional* prop/model. Insofar as it is a model, it exemplifies some salient features of the target which are relevant to the way it is taken up and manipulated, but insofar as it is fictional, we may say that it also fails to exhibit some other significant and relevant similarity to its target. What is required, therefore, for the prop/model to play itself out is that conditions be in place that can prevent this failure from becoming conspicuous. This is essential to maintaining an absorbed engagement with the prop/model that is relatively passive/reactive in comparison to the explicit effort required for the designation of a representative relation by fiat. Note again, that the role of convention and the activity of the subject are not eliminated here, but they merely become transparent in the experience of fluid performance. And this, in turn, is crucial if we hope to take surprising and novel revelations of interesting features of the target seriously when they are the product of fiction.

The revelation of novel relevant features of a target system often requires that we unsettle entrenched assumptions about how the world is, and exploring fictional models facilitates this. But one would be correct in saying that this puts the intentional attitude required for make-believe in direct tension with the demand for relevant similarity required for exemplification. This is the unique problem that is posed by fictional scientific models. A very delicate balance must be struck between relevant similarity and the sort of significant dissimilarity that enables discovery. It is difficult to come by, and more often than not we may either succumb to the inertia of our sedimented world view or dally in fantasies which reveal nothing of consequence about the world. Furthermore, discovery is not the only end served by fictional models (as is discussed in chapter five) so this does not exhaust the full extent of the usefulness of scientific fictions

Models, theories, and the world

To briefly recap, models fulfill two distinct representative functions; they model theories, and they model the world. A perfect identity or even isomorphism between the model and “real” target system is too strong a relationship to cover all models.[20] Instead, a looser unspecified similarity is what a modeler aims for between a model and the world, which Giere calls “fit.” (Giere 1999, 122-123) So, theories are literally true descriptions of models (provided they are satisfactory models of the theory), and models “fit” real world systems to varying degrees. The better the fit, the better the model, and if that model is a satifactory model of the theory from which it is constructed, the better the theory as well.[21]

Thus there is what appears to be a two-fold role for models to play and according to which they can be evaluated. An effective model must instantiate the conditions in which a theory is true, and also fit actual, real-world target phenomena by instantiating and exemplifying the relevant features of real-world systems. A model can conceivably

[20] In fact, a perfect identity between a model and its target would render the model completely useless, offering no benefit whatsoever over studying the target itself.

[21] This provides a clue to understanding how someone like Nancy Cartwright can claim to be an anti-realist about theoretical laws, but a realist when those same laws are regarded as phenomenological explications of particular actual systems.

fail on either count, though to pitch a “failed” model into the rubbish bin just because it does not manage to fulfill both obligations is a bit rash. It is possible to see this two-fold role as a single role where the activity of “fitting” is not only applied to the relationship between the model and targets in the world, but to the activity of negotiating the demands of both roles – effectively “fitting” theories to phenomena via models. This sort of negotiation is no easy task, and must begin from imperfect but promising models which may model a theory better than they do the world or vice versa.

This is what Mary Morgan and Margaret Morrisson seem to have in mind when they describe models as “mediators,” that is to say their function is to mediate between theories and the world. However, this characterization as it stands leaves models as something “between” theories and the world, a depiction which Morgan and Morrisson offer convincing arguments against in favor of one where models stand apart in a semi-autonomous relation to theories and the world, “outside the theory-world axis.” On their account, models are technological tools, “instruments of investigation,” with a wide variety of functions. They are used, not only to represent theories, but in their construction as well, not only as representations of the world, but as tools which draw our attention to certain features of the world, and inform and facilitate our ability to collect data from the world as well as plan our interventions in actual systems. Their ability to perform these functions requires that they are not completely dependent on either theories or the world for their construction and use (Morgan and Morrisson 1999, 10-25).

Deborah Dowling (1999) makes a similar point about the autonomy of models, arguing that the activity involved in working with models and simulations is more like experimentation than theory. Thus the practice of modeling and simulating is characterized by “fiddling with machines,” “trying things out, “watching to see what happens,” and relies heavily on “tinkering, noticing, and intuition” in ways that are uncharacteristic of theorizing, and strikingly similar to experimentation. Furthermore, Dowling comments on the “strategic black-boxing” of models/simulations, such that scientists are invited to “treat the system as a unitary entity, and to *interact* with it, the way one would interact with a person or object.” This black-boxing is, as Dowling is careful to point out, to be distinguished from the sort of black-boxing that characterizes tools whose reliability hinges solely upon their transparent functioning in practical projects. Dowling stresses that it is temporary, and goes on to elaborate on this by explaining that it is considered professionally irresponsible for computational modelers to not understand the inner workings of their models.

Furthermore, Morgan and Morrisson emphasize that while the construction of models as representations of theories and/or the world does offer significant opportunities to learn about those targets, the technological and independent character of models is more pronounced when we learn from models by “using them,” that is to say when we begin to manipulate these models. In this sense, models are a “representative” rather than a “representation” of their target, a surrogate upon which interventions may be made. This gesture toward the manipulability of models serves to underscore their autonomy (since these manipulations are interventions performed on the models themselves apart from their relationships to theories and the world), but also redirects our attention towards a performative dimension of models and modeling apart from their representational function (Morgan and Morrisson 1999, 25-37).

It may be argued that while the manipulability of models is a particularly interesting feature of some models, it doesn't apply to all models. Perhaps this is true, but I would respond that exceptions are fewer than we might expect, and those models which aren't or even cannot be manipulated are models in a deficient sense of the term. Mathematical functions, for instance, essentially involve a manipulation in which their use requires the input of an argument to yield a value. Proceeding from the input to the output of a mathematical function doesn't happen automatically; frequently the function is portrayed graphically or instantiated materially in some sort of calculating machine that must be manipulated in order to yield output values from input values. Even working the solution to an equation out with pencil and paper requires a stepwise manipulation of the arrangement of symbols, an activity that seems more focused on the objects on the page than on straightforward formal inference. It is rare that one is able to make use of a mathematical model without some form of concrete manipulation of something, and when we are able to do without such manipulations, it would seem that we are in the territory of the derivation of formal consequences from theoretical statements rather than working with models at all.

Likewise, many other uses of abstract models seem to defy the suggestion that they are manipulated in a concrete way, in particular because they are never constructed in a way that goes beyond mental modeling. Even with these mental models, we do imaginitively manipulate them, usually calling such activities "thought experiments." I'm inclined to include such imaginative manipulations along with material ones as genuine manipulations of models, but as is the case with finding solutions for all but the most basic mathematical functions, most people cannot do even moderately sophisticated thought experiments entirely in their imagination. Take a very simple exercise like the following for example: imagine a cube that is turned so that it stands balanced on one of its vertices with the opposite vertex directly above it, and then imagine that a horizontal cross section is taken from this cube exactly 1/3 of the distance from the bottom vertex to the highest point. What shape is this cross section? The instructions for this task are easy enough to understand, and with a physical cube to turn in one's hands, or a diagram, the solution is not difficult to find. But to do this entirely in one's head, eyes closed in an armchair, requires quite a bit of concentration if one can do it at all.[22] It is no surprise, then, that most fairly sophisticated thought experiments, say any of Einstein's famous thought experiments, do involve material props (e.g. diagrams) in order to perform. Other mental exercises frequently referred to as thought experiments, say Galileo's conversations between Salviati and Simplicio work fine without props, but on closer inspection work by formal deductive inference, not by imaginative manipulation, and are better characterized as hypothetical theoretic deductions rather than manipulations of non-linguisitic mental representations.

Accordingly, an adequate account of how scientific models work would have to accommodate not only the way they represent, but the way that they are manipulated. For this, we must not forget that models are frequently (if not always) concretely (and even materially) instantiated objects that, despite their representational relationships to theory and the world, are also vaguely autonomous from them.

[22] The cross section is a triangle, and from my anecdotal experience, most people I know who can figure this out without the aid of material props do so by way of a combination of imagination and formal deduction (i.e. "I can imagine the vertex of a cube, and can 'see' that three faces/three edges meet at it, therefore a cross section 1/3 the way up from the vertex will have to cut through three faces/edges, and therefore it will have three sides/angles, and therefore it is a triangle).

Ontology, again

Both the representational structure of models (in relation to theories as well as the world) and this additional performative dimension provide a clue for the questions raised at the beginning of this chapter regarding the ontology of models. We struggled with the question of whether models are abstract or concrete. Since models represent both theories and real-world systems, their relative concreteness or abstractness is relative to what we compare them to. Since they instantiate conditions under which a theory is true, they seem to make theoretical abstractions more concrete. This relative concreteness in relation to theory is further elaborated in the semantic view of theories, which defines a theory by identifying it with a group of models that successfully represent it.

However, since a single model may fit a wide variety of actual systems, such that they represent classes of real-world systems for which theories are applicable, it would seem that models are more abstract than particular targets in the world. But even with respect to the way models represent the world, there is a notable concreteness.

Models of theories instantiate conditions in which a theory is true. Models of the world exemplify certain salient features of real systems and phenomena, and do so by instantiating them. In both cases, there is an essential instantiation, a making manifest of some system that will facilitate making a fit between theories and the world. This requires a certain concreteness of models, an effort to make even a non-existent system "real" in some non-trivial sense. Additionally, the manipulability of models indicates that they be the sort of thing can be manipulated, which in turn demands some level of concreteness. This concreteness need not be of the material variety; a model must be concrete enough to be manipulated and this requires particularity (though given how feeble our imaginations are when it comes to manipulating mental models, some sort of material prop is frequently employed). So, it would seem that models do require a specific sort of concreteness.

But we also considered whether models were real or unreal. Recall that reliable manipulability was Hacking's standard for reality, the guiding sentiment behind the tagline "if you can spray them, they're real." However, it is doubtful that reliable manipulability of a model was what Hacking had in mind when developing his instrumental realism,[23] and it would seem that we are playing pretty fast and loose with the term "real" if we attribute it to models without some sort of qualification. Between their concrete manipulability and their instantiation and exemplification of what one might take to be the "real" features of actual phenomena, there is something undeniably real about a model.

Yet, this is only insofar as we consider models in isolation from their targets, and to call models real would be to ignore the representational function which is just as essential to models as is concrete manipulability. When a model instantiates a world where the universal laws of a theory are true, even when those laws are not true of the real world, it is clear that the model world is not real. Even with respect to the way models represent the world, we should recall that when models exemplify certain features of their targets, they do so by representing them *as if* they were a certain way, and this certain way need not be realistic. When we ask what sorts of things models are, our answer is that they are *virtual* constructions, with many of the trappings of the real yet somehow still being potentially fictional.

[23] He is, in fact, quite clear that modeling belongs primarily to the realm of "representing" rather than "intervening" and that models are "most generally... something you hold in your head rather than your hands." (Hacking 1983, 210-219)

Chapter 3
Fictional Scientific Models: From Representation to Performance

Specifying what we mean when we say that some particular scientific model is a "fiction" requires some care. The mention of fictional scientific models generally draws two sorts of negative responses. Those with instincts leaning toward scientific realism have a tendency to protest that calling so-called "fictional" models *fictions* amounts to an exaggeration about the ways that science frequently employs representations that are abstractions, or approximately true. On the other hand, those who find themselves aligned with scientific anti-realism, whether in its logical empiricist or social constructivist forms, may find the prospect of fictional models to be obvious and even unavoidable.

In this essay, I aim to accomplish two things. The first is to offer a narrow conception of "fiction" in the context of scientific modeling that avoids both of the objections just mentioned. After entertaining some options from existing efforts to distinguish between narrow and wide conceptions of scientific fictions, I draw from concepts developed in the previous chapter to articulate a narrow conception based on the discontinuity between fictional models and their targets evident in our inability to gracefully de-fictionalize them.

The second is to proceed from this primarily representationalist conception to an account of fictional models and simulations that can accommodate their performative and technological character. This does not amount to disregarding the representational character of models, for it is essential to scientific models (even fictional ones) that they are representative of some target phenomenon. Nonetheless, I argue that we should think of scientific fictions as fictions first, and scientific second, implying that an effort to understand them demands that we place their functions in make-believe performances center-stage. Drawing on insights from Kendall Walton on the subject of mimetic performance in the arts and games of make-believe, I suggest that we think of scientific models (especially fictional ones) as props for a unique sort of scientific performance.

Vaihinger's Fictionalism

The vast majority of philosophers addressing the putative epistemological role of fictions, specifically scientific fictions, take Hans Vaihinger's *Philosophy of 'As if'* as a major historical touchstone. It is through Vaihinger that some initial efforts to narrow the concept of 'fiction' can be made. According to Vaihinger, fictions are propositions that are in conflict with reality in some way, and are employed or asserted in spite of full knowledge of this conflict. Vaihinger makes a distinction between semi-fictions and "real" fictions, stipulating that a true fiction is not only in conflict with reality, but is self-contradictory. For example, Vaihinger takes statements involving entities such as atoms (taken by him to be extensionless; we might today identify the offending concpet as point-masses, or point charges) and voids to be fictional, in that he finds them to be logically incoherent, yet frequently employed and expedient in our practical and theoretical activities. We might add to this concepts such as instantaneous velocity, action at a distance, or any of the mind-bending notions found in various interpretations of quantum mechanics that have remarkable utility (not only in making predictions and reliable technologies, but also in developing further research programs), yet on close consideration seem to make logically contradictory claims.[24] As examples of semi-fictions, Vaihinger offers the use of Ptolemaic astronomy by Arab scholars in the Middle ages, artificial taxonomical classifications, as well as certain abstractions and idealizations which simplify real phenomena into reduced versions that are more easily conceptually or mathematically analyzable.

While these claims may seem an attack on the legitimacy of science, Vaihinger does not mean it to be so. Rather, his goal is to use such examples to vindicate the role of fictions in practical matters. With strong influences from Darwin and Schopenhauer, Vaihinger stresses the natural function of mental activity as the preservation of the organism (Vaihinger 1-13). The justification of fictions thus lies in its indirect (and not necessarily conscious) accomplishment of the goal of proficient activity in one's environment. To quote Vaihinger, "Instead of remaining content with the material given, the logical function introduces these hybrid and ambiguous thought-structures [namely fictional ones], in order with their help to attain its purpose indirectly, if the material which it encounters resists a direct procedure" (Vaihinger 13). Consequently, his emphasis on practice and citation of the practical utility as a justification for why fictions should be tolerated, situates Vaihinger under the big tent of pragmatist thinkers, without going so far as to advocate a Jamesian reduction of "truth" to "what works" (Fine 1993, 34). This allows him to embrace fictions and the role they play while still allowing them to be fictions, his central aim being "to undo the opinion that if constructs are devoid of reality they are devoid of utility," or, to articulate this in the contrapositive formulation, that "the inference from utility to reality" is flawed (Fine 1993, 26). In fact, the hallmark

[24] It's worth noting here that Vaihinger seems to take it on faith that self contradictory statements or concepts are also in conflict with reality.

of a truly productive fiction is that it provides a scaffold for building non-fictional theories and models, a scaffolding that can ultimately be discarded. Those fictions that are expedient means to certain practical and/or non-fictional ends are considered "scientific" fictions, while those without such utility are not. Though this indirectness might obscure the utility of fictions while they are being used, it hardly diminishes it (Vaihinger 88).

It is this indirect relationship to the accomplishment of its justification that helps us to understand the distinction Vaihinger draws between a fiction and an hypothesis. Vahinger devotes a chapter to untangling these two, acknowledging that they are frequently mistaken for one another. In short, hypotheses are intended as accurate representations of reality and demand verification by empirical observation. Fictions are not, and in fact are held despite the supposition that they are not accurate representations of reality. As mentioned above, the vindication of a scientific fiction lies in its utility, even its ultimate abandonment, while the vindication of a scientific hypotheses lies in corroborating evidence. To be sure, part of the confusion between the two lies in the instability of their respective statuses. Hypotheses may become fictions if, once they are demonstrated as failed, they are nonetheless retained. Likewise, a fiction, once held to be an untrue though expedient construct, may later be found to be plausible and adopted as an hypothesis (Vaihinger 85-90).[25]

Nor ought we confuse fictions with misunderstandings or erroneous judgments. Though it is true that such errors are false, and it might even be true that errors can be

[25] This seems significantly more likely in the case of semi-fictions, but can be the case for full fictions as well, if what was taken earlier as self contradiction is later shown to be non-contradictory. This possibility is acknowledged by Vaihinger as also destabilizing the distinction between "real" fictions and semi-fictions, as our categories are not stable over time, and may change a particular conjunction of ideas from contradictory to coherent or vice versa.

instrumental in helping us to progress (if for no other reason than that they allow us to rule out unproductive lines of thinking), the person who misunderstands or judges erroneously does so unwittingly. The person who puts forth a fiction, on the other hand, does so knowingly.[26] The fiction is untrue in some significant way, we know that it is untrue, and we use it anyway. If the fiction is scientific, then using it is expedient for the goals of scientific practice.

Narrow and Wide Conceptions of Fictions

Vaihinger's definition of genuine or full fictions as being self contradictory is a provocative move on his part, but it seems possible that he sets the bar a bit too high. Amongst his contemporaries, the prospect of concepts that are clearly both useful and self –contradictory elicited suspicion and even outright dismissal. Morris Cohen remarks on this, saying that "With amazing industry [Vaihinger] has gathered a most imposing list of what he calls genuine fictions. Nevertheless, there ought to be no hesitation in flatly denying that any of them do involve self-contradiction. If they did, no fruitful consequences could be drawn from them and they would not have the explanatory power which makes them so useful in science." (Cohen 1923, 485) Many cases of what Vaihinger terms "genuine" fictions are susceptible to *ad hoc* reinterpretations of what the putatively fictional concept actually represents. Point masses or instantaneous velocities, the physicist might say, are only limits that are approached, never reached, and to take them as self-contradictory concepts is to misunderstand them. Cohen offers a critique of the supposition that the "imaginary" value *i*, whose square is –1, is a self-contradictory

[26] The extent to which this is "known" is ambiguous and requires further elaboration. This is addressed in Ch 6 of this dissertation.

concept (given that there is no real quantity whose square is negative) by pointing out that this mathematical expression is “not a thing nor the property of a thing, but a relation or transformation of things” (Cohen 1923, 485-486).

Still, Vaihinger’s instinct to differentiate between a wide and narrow conception of fiction is laudable. It would seem that the term “fiction” is used in a variety of ways and it is helpful to not only categorize these uses, but to do so in a way that focuses our attention on essential features of scientific fictions. I suggest that we are looking for three qualities when we discriminate between wide and narrow conceptions of fiction (or any term for that matter): 1) we want to be sure that both the wide and narrow senses of a term are clearly applicable in real cases, 2) we want to be able to apply the distinction between wide and narrow senses of a term with as little confusion between the two as possible and 3) we want the narrow sense to identify cases that are more philosophically problematic than the wide sense. From the above objections, it seems that it isn’t clear that any non-controversial examples of Vaihinger’s narrow sense of scientific fiction (*qua* self-contradictory idea that is useful for progressing scientific inquiry) exist, so this way of differentiating a wide and narrow view fails my first criterion.[27] Accordingly, I prefer to pursue alternative ways of addressing “wide” vs “narrow” views of fictions.

[27] Mauricio Suarez (2004) argues that there are mutually contradictory assumptions articulated in insolubility proofs of quantum measurement. Such proofs can also be found in Wigner (1964), Earman and Shimony (1968), Fine (1970), Brown (1985), and Stein (1996). It is notable that the four assumptions Suarez claims as contradictory (which he identifies as the “transfer of probability condition,” the “occurrence of outcomes condition,” the assumption that the measurement system is closed, and the assumption of “real unitary evolution”) are only contradictory under certain interpretations of quantum observables. More noteworthy is the fact that it is difficult to pin down where the offending “fiction” lies, since the contradiction is the result of all four assumptions being held at once. I would not count this as an uncontroversial case of a Vaihinger-ian “genuine” fiction, though it is a compelling enough example as to suggest that we ought not dismiss the possibility of such fictions altogether as Cohen suggests.

Mauricio Suarez proposes doing so along the lines of what he terms the "fictional" and "fictive," where fictional models represent objects that don't exist, and fictive models represent existing objects falsely (2009, 13, 173). Hence, a model of a star that portrays the star as a sharply defined closed system unaffected by external gravitational or electromagnetic forces, composed of 70% hydrogen and 30% helium, in a permanent state of thermal equilibrium and spherically symmetrical (Suarez 2009, p165) is *fictive*, since it is employed to represent a real star, but does so in a way that contradicts empirical facts about that star.

For an example of a *fictional* model, we may consider the famous "ether vortex" model proposed by James Clerk Maxwell in his treatment of electromagnetic fields, articulated in these excerpts from Maxwell's 1861 paper "On Physical Lines of Force":

> We are dissatisfied with the explanation founded on the hypothesis of attractive and repellent forces directed towards the magnetic poles, even though we may have satisfied ourselves that the phenomenon is in strict accordance with that hypothesis, and we cannot help thinking that in every place where we find these lines of force, some physical state or action must exist in sufficient energy to produce the actual phenomena. My object in this paper is to clear the way for speculation in this direction, by investigating the mechanical results of certain states of tension and motion in a medium, and comparing these with the observed phenomena of magnetism and electricity. (Maxwell 1861, 161-162)

> Let us now suppose that the phenomena of magnetism depend on the existence of a tension in the direction of the lines of force, combined with a hydrostatic pressure; or in other words, a pressure greater in the equatorial than in the axial direction: the next question is, what mechanical explanation can we give of the inequality of pressures in a fluid or mobile medium? The explanation which most readily occurs to the mind is that the excess of pressure in the equatorial direction arises from the centrifugal force of vortices or eddies in the medium having their axes in directions parallel to the lines of force. (Maxwell 1861, 165)

> I have found great difficulty in conceiving of the existence of vortices in a medium, side by side, revolving in the same direction about a parallel axes. The contiguous portions of the consecutive vortices must be moving in opposite

directions: and it is difficult to understand how the motion of one part of the medium can coexist with, and even produce, an opposite motion of a part in contact with it. The only conception which has at all aided me in conceiving of this kind of motion is that of the vortices being separated by a layer of particles, revolving each on its own axis in the opposite direction to that of the vortices, so that the contiguous surfaces of the particles and of the vortices have the same motion. (Maxwell 1861, 283)

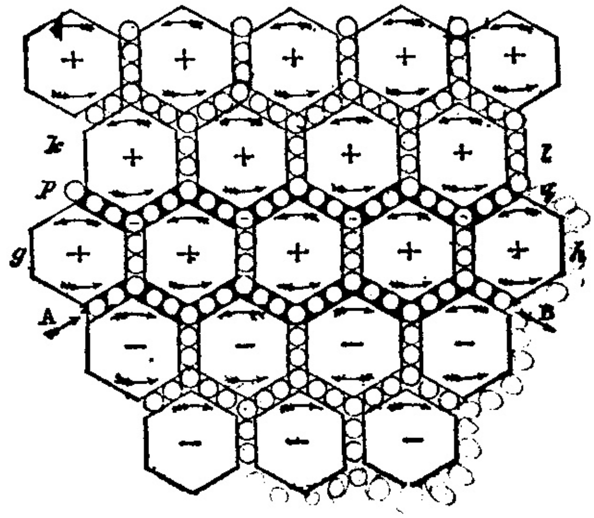

Diagram of Maxwell's model, from Maxwell 1861, "Fig. 2"

Maxwell's insistence on a causal mechanism producing magnetic force fields leads him to entertain an ethereal medium in which this mechanism operates in order to "clear the way" for further speculation. This leads him to entertain a mechanism that operates through an analogy to a hydraulic system, creating pressure via vortices in a fluid medium that rotate perpendicular to the lines of magnetic force. And this, in turn leads him to note that this would require ball-bearing-like particles between each vortex in order to allow for adjacent vortices to rotate in the same direction. The ether vortex model was instrumental in the development of the equations Maxwell developed for the governing laws of electromagnetic phenomena, yet insofar as there is no ether that serves as a medium for the vortices he describes, the model operates not merely by way of an analogy to hydraulic systems, but by way of an analogy to a non-existing mechanical system. Though the extent to which Maxwell himself believed in the existence of vortices in a medium as the cause of electromagnetic forces is somewhat contentious,[28] the sheer complexity of his mechanical model, complete with ball-bearing-like non-etherial particles existing between the vortices strains credulity. It seems safe enough to say that Maxwell allowed his speculation of the mechanism that gives rise to magnetic forces to stray into territory that was more fictional than hypothetical, making his ether vortex model a relatively clear case of a Vaihinger-ian "scientific fiction," and a fairly plausible example of Suarez's "fictional objects."

That some physical cause exists for electromagnetic forces at a distance seems to be a given for Maxwell, something "we cannot help thinking." The ether vortex model he produces is intended to explore the sorts of "tensions and motions" in a medium that could give rise to those forces. Now whether that model is a system that we imagine in our minds (prompted by Maxwell's text), a diagram drawn on paper, or mechanical structure of gears and cranks, I would take this to be the same model. One could, in principle, I suppose, make an argument that the imaginary model, the diagram, and the machine are all different models, or even that some of these are, in fact, models of others, but I fail to see the value in doing so. All have the same sorts of features that behave in the same sorts of ways and bear the same representational relationships to their targets.

[28] The evolution of Maxwell's own language in his 1861 "On Physical Lines of Force," 1865 "Dynamical Theory of the Electromagnetic Field," and his 1873 "Treatise on Electricity and Magnetism" displays a diminishing ontological commitment to the existence of the ether or the literal truth of his mechanical model as an hypothesis of the cause of electromagnetic phenomena, but to the consternation of even his contemporary readers, he is never clear about how one should view this model. Michelson and Morley's null results regarding the existence of the ether would not occur until after Maxwell's death in 1879 and even then would take some time to be interpreted and gain acceptance as definitive evidence against the existence of the ether. Nonetheless there does seem to be a consensus amongst contemporary scholars (Morrisson 2001, 2009, Suarez 2009, Crease 2008, Warwick 2003), that the ether was a dubious entity to tie a causal theory of electromagnetism at the time, and that even if Maxwell himself started out hypothesizing/believing in the literal truth of his model, he continued to employ and elaborate on it well after abandoning such beliefs.

The construction of models in different media can introduce significant differences in modeling decisions, such that wholly distinct models are created in one medium as opposed to another, but a difference in medium does not always yield a significant difference in the model (e. g. a diagram made with paper and pencil is, for all intents and purposes, no different than the same diagram made with chalk and slate).

But such differences in medium, while not significant enough to require that one version of a model be a wholly different model than another version of that same model, may yield differences in the way that various versions get taken up and draw our attention to some insight. For example, if we take a particularly interesting step in the construction of Maxwell's model, the introduction of interstitial particles between vortices, the problem that precipitates their introduction reveals itself in different ways depending on whether we imagine the system, diagram it, or build it out of gears and cranks. The problem is that the vortices cannot all spin in the same direction without running in opposite directions where neighboring vortices meet. Noticing this problem imaginatively is difficult, and I would even suggest that it may even be possible to picture a collection of tightly packed vortices all rotating clockwise in one's mind. Drawing the model out makes the problem more conspicuous, as one would more easily notice that arrows designating clockwise rotation for neighboring vortices run in opposite directions where the vortices come into contact with one another. Still, there is nothing in the affordances of the paper and pencil that would prevent me from drawing such arrows. In a machine built of gears and cranks, however, the problem is unavoidably conspicuous, as one would not be able to turn the gears at all.

Yet, there is potential ambiguity here regarding the extent to which Maxwell is denoting a fictional (i.e. non-existing) object in these models, since he is also representing a field of electromagnetic force (which he was earnestly ontologically committed to) by way of a less serious analogy to a non-existent surrogate that represents this field incorrectly. This ambiguity is shown more clearly in an example of a scientific fiction put forward by Eric Winsberg (2006, 2009). Winsberg describes a modeling practice used in nanomechanics for studying the propagation of cracks in a solid. There is a problem with modeling the phenomenon in question due to the fact that it bridges scales that are represented by three distinct theoretical models. Continuum mechanics (CM) is typically used for describing solids at or near equilibrium, but breaks down when the system strays too far from equilibrium. For systems further from equilibrium, classical molecular dynamics (MD) offers a much better description, but the computational demands of modeling a solid this way limit applications to regions of about 10^7-10^8 molecules (~50 nanometers). Furthermore, the formation of cracks in a solid involves the breaking of chemical bonds, which, in turn, involves changes in the electronic structure of individual atoms. This renders classical molecular dynamics inaccurate at the very tip of the propagating crack where these bonds are breaking. Quantum mechanical (QM) approaches handle this region better, but are even more computationally expensive (limited to ~ 250 atoms at a time). “The upshot,” as Winsberg tells us, “is that it takes three separate theoretical frameworks to model the mechanics of crack propagation in solid structures on the order of one micron in size” requiring that the region being modeled be divided into three regions where three different models are applied simultaneously and where information is continuously passed between regions.

The difficulty with this lies in what Winsberg calls the "handshaking regions" on the boundaries of the domains where each of the three models is applied, and where the transfer of information from one model to another occurs. Each model has distinct fundamental elements.

> In the CM region, the elements are vertices on a grid over which the continuous equations have been discretized. The energy function comes from the elastic forces. In the MD region, the elements are molecules, and the energy function comes from a classical force function. And in the QM region, the elements are atoms, and the energy function is a quantum Hamiltonian. (Winsberg 2009,185)

In order to pass information back and forth across the borders of the regions, algorithms are required to "interpret" information from one model to another. And there is a particularly nasty problem for the "handshaking" algorithm that the MD region uses to interpret the energetic interactions with neighboring atoms in the QM region. Since the QM region calculates energy globally rather than locally, there is no straightforward way to calculate the energetics of the interaction between the outermost QM atom, and the innermost MD atom. As a solution to this problem, researchers assign a hybrid character to the atoms at the boundary of the QM region that allows them to localize their quantum mechanical energy. For scenarios where the material being studied is silicon, this involves positing fictional atoms called "silogens" which have some properties of silicon and some properties of hydrogen.

As Winsberg argues, these "silogen" atoms are clearly fictions, since no such atoms exist in nature or anywhere else. Yet, it is unclear whether they should count as *fictive* or *fictional* according to Suarez's description. There is *something* there in the "handshaking" region, it simply doesn't have the properties that the model represents. Fictional objects are fairly unambiguous when we are considering characters in a novel, or paintings of unicorns, and it is easy enough to distinguish them from fictional depictions of real objects, but scientific models seem to problematize this distinction since they aim, however indirectly, at some apparently real phenomenon as a target.[29]

This ambiguity is characteristic of the "representation-as" structure of models described in the previous chapter (as articulated by Elgin and Goodman). Insofar as representations-as are a two stage representation — a "p-representation" that does not necessarily denote any existing object, and a "representation of p" that does denote an existing object — it is not surprising that we would have difficulty in applying Suarez's distinction between fictions that represent non-existing objects and those that represent existing objects as other than they actually are. Both Maxwell's ether vortices and Winsberg's "silogens" fit this structure, and consequently are ambiguously "fictional" and "fictive" in Suarez's sense. We may try to rescue Suarez's version of a narrow and wide sense of the concept of fiction by stipulating that it hinges on the first stage of a representation-as. If a model *x* represents *y* as *z*, perhaps we could say that the model is merely fictive if *z* exists, and fictional when it does not. But this still does not get us clear of a deeper problem in Suarez's distinction. How are we to distinguish between representations of existing objects as other than they actually are from representations of non-existing "objects?" Is a "silogen" atom simply a silicon atom with unrealistic properties, or is it some distinct and wholly fictional thing?

[29] Winsberg's "silogen" example suggests a subtly different to the ambiguity in addition to the one regarding the fictional/fictive status of elements of models that represent existing targets *as* non-existing objects; namely, a possible confusion regarding the fictional/fictive status of parts of a model when the greater model-at-large is arguably non-fictional. This potential for confusion is not necessarily problematic, and it may well be that whether or not a model is fictional or fictive depends upon our mereological focus. However, this focus is somewhat arbitrary, and if a clearer and less arbitrary means for distinguishing between a narrow and wide conception of fictions is available, it would be preferable.

Suarez claims that Winsberg's efforts to establish "silogens" as a fictional part of a greater non-fictional model/simulation indicates a narrow conception of fiction qua fictional object that excludes fictive description (Suarez 2009), but it seems that Winsberg is actually after something slightly different. He does favor a narrow view of fictions, but it hinges less on excluding fictive descriptions than excluding approximation and abstraction. Winsberg is careful to emphasize that the simulation of crack propagation in a silicon brick, at large, is not a fiction. Nor, for that matter, are the three component models (QM, MD, and CM) that govern its three regions.[30] "We are deliberately getting things wrong locally *so that we get things right globally*." (Winsberg 2009, 186, emphasis mine)

While Winsberg's gesture toward discriminating between fictions and non-fictional attempts to represent the target with a reasonable degree of accuracy appropriate to a modeler's aims is a bit vague, I believe we can strengthen it by way of reference to the concept of de-fictionalization described in chapter two. Recall that within the broad class of representations that represent inaccurately, a distinction was made between those that can be gracefully de-fictionalized, and those that can. To reiterate that standard, among those representations that represent inaccurately, some, like idealizations and abstractions, do so with "approximate" accuracy. Abstractions and idealizations may be approximately true, even though they are not strictly true. They can be de-idealized/de-fictionalized through a stepwise series of continuous corrections and additions that make

[30] The extent to which the incompatibility of the three models employed in this simulation suggests something fictional is an interesting question; one that calls to mind Michael Weisberg's discussion of "multiple model idealizations." This is certainly relevant to the subject matter of this dissertation, but is just beyond its scope. I am prepared, for the time being, to agree with Winsberg (who I believe is drawing from Giere here) that insofar as each model yields an approximately true depiction of its target *when restricted to appropriate contexts*, it is non-fictional.

a model as accurate as our needs demand. Other models, however, like those that represent their targets via analogy, surrogates, or caricature, cannot be de-fictionalized in this way. Attempts to gradually square these representations with their targets have a tendency to do so in larger discrete steps and get worse before they get better.

The latter seems to be the case both for Mawell's ether vortices and Winsberg's "silogens." For instance, if we remove the ball-bearing-like molecules between the vortices, in Maxwell's model, it becomes less fictional but unworkable, as it no longer allows for adjacent vortices to spin in the same direction. De-fictionalizing this model while simultaneously preserving its utility requires such radical changes that we would effectively be constructing a whole new model rather than adjusting a distorted one. For the "silogens," a more realistic portrayal of the silicon atoms becomes computationally problematic absent a workable "handshaking" algorithm. Again, this would be more accurate "locally," but at the expense of the "global" utility of the model, and de-fictionalizing the model while simultaneously preserving its utility involves significant, discrete and qualitative changes.

I suggest that this distinction, one that specifies a narrow and wide conception of fictions on the basis of whether or not the inaccuracies of a fictional representation can be "de-fictionalized" without sacrificing the overall adequacy of the representation, is preferable to the others that have been proposed above. Following a conservative principle of preferring narrower conceptions to wider ones, I will henceforth refer to distorted representations that cannot be gracefully de-fictionalized as "fictions," and those that can be gracefully de-fictionalized as "approximations." Unlike Vaihinger's "true" or full fictions which are self-contradictory, there is little doubt as to the existence and utility of non-de-fictionalizable models in scientific practice (I have just discussed two of them). And this schema does a better job of avoiding vagueness and confusion than Suarez's distinction between fictional objects and fictive representations. Perhaps more importantly, it gets closer to the heart of what is interesting about fictional models. It is of questionable significance whether they represent non-existing objects or existing objects as other than they actually are (presuming that we could reliably discriminate between the two). It is far more important whether or not they are part of a continuum of gradual distortions with their targets.

A Brief Excursus on the Question of Models in Quantum Mechanics

On of the more vexing examples of the putative use of fictional models in physics concerns the question of models in quantum mechanics. Here, the phenomena to be modeled are so contrary to our intuitions (intuitions which serve us well in "classically" modelable phenomena) that it is difficult to see how we might model them in a way that is not fictional. This is a very broad topic, with a great deal of commentary, but it would be a mistake to not address it at all. For the sake of brevity, I will focus on Erwin Schroedinger's 1935 paper, in which he deploys his now famous cat thought experiment, and in which the question of whether the psi-function is a model.

Schroedinger seems pretty set on maintaining that the psi-function is not a model (*c.f.*, specifically, section 10 of Schroedinger 1935). Part of this is due to the fact that he understands "model" as classical, deterministic models (I could not find one instance in this paper where he employs the concept of a "quantum model," instead referring to "the model," by which he means a classical model). Admittedly, this is an idiosyncratic understanding of what a model, in general, is. Crucial to it is distinguishing between a “determinate model” of a phenomenon, its “determining parts,” and a “determinate state” of that model. For example, if we consider the Rutherford model of the Hydrogen atom, we may describe the model in terms of the values of certain constants (m, M, and e) and the relationships between variables (*e.g.* the positions of the proton and electron along with their velocities described in a rectangular coordinate system), which variables are its determining parts. The model does not specify what the values of these variable are, but insofar as they are determining parts of the model, when they are set, the model is said to be in a determinate state. Thus, the model provides a description of all the possible states of the system. But, Schroedinger is careful to point out, this involves more than simply a description of possible positions and velocities for the proton and electron in a Hydrogen atom. “It embodies also knowledge for *every* state how it will change with time in absence of outside interference.” (Schroedinger 1935, 3)

As we know, quantum mechanics represents a fairly stark break from this approach to modeling its phenomena. Specifically, the notion that all the determining parts of a model that adequately describes the phenomenon can all be known, in principle, to an arbitrary degree of precision. As expressed in Heisenberg’s uncertainty principle, the more precise our determination of the position of a particle, the less certain its momentum must be. Therefore some aspects of the system must necessarily be, at best, known probabilistically.

Yet, Schroedinger holds that the classical model, though our understanding of it must drastically change, is not to be abandoned altogether. For one thing, the classical model, does, in fact, tell us which features of an object can be measured (even measured to arbitrary precision if taken individually), it simply fails to capture the nature of the mutual interrelationship of the determining parts. As such, one might suppose that the deviation of quantum mechanics from the classical model is simply a matter of predicting probabilities rather than precise values of the determining parts of the model.

In truth, the break from the classical model is much more radical, but even this points to a peculiar value in retaining the classical model, if only for the purpose of illustrating this break. Schroedinger demonstrates this through an examination of angular momentum. Showing that the classical model dictates the determining variables to be the linear momentum of an object and the distance from a geometric reference point, while in quantum theory angular momentum is "quantized," allowing only certain values and forbidding others. As Schroedinger points out, this conflicts with the classical model, which allows the angular momentum of an object to vary continuously as the distance from the geometric reference point varies. Similar deep conflicts exist between classical and quantum treatments of harmonic oscillators and radioactivity, and in all cases the quantum treatment describes the relevant phenomena more accurately, underscoring that depth of the difference between the quantum world and classical models extends beyond a simple matter of thinking probabilistically rather than deterministically.

Thus the classical model, in its sharp divergence from quantum theory and quantum phenomena, can retain at least this, arguably non-trivial, consolation prize – it helps to illustrate important aspects of the quantum theory, specifically the ways that it bucks our classical intuitions. We can regard the classical model as a fictional model of those quantum phenomena to which it fits poorly, and this fictional model exhibits a certain pedagogical utility. But surely, we would want to move beyond this mere pedagogical utility, beyond a model whose primary virtue is illustrating how quantum phenomena are *not*. We desire a quantum model, and it seems as if the psi-function is a contender for this.

Why, then, does Schroedinger insist that the psi-function is not a model, one that portrays an entity whose variables are not precisely determined, but instead "blurry?" Again, the reason for this appears to hinge on Schroedinger's understanding of a model as a classical model. If what makes the classical model a model is that it represents all the possible states of a system, and what makes it an inadequate model of quantum phenomena is that it fails to capture the non-deterministic relationships between determinng features while also failing to accurately capture the uniquely quantized nature of those phenomena, the psi-function suffers the opposite problems. It does manage to represent quantum phenomena in ways that make accurate predictions and also captures the interesting features of those phenomena, but it does not represent all the possible states of the corresponding system.

This is a peculiar statement, particularly if we understand the psi-function to do precisely what I have just said it does not. What else is it but a representation of the determining features of a system (*e.g.* positions and momenta of objects) in terms of a probability distribution of all possible values? It becomes less peculiar, however, if we attend to two things: 1) the difficulties associated with making measurements of quantum systems, and 2) the previously mentioned distinction between "model" and "state" that Schroedinger is so careful to make.

Schroedinger describes the psi-function as an "expectation-" or "prediction-catalog" that provides "the relation- and determinacy-bridge between measurements and measurements, as in the classical theory the model and its state were." This seems to suggest that the psi-function is the quantum theoretic alternative to the classical model, and this is true in one sense. But the fact that measurements influence a quantum system severely complicates any prospective analogy that would allow the psi-function to serve in the same way as a model (i.e. a *classical* model) does. When a classical model is fleshed out to include definite variables for a determinate state, it provides us with knowledge of all subsequent states of the system. In order to define a state through the psi-function, one must make a measurement whose certainty is inversely proportional to the certainty of its conjugate variables. But this is not to be understood simply as missing knowledge about the state, *e.g.* wherein a particle whose position is known *does* actually have some determinate momentum but we simply don't know what it is. When we have the psi-function for a system, we know all that can be known about it. If we were to perform a subsequent measurement on the system, in order to become more certain of the uncertain variable, we introduce an abrupt change in the state of the system, such that measurement has now rendered *other* variables of the system (*e.g.* those whose values became known in the previous measurement) uncertain. Statements about the system which were previously correct have, in the course of taking new measurements, become incorrect, which can only be possible if the object described by those statements has changed. What this boils down to is that any given psi-function only applies to one state of a system, and since subsequent measurement changes the state of that system, a new psi-function is required each time a measurement is made. And if a (classical) model provides knowledge of all possible states when maximally determined, then we can see that the psi-function is incapable of doing this work.

To be sure, the preceding argument is complicated and begs some fairly crucial questions. Buried within it is a radically empiricist epistemology which holds that there is no underlying reality to an object beyond that which can be observed (or extrapolated from those observations) (see section 6 of Schroedinger 1935). Consequently, the ease with which Schroedinger asserts that the psi-function is not a model is more a consequence of his inclination to "*not* relate our thinking to any longer to any other kind of reality or model" than it is a justification for such a view. Additionally, it has already been mentioned that a "model" is defined in a particularly narrow, and even peculiar way for Schroedinger. Beginning from these premises, not only is the psi-function not a model for a quantum system, but it is difficult to see how there could be *any* sort of quantum model. But we may take him to be making a slightly different point, even if we do not follow him in insisting that a quantum model is, effectively, an incoherent concept, namely that our idea of what a model is must radically change if we are to entertain the possibility of modeling quantum phenomena. Such a model would require us to articulate concepts that are unintuitive, and even absurd, within the conceptual framework that we have acquired through millennia of interactions with classical, meso-scale phenomena. This is the point of his famous thought experiment involving a cat that is both living and dead until observed. Any model of quantum systems will be as unintuitive as the notion of a living and dead cat "mixed or smeared out in equal parts" (it is worth pointing out that this thought experiment, by its very nature, yields a result that cannot be said to be empirically false, but is instead so counterintuitive to our understanding of meso-scale objects as to be ungraspable). Until such time as that idea becomes more palatable, we may be inclined to dispense with the idea of a quantum model altogether.

But I find it difficult to see how it is that the psi-function could be used in the way that Schroedinger prescribes and not be a model. In order for something to be a model it must be interpreted as such, bearing a non-arbitrary representational relationship to some target; and if someone refuses to interpret it in this way, then it is not a model. But even under the most bare-bones interpretations of the radical empiricist, it would seem that a mathematical function *does* model the phenomenon it describes. The machine-like character of a mathematical function by which arguments are input in order to solve for a unique output serves as a model so long as those inputs and outputs are understood to correspond to something else. This is satisfied by the psi-function insofar as its input correspond to observed measurements and the output corresponds to the probabilities that future observed measurements will obtain. The fact that a new state is produced each time a measurement is made complicates matters but does not obliterate the representational relationship of the function to the system in a particular state.

Additionally, rejecting the supposition that the function *cum* model represents any underlying reality outside of these observed measurements, or even that such an underlying reality exists at all fails to escape from this representational relationship. In doing so we may limit the ways that the function is to be interpreted to correspond to its target, but we do not eliminate such an interpretation altogether. Furthermore, the "proper" use of the psi-function does not appear to demand this degree of metaphysical agnosticism. One may, as many have tried to, interpret the psi-function as representing particles with determinate positions and momenta in many-worlds, as particles that simply behave probabilistically, as wave-packets that collapse when interacting with a measuring device, or as objects governed by non-local "pilot-waves." The function itself is multistable for the interpretations it supports, but I would argue that in order for it to do what it does, it must be interpreted in *some* way to correspond to a target system, which is to say that it must be a model.

The extent to which we may classify such models as fictional or non-fictional is difficult to assess, however, and this is what makes the prospect of modeling the quantum world so confoundingly peculiar. Certain models of quantum systems, for example classical ones, are clearly fictional in the way they conflict with empirical data. Others meet the demands of empirical adequacy, but are so offensive to our intuitions as to border on incoherence (as illustrated with Schroedinger's cat). These may fall roughly under (or near) Vaihinger's notion of fiction as internally contradictory in that they demand that we think contrary to our ordinary conceptual tools in order to interpret them. This is a domain of fictional models that I have not spent much time addressing, mostly because they are both peculiar and rare. This alone does not justify ignoring them, and they do present a vexing problem that should be explored. However, I forego that exploration in order to address cases where models are straightforwardly comprehensible but known to be false, and are, nonetheless, employed as a means to some epistemic end. Fitting quantum models into that discussion is difficult, not only because their interpretations are difficult to comprehend, but also because we possess no non-fictional understanding of their targets against which we might say that they are either true or false.

But, these two types of fictional models (those which are taken to be false because they conflict with empirical data, and those which are so conceptually bizarre that they conflict with our intuitions) are more similar under a different approach to fiction, one which I would argue captures the essence of fiction better than what we might call a truth-conditional approach. In both sorts of cases, what is crucial is not that we take the model to be false, but that we simply do not care about its truth or falsity when we put it to work. Questions of whether we genuinely believe that the model represents its target the way it actually is are suspended when we make-believe in a fictional target-as-represented-by-the-model.

Getting Away From a Truth-Conditional Approach

The preceding efforts at drawing a distinction between a narrow and wide conception of fictions are instructive, both for allowing some increased precision in articulating precisely what sort of models we would consider to be fictional, and also for illustrating that these sorts of fictional models do, in fact, exist in scientific practice. Nonetheless, there is something about this approach to defining scientific fictions that misses an essential point. There is an ambivalence in Vaihinger's account of fictions that has been carried on by contemporary philosophers concerning what Suarez (2009, 12) refers to as "truth conditional" and "functional" trajectories of analysis. It would seem that fictions must be false in some significant way (which ways we have just specified) in order to distinguish them from non-fictional scientific facts; but despite the fact that most philosophers who address the topic of scientific fictions, from Vaihinger onward, maintain that fictions are false and are known to be false by those who employ them, nearly all of them also acknowledge that fictions are employed without any regard to whether they are true or false, so long as they possess some pragmatic utility. In one breath we say that fictions are not true and specify the ways in which they deviate from truth, and in the next we say that their truth or falsity is irrelevant for how they function.

It is worth pointing out that there is no necessary tension between these two statements, and the need to address both the truth conditional and functional aspects of fictions is a reasonable one. To say that fictions are employed for their pragmatic function without regard to whether they are true or false does imply that they *may* be true, but any expectation of truth emerging directly from an exclusive concern with pragmatic utility is unwarranted. This is precisely the heart of Vaihinger's objection to a Jamesian reduction of truth to "what works" and a key argument put forth by anti-realists from Duhem to Van Fraassen. Nonetheless, when a model that is known to be false does yield some pragmatic utility, this is surprising, and demands some attention. Furthermore, when we have reason to believe that a model which is known to be false is instrumental in developing ones which are thought to be true, our interest ought to be even more piqued. Thus, while an essential aspect of fictions may be that we do not care whether they are true or false, they are particularly interesting because they are false.

Additionally, articulating a clear account of the falsity of scientific fictions is essential for demonstrating that scientific fictions do, in fact, exist. Attending to their falseness, particularly by adhering to a narrow conception of what counts as fictional, provides answers to skeptics who might suggest that those who speak of fictions in science are simply making too much of a fuss over the fact that many scientific models are only approximately true. Even for empiricist anti-realists who seem to think science is always practiced with an eye to empirical adequacy and more or less of a disregard for metaphysical truth, the preceding analysis seems to suggest that there are significant distinctions between models which disregard truth more and those that do so less, to the extent that even empirical adequacy may be sacrificed.

I would argue that despite the consistent recognition that fictions are false *and* that we don't care whether they are true or false, the latter has been somewhat obscured by the former. Accordingly, discussions of fictional models have been dominated by themes of representation, correspondence to targets, and the mental, conceptual, and linguistic character of models at the expense of exploring themes of performance, ontological autonomy, and the technological character of models. Taking the functional trajectory of analysis more seriously involves giving precedence to the notion that we don't care about the truth or falsity of fictions, over the notion that fictions are false.

This seems to fit a more general approach to fictions, including non-scientific ones like those found in literature. *Alice's Adventures in Wonderland* may not be true, but to point this out, let alone to catalogue the ways in which Carroll's story succeeds and fails at representing reality, is anathema to the essence and function of the story as literary fiction. In this way, a focus on the representational character, specifically the ways in which a fiction represents the world inaccurately (or fails to represent any existing objects whatsoever) not only covers over this function, but misses the point of the fiction.

To inquire into the function of scientific models (fictional or otherwise) is to emphasize their technological character. This emphasis is not unprecedented, and some indications of this were given in the previous chapter in the discussion of Morgan and

Morrisson's conception of models as semi-autonomous technologies. Morgan and Morrisson stress that models are tools for investigation that render target phenomena epistemically accessible, and this distinguishes them from other sorts of technologies. They note that hammers, while excellent tools for driving nails, don't necessarily tell us much about the nails they drive. Instead, they say, models are more like thermometers which are used to explore some aspect (i.e. the temperature) of an object and render it epistemically accessible to the one who uses it (Morgan and Morrisson 1999, 11).

As helpful as this analogy is, it simply isn't true that we learn nothing about nails when we drive them with a hammer. We learn whether the nails are too soft, too short, or too long, or adequate for the job, as well as a broad range of subtle details about how to use nails in conjunction with hammers. Additionally, the example of a hammer is particularly evocative if for no other reason than it brings to mind Heidegger's use of hammers in his phenomenological analysis of tool use, with the hammer disclosing the nail in a particular way — as implicated in the social and practical nexus of construction projects. Thinking of the investigative capacities of hammers may seem a bit unconventional, but this is a real dimension of the normal use of hammers.

Measuring devices like thermometers seem somehow "more" or "better" suited to investigative projects, but does this mean that a thermometer is a model of my body when I stick it in my mouth, one that represents my body as a rising column of mercury in order to exemplify its thermal features? These examples stretch beyond the limits of how we usually think of models, and it's unclear whether this is a virtue or a vice for approaching models in a primarily technological way. It should be clear, however, that for better or worse, more must be said about the function of scientific models in order to distinguish them from other technologies.

Even for technologies that may fall within the common usage of terms like ‘model’ or ‘simulation‘ we may want to be on guard against an exclusively technological perspective. Consider Tarja Knuuttila’s account of a so-called “simulation” of language parsing in the *Constraint Grammar Parser for English* (EngCG) (Knuuttila 2006, 41-55). The EngCG disambiguates morphologically ambiguous words by tagging each word in a sentence with all its possible syntactic functions and discards those functions that are contextually illegitimate. A key feature of this parsing strategy is that it is a “reductive grammar” that operates by rejecting possibilities, contrary to traditional theories of natural linguistic performance that hypothesize “licensing” grammars that make positive conjectures of possible meanings. Constraint grammar parsing proves particularly robust for computers despite the fact that its method is unrealistic if taken as a simulation of human language use, and as such, Knuuttila concedes that “the EngCG-parser cannot be seen as a clear-cut representation of any part of the world that exists independently from our representational endeavors” (Knuuttila 2006, 47). I would go a step farther and suggest that this language parsing program has renounced its representational character altogether in exchange for technological reliability. Its ability to fit theory to target reality is so compromised as to make it an unlikely tool for scientific inquiry into natural language use, even if it is an impressive technology that computationally serves a function that a competent language user performs non-computationally. We have crossed over from representing inaccurately, even in the extreme ways that a narrow conception of fiction outlines, to not representing at all.[31] Without a convincing argument for how this technology can be used as a tool for studying the process that it "simulates," I would say that calling it a "model" or a "simulation" is to use these terms in a way that is, at best, inconsistent with how I am using them, and at worst, too loose to effectively discriminate between models/ simulations and other sorts of technologies.

It is not only their functions (i.e. to render target phenomena epistemically accessible) but also the *way* they achieve these functions that distinguishes models from other investigative tools. We cannot disregard the representational aspects of models altogether in favor of focusing on their technological functions, for models achieve their function *by* representing. Furthermore, models do not simply *make* representations (for thermometers and all manner of laboratory instruments make representations of phenomena without qualifying as models in any strict sense). Models *are* representations, or to use Morgan and Morrison's phrasing they are *representatives* of their targets (Morgan and Morrisson 1999). It is this special way of representing, being a "representative," or a "surrogate," or "stand-in" for actual target systems, that reveals how it is essential *that* fictional models represent their targets, while allowing the question of whether or not that representation is an accurate or truthful one to be secondary.

Again, insofar as fictional models effect a representation of their targets *as* other than they actually are, this is a two-stage representation. The first stage (the p-representation) may be utterly fictitious while the second (the representation of p) involves articulating the degree of fit between the potentially fictional system represented

[31] A considerably deeper discussion of this line between models and technologies that don't represent at all is taken up in the next chapter.

by the model and an actual target. The first stage is characteristic of all fictions, including literary and artistic creations. The second concerns the scientific function of affording epistemic access to actual targets. Thus, we can see that a truth-conditional analysis, which is focused on the adequacy of the representational relationship between a model and its target, makes the second stage primary. This approach tends to think of scientific fictions as scientific first and fictional second, as already possessing some degree of fit to actual phenomena and then inquiring into scenarios where this fit is strained. In order to get away from this truth-conditional approach and get more to the heart of how fictional scientific models are taken up and used in order to perform their unique technological functions, I suggest an inversion of this priority. It is the first stage of the representation-as structure that is primary, and scientific fictions are better understood as a particular type of fiction first, and scientific second.

A Performative, Mimetic Approach

The preceding section suggests that a truth-conditional analysis of scientific fictions obscures a more essential character of fictions, namely that these fictions are (insofar as they are fictions) employed with no regard to their truth value or the accuracy with which they represent their targets. Accordingly, as a corrective to this, I have recommended an approach that treats scientific fictions as fictions first and scientific second. On this approach, it would seem prudent to take some significant cues from thinkers who have treated fictions in a non-scientific context with an eye to adapting general theories of fiction to a specifically scientific context, rather than beginning from a general theory of scientific models and then adapting that to cases where models represent their targets fictionally.

While Vaihinger's theory of fictions *is* a general theory of fictions, and while he does regard fictions as having a potentially technological value, he does so within a problematically thin conception of technology. Fictions are technological for Vaihinger insofar as they are ideas, and he regards "thought as art" and "logic as technology." The result is a particularly mentalistic and subjectivistic theory of fictions, one which is easily taken up by theories of scientific modeling which have grown from extensions of thinking about scientific theories, and which tend to portray models as mental constructions that when applied to the world either work or don't. A thicker philosophical approach to technologies recognizes that they almost never go straight from invention to successful application, that technologies are picked up and used through a complex negotiation between the technology and user, and that this negotiation takes place through performances that are co-constituted by the technology, subject, and environment. Accordingly, a preferable theory of fictions would be one that acknowledges the affective impact of fictions on their audience instead of, or in addition to how they are invented by free subjects.

In the burgeoning conversation about fictions in science, such approaches are in the minority, but not unheard of. In particular, Anouk Barberousse and Pascal Ludwig (2009) forgo the more popular route to scientific fictions through Vaihinger and opt instead to base their analysis of fictional models on theories that have emerged from treatments of fiction in the arts. Specifically, they follow Gregory Currie and Kendall Walton by defining the function of a fiction as making "its interpreters imagine a certain intentional content, or to make-believe that some proposition is true." (Barberousse & Ludwig 2009, 58)

Following Walton (1990) in particular, we can say a bit more about such acts of imagination and make-believe. These acts follow loose rules of interpretation – what Walton calls "principles of generation" – which enable the capable interpreter to understand a fiction even when it deviates significantly from "serious," non-fictional experience. Some of these principles of generation are entirely fabricated and idiosyncratic to the fictional world in question, others are drawn from natural laws and social conventions that apply to the real world. Still others, and these are particularly important, emerge through the process of interpretation itself as *ad hoc* solutions for avoiding contradiction between fabricated principles and those borrowed from "real life." This means that the sort of contradictions that spring to mind when thinking of Vaihinger's genuine fictions are not simply overlooked (as Vaihinger suggests) or are barriers to practical utility (as Cohen argues), but, in Walton's account, challenges to be overcome in order for a mimetic performance to maintain internal coherence and not fall apart.

Many literary and cinematic science fictions (not to be confused with scientific fictions) exploit precisely this feature (*e.g.* the central dilemma of the movie *Back to the Future*, which finds the protagonist attempting to avoid erasing his own existence as a consequence of traveling back in time and inadvertently interfering with his own parents' romantic relationship). In addition to complicating and steering away from a narrow conception of fiction that demands genuine fictions to be self-contradictory, this approach also eschews a narrow conception that demands fictions to involve non-existent entities or physically impossible scenarios. All that is required is that the interpreter must imagine things other than they are, in a way that conflicts with empirical fact, in order to understand them. Counterfactual literary fictions, such as Len Deighton's *SS-GB*, which explores post-World War II Britain supposing that the Axis powers had won, involve scores of real (or at least realistic) entities playing out an alternate history. The central fictional twist in such works does not require the non-existence of its central characters. It simply creates a sharp discrepancy between the fictional world and the "real" world that is neither approximately true nor gracefully de-fictionalizable.

Additionally, Walton provides a thoroughly technological approach to fictions and make-believe — one that not only addresses the *ad-hoc* negotiations involved in putting principles of generation into action, but also takes up the material dimensions of mimetic practices.[32] His theory is, at bottom, one that identifies fictions a "props" for imaginative performances (or to use a more familiar vocabulary, "representatives," "surrogates," or "stand-ins"). In an instructive example, Walton describes a game of make-believe between two friends in the woods who decide to pretend that any stump they see is a bear. A stump, within the context of the game, is a prop, or stand-in for a "bear" that represents bears *as* stump-bears/bear-stumps. Bears, do, in fact exist, and behave in certain ways. Perhaps more importantly, people ought to behave in certain ways around bears, so there are principles of generation for the game that come from real life. But stumps are not, in fact, bears, and if one of the players were to point to a stump and yell, "Hey! There's a bear!" this would, strictly speaking, be false.

Yet within the context of the game, something very close to truth and falsity exists. As Walton explains, if the players were to examine what was initially thought to

[32] Interestingly, Barberousse and Pascal do not explore this dimension of Walton's theory of fictions.

be a bear more closely only to find that it was not a stump but a moss covered rock, then the previous warning would have been false not only with respect to serious contexts, but "false" within the playful one as well. And if, unbeknownst to either player there was an actual stump obscured where they could not see it, then within the context of the game, it is "true" that there is a bear waiting to be discovered. What is particularly provocative here is that the props as technological objects are not reducible to the imaginative faculty of the interpreter, but exist as semi-autonomous fictional objects on their own. What makes them fictional are the prescriptive rules that are to be followed when a player takes them up and does something with them; at least some of these rules must prescribe an act of make-believe.

Furthermore, as mentioned earlier, many of these rules, or principles of generation, are sorted out on the fly, in a negotiation between agent, prop and environment, and usually with some attention to alleviating some of the conflicts that arise between those principles of generation that are drawn from reality, and those which are fabricated. Walton's make-believers in the woods will soon realize that none of their "bears" ever move, which is unusual for bears (and not very fun), so some new rule will have to be generated to keep the game going.

It should be clear here that this theory of fictions, in addition to being technological, is also thoroughly performative. It is as much a theory of play as it is a theory of fiction. And insofar as scientific models can be taken as technologies, as objects that we take up and use, and which behave in sometimes surprising and unpredictable ways, a theory of models should be performative as well, and a theory of fictional models should be as much a theory of play as of fiction. Looking back on Maxwell's ether vortices, we may note how new fictional elements emerge as Maxwell begins to try and do something with his fictional model and adapt to difficulties on the fly. He may fictitiously posit an ethereal fluid as a medium for electromagnetic forces, but in order for that fluid to behave in the ways that fluids should behave, he must posit non-etherial "ball-bearing" molecules between the vortices. It becomes easy to see how, as Suarez notes, we "sometimes lose track of the fictional nature of the entity – in fact there is a sense in which for it to perform its function correctly, it is essential that the fictional nature of the entity be in some ways suppressed.[…] Something very similar operates in the case of scientific fictional representations." (Suarez 2009b, 169-170).

For the fictional scientific model, this imaginative performance constitutes the first of the two-stage representational scheme discussed earlier in this chapter as well as the previous one. It is a "*p*-representation" that represents something that does not necessarily have a referent. Insofar as a model is fictional, its use and interpretation combines real and fabricated principles of generation. Insofar as the model is scientific, however, it ultimately refers to a real target phenomenon (this is the second, "representation of *p*" stage), and it serves some practical utility that progresses a program of inquiry into that phenomenon.

A Theory of Fictional Models vs A Fictionalist Theory of Models

That fictions on this view borrow many of their principles of generation from non-fictional theoretical laws should square with our understanding of scientific models as generated from theoretical principles. A fictional scientific model will have some additional principles that would qualify as fictional twists on otherwise non-fictional hypothetical scenarios. Barberousse and Ludwig maintain that all scientific models involve such fictional twists that allow/demand their interpreters to imagine a scenario that conflicts with theoretical generalizations and cannot really exist. In a particularly telling example, they reject the prospect that such models could represent physically possible situations. "When using the billiard ball model of a gas in order to predict when it will reach equilibrium from such and such out-of-equilibrium state, a physicist by no means claims that quantum laws do not hold in some physically possible sample of some gas. The very notion of physically possible sample of a gas necessarily implies that it is subjected to quantum laws." (Barberousse and Ludwig 2009, 62)

There is something of the flavor of Cartwright's anti-realist position here, that theoretical laws "lie" and are only literally applicable in fictional models, and it is unsurprising that those who would paint all scientific models with the broad brush of fictions find resistance from even moderate realists. What Barberousse and Ludwig seem to be arguing for goes beyond the existence of legitimately fictional models in legitimate scientific practice, but rather a full blown *fictionalism* that runs through all of modeling and, insofar as models are an indispensable part of scientific practice through all of science.[33]

Ronald Giere in particular lodges a reactionary critique against a fictionalist approach to models in which he articulates three arguments against this view (Giere 2009). First, he points out that while scientists themselves invoke the notion of fictions, and though there may be "fictional" elements in many models, the models themselves are not fictions. This runs parallel to the argument we saw from Winsberg regarding how the

[33] Fine's essay on Vaihinger, tellingly titled "Fictionalism" makes a similar move, and Van Fraassen also appropriates Vaihinger in an effort to buttress his own anti-realism, claiming that the attitude which I say "is something like belief" characterizes all scientific models and attempts at explanation.

inclusion of fictional “silogens” does not compromise the non-fictional status of the larger model. I am uncertain about how compelling this argument is, both because I find it difficult to disentangle the parts of a model from the whole, and because the fictional parts are no less noteworthy even if they are part of a greater non-fictional model.

Giere’s second argument is far more compelling; many scientific models offer poor fit to their targets, or even no physically possible target, but these are abstractions and idealizations, not fictions, and trade on an exaggerated conception of non-fiction that would exclude approximations as realistic portrayals of phenomena. This serves as an argument against a full blown fictionalism, and it is noteworthy that contrary to Barberousse and Ludwig’s claim in their example of a classical “billiard ball” model of a gas, the realist who claims that such models are non-fictional is not suggesting a possible world without quantum laws. They are merely pointing out that these laws are negligible for the purposes at hand and can be omitted from the model. A standard for non-fiction that does not allow for realistic abstractions and approximations is impossibly high (or at least requires that our only candidates for non-fictions be the unwieldy and complex phenomenological laws that Cartwright describes).

Giere’s third argument runs dangerously close to a genetic fallacy, essentially claiming that fictionalism is frequently deployed as part of a broader anti-realist position, which is not necessarily justifiable. As sympathetic as I am to this argument, like many arguments regarding broad realist/anti-realist postions, Giere is begging the question here. Without an accompanying knock-down argument against scientific anti-realism, the association of a fictionalist approach to models with anti-realism is not a compelling argument against it.

Giere's somewhat reactionary take on fictionalism is somewhat warranted. He is concerned less with whether or not we could call models fictions than whether we should, and he is correct that a less than careful trend in referring to scientific models as fictions opens the door too wide to anti-realist tendencies. His own "perspictival realism" outlined in *Scientific Perspectivism* (2006), runs contrary to such tendencies. Like myself and many others, Giere is aiming to find a third option between the metaphysical excesses of capital-R-Realism and the aspects of anti-realisms that run counter to common-sense intuitions regarding the rationality of scientific claims about theoretical entities that allows us to take the claims of scientists at face value and proceed philosophically from there. In particular, Giere has the relativist/subjectivist excesses of constuctivist anti-realism in his sights. While he concedes the instrumental and even socially constructed contingencies of science, he maintains that these contingencies constitute an essentially human perspective on the world, one that is consistent with metaphysical hypotheses that various perspectives are, in fact, perspectives on unique, external events and entities, but does not necessitate such an hypothesis. Basing his argument on an analysis of color perception that reveals the "objects" of our perception as neither objective nor subjective, but rather products of "*interactions* between aspects of the environment and the evolved human [perceptual] system," he lays out an asymmetric, perspectivalist view of this interaction, one which is not an objectively sanctioned privileging of an evolving human perspective, but which is the only available perspective that ought to be of special interest to a scientific inquiry done by humans, for humans. As Giere asserts, an objectivist, capital-R metaphysical Realism isn't a viable ideal for this science. "We humans have a particularly human perspective on the world.

The world has no particular perspective on us. It does not care about us." (Giere 2006, 32) But to withhold the status of the "real" (in the sense that we apply it to everyday objects) from the picture of the world yielded by competent scientific inquiry, does a disservice to science. Building from a perspectivalist analysis of color perception, to extend this position to instrumental scientific observation, and finally a perspectivalist account of scientific theories and models that provides us with a moderate scientific realism that is "as much realism as science can provide." (Giere 2006, 16)

If Giere's critique is against fictionalism, then I believe it to be on target, even despite its tendency to beg the question. Previously I argued that a broad anti-realist stance is problematic for a progressivist theory of scientific fictions due to the fact that it leaves little to no room for scientific non-fictions. But Giere goes farther, denying that any scientific models should be regarded as fictions —"Should scientific models be regarded as fictions? […] I think the answer is decidedly negative" (Giere 2009, p248) — and if this is his goal, then I think he is wrong. The approach I have laid out sketches elements of a *theory of fictional scientific models*, not a *fictionalist theory of scientific models*. I do not hold that all models are fictions, offering a narrow view of fiction that excludes abstractions, approximations, and idealizations, but includes what I have called "non-de-fictionalizable" models.

These fictional models do not demonstrate an acceptable level of fit to their targets, but nonetheless, they can still be taken up as props in a thoroughly scientific practice. But this is a unique sort of practice, one characterized by make-believe. Due to the discontinuities between non-de-fictionalizable models and their target phenomena, the performances involved in this practice are not entirely continuous with ordinary, “serious” scientific practices. In the midst of these performances, we do not care whether fictional models represent their targets faithfully. But this should not be taken to indicate that their fictional character is without any epistemological significance.

Chapter 4
Simulations and Experiments

In chapter three, I gave an account of fictional scientific models as props for make-believe. Models have both a representational and a technological and performative character to them, and it was argued that a prop theory fares better at capturing both of these aspects than approaches that focus primarily on the representational and truth-conditional aspects of scientific models.

Any prop, insofar as it is taken up in an imaginative performance and used as a surrogate for something else, has some distinctive fictional character; and any model or simulation, insofar as it is used as a prop in some virtual performance and is a stand-in for some target phenomenon, also has the same sort of fictional character. Thus, *in a wide sense*, all models (provided that they are not, in fact, earnest theoretical hypotheses about the nature of their targets – in which case they are arguably not models in a strict sense) are fictions. However, *in a narrower sense*, where the distinction between the narrow and wide sense hinges on whether or not the "fiction" is question is gracefully de-fictionalizable to a more accurate representation of actual systems,[34] fictional models are somewhat less common.

One consequence of the shift to a prop-theory for models in general (and fictional models in particular) is a shift from attending to the model as something that is constructed and then just sits there representing, to attending to models as objects that we take up and do things with. This is not that radical, as it is in keeping with a conception

[34] This focus on de-idealization offers a further commitment to praxical dimensions of modeling, as de-fictionalization does require a relatively "hands-on" intervention on models, even if the prospect of de-fictionalization is only considered as only potentially executable.

of models as technologies – but it does signal a subtle shift in terminology. Insofar as we are concerned with what scientific models are like when taken up and put to work in a mimetic performance, we are concerned with *simulations*, and I will use these two terms more or less interchangeably in this chapter.[35] By insisting that we pay attention to the technological character of models and simulations we frame modeling/simulational practices as analogous to experimental practices, rather than as extensions of theory.[36] By applying a conception of models as props that are used in simulational performances, we open up a domain of quasi-experimental praxis in which models and simulations can be studied.

This performative shift is not new in philosophy of science. Writers such as Ian Hacking (1983) have offered ballast to the historically representationalist trend in philosophy of science by focusing on experimental intervention, and posit that the epistemological fruits of scientific inquiry hinge on the reliable use of "theoretical" entities in investigative practices. Others, like Joseph Rouse (1987) and Robert Crease (1993) have articulated the ways in which experimental

[35] This equivocation is fairly typical (and near as I can tell, fairly harmless) in literature on scientific modeling and simulation. If we are looking for a way to distinguish between the two, it seems straightforward enough to stipulate that a model offers a static and/or synchronic representation of its target, while a simulation represents its target diachronically as a dynamic and evolving system over time. I take it that within the context of this dissertation, nearly everything that can be said about modeling practices can also be said about simulation practices and *vice versa*. Wendy Parker (2008) makes an additional distinction between *simulations* and *simulation studies*, stressing that simulations are simply representations of "time ordered sequence[s] of states," while simulation studies involve interventions on a simulated system and therefore qualify as experiments. I am sympathetic to a distinction between simulations and simulation studies, but prefer to collapse both under the umbrella term "simulational practice," which includes the building, intervention, and broader use of simulations in a research program. I will not insist upon marking the difference with separate terms. Additionally, I resist Parker's suggestion that a simulation study is the same thing as an experiment (for reasons that will become clear) even though the two are admittedly similar in some important respects.
[36] In truth, models and simulations are both extensions of theory and quasi-experimental artifacts, but (as was the case for chapter three's attempts to balance questions of representational and technological character) I am concerned with which of these two is emphasized, and I am responding to what I take to be an historical overemphasis of simulation's relationship to theory and underemphasis of its relationship to experimental practice.

laboratory performances are self-contained productions in addition to/rather than representations of the natural world. And scholars taking more sociological approaches to laboratory practices such as Andrew Pickering (1995) and Bruno Latour and Steve Woolgar (1979) have illuminated the ways that experimental practices involve the complex coordination of diverse arrays of praxical contexts.

These broadly pragmatic approaches to studying science have provided ways of answering to or avoiding many metaphysical and epistemological anxieties, such as those concerning scientific realism/anti-realism or the relativistic slippery slope of social constructionism, by assuming (explicitly or tacitly) that these issues ultimately get ironed out in material practice. Provided that we attend to the material interventions of laboratory life, then, science ought not present us with any special metaphysical and epistemological problems beyond those that we might encounter with everyday objects.

For the most part, this approach has been relatively successful so long as the paradigm of scientific performance is traditional, “wet-lab” experimentation. However, serious concerns should be raised as to whether the performative idiom fairs as well for the “dry-lab” of computer simulations. Simulational practice involves a peculiar sort of virtual performance that reasserts many of the epistemological and metaphysical issues that are avoided by the pragmatic focus of a performative (rather than representational) idiom. Hacking’s instrumental realism, if adequately summed up in the maxim “if you can spray them, they’re real” seems to say precious little about the place that simulational practices occupy in science. When we open the range of investigative tools to include virtual “spraying,” *i.e.* using fictional props to learn about real objects, surely we need to be careful about the sense in which we consider these things to be “real.” On the other hand, Pickering and Latour’s affinities for the symmetries of actor-network theory may include simulations and their products all too unproblematically, as nothing more than another node in the mangle of the network, thus failing to account for the peculiarity of their imaginary character. For Crease’s part, while he seems interested in extending his performance metaphor to simulations (Crease 2006), his suggestion to include the “cyberstage” under his performance framework represents the major difficulty of assimilating simulations into research programs as a social one centered on cultural differences between computational and experimental scientific traditions, rather than the metaphysical, ontological, and epistemological difficulties surrounding the radical artificiality of simulated phenomena. Similarly, Rouse seems interested in drawing a strong analogy between the “micro-worlds” of the laboratory and the virtual world of simulations, but this analogy comes at the expense of losing any distinction between real interventions on real systems and virtual interventions on simulational ones.

For my part, I support the attempt to carry the intuitions of the performative idiom over to simulational practices, but I am only comfortable doing so as long as a distinction between experiments and simulations can be maintained. This distinction is a subtle one, but it allows for a discussion of how fictional constructions may give way to non-fictional scientific insights.

What sort of technology is a fictional model under a prop theory?

Taken as props for simulational practices, models are a distinct from other sorts of technologies. For one thing, not all technologies are tools for representation. Morgan and

Morrisson (1999) stress that models are tools for investigation. They draw our attention to thermometers to help explain what they mean by a technology used in an investigation, but, they quickly add, thermometers are not models. In the previous chapter, I questioned whether any technology could really be excluded from the class of investigative tools, noting that we need not think too creatively to see that a hammer is an investigative tool, even when used for its primary function of driving nails. What differentiates models from hammers and thermometers is not their investigative function (which is arguably common to all technologies), but their representational function.

However, this representational function alone is insufficient for distinguishing models as props from all other technologies. We may note that imaging technologies fulfill a representational function in a strong sense that hammers (and perhaps thermometers) seem to lack, but imaging technologies and imaging practices are quite different from models and simulational practices.

One facile distinction mentioned previously is that imaging technologies *make* representations while models *are* representations. But if we are shifting our attention away from models themselves and toward simulational practices, then this distinction gets a bit hazier; both imaging practices and simulational practices are staged performances that make representations of phenomena. However, we may distinguish these two practices from one another *at the level of practice* rather than at the level of the nature of the technological objects at the center of those practices.

Here, I offer a suggestion for provisional distinction[37] between what I would call *experimental representations* and *simulational representations*. Technologies that make

[37] The problematization of this distinction is a partial theme of this chapter, so I am, in effect, setting it up for the purpose of trying to knock it down.

experimental representations mediate relatively directly between humans and their objects of inquiry, and practices involving imaging technologies constitute a paradigmatic class of these types of representations. These technological practices fall broadly under what Don Ihde (2012, 155-170) has labeled variations on the *camera obscura* as an exemplar for modernist conceptions of scientific representation.[38]

In its most basic form, the *camera obscura* is simply a dark room with a small aperture through which light from the outside world enters and projects an image on the wall. A somewhat more complicated version is a photographic camera that focuses the image with a lens and fixes it on a chemical film, but the result is effectively the same; both produce a more or less isomorphic image of an object by interceding between the object and a subject who views the image.[39] The use of such technologies offer some advantages, e.g. in the case of photography, the relatively permanent "fixing" of the image so that it can be viewed by people at any time rather than only by those present when the image is made.

Variations on this become more complicated as the images produced become more distorted and less isomorphic to their targets. False color imaging, time lapse photography, passive spectrometry, and data models all do the same thing as a simple lens – intercede between the investigator and raw empirical data taken from the object of their investigation to produce a distorted image that is observed by the investigator. In exchange for the distortion, the investigator is afforded novel perspectives and access to

[38] Ihde also identifies the *camera obscura* as a theoretical model for epistemology at large, particularly one that is typical of a particularly modernistic, Lockean theory of perception.

[39] Use of terms like "isomorphic" or "similar" demand that we supply some basis for comparison. Insofar as a metaphysically realist referral of representations to objects in-and-of themselves requires an impossible (and even nonsensical) non-perspectival access to objects, this cannot be the basis of comparison. For the sake of the present discussion, I am content to understand "isomorphism" as a comparison to un-mediated experience (though this, too, is not without complications).

profiles of the object that are not afforded by naked observation or isomorphic imaging. Even more complicated are what Ihde calls "postmodern" imaging technologies, like active spectrometry and interferometry, where we have to poke and prod the object of inquiry in order to elicit an observable response. For all of these, there is what appears to be a relatively direct mediation between investigator and what's being investigated.

Technologies that produce simulated representations, on the other hand, do not intercede between investigators and the targets of their investigations. They are built by the investigator based (in part) on their theoretical understanding of the target phenomenon. At its most basic level, the distinction between the production of an experimental image and a simulated one is the difference between being presented with an apple and being asked to draw what you see, and simply being asked to draw an apple based on your prior experience and theoretical understanding of apples. In the technologized version of this, the human drawer is replaced by a machine that produces an image; and while it would be rightly noted that these machines are generally built according to some degree of theoretical understanding of the phenomena they represent, the fact remains that in the experimental case we can trace a relatively direct causal relationship from the produced image back to the target phenomenon, while in the simulational case we cannot. This is, admittedly, an oversimplification, and it may be objected that neither drawers nor builders of simulations don't typically simulate from first principles, but rather from a theoretical understanding that is causally traceable to empirical experience of the simulational target. Nonetheless, a clear enough difference in kind between technologies that manipulate empirical data and those that extrapolate from theory should be evident here.

It is, therefore, understandable that some critics of simulational practices in the sciences express anxieties that simulations do not tell us anything we didn't already know (which is obviously false) or that they don't really tell us about anything other than the virtual world they represent (which is not so obviously false, and a warranted anxiety). This is well expressed by John Maynard Smith who writes, "…I have a general feeling of unease when contemplating complex systems dynamics. Its devotees are practicing fact-free science. A fact for them is, at best, the output of a computer simulation: it is rarely a fact about the world" (Smith 1995, 30). Or perhaps we could quote Nigel Gilbert and Klaus Troitzsch, who write "The major difference [between an experiment and a simulation] is that while in an experiment, one is controlling the actual object of interest (for example, in a chemistry experiment, the chemicals under investigation), in a simulation one is experimenting with a model rather than the phenomenon itself" (Gilbert and Troitzsch 1999, 13). The simulation operates without an anchor to its target, as it were, and can't be reliably traced back to the object of inquiry itself. If we catalogue various types of technologies that mediate between humans and the world in order to construct a representation of part of the world, we can see that models and props are of a peculiar sort because they demand that their users engage in an act of make-believe that never quite comes into direct contact with their target of a mimetic performance.

The materiality of simulations

This is, of course, not to suggest that fictional models are entirely products of our imagination, or that we use them purely imaginatively. The notion of a prop conjures a sense of a concrete object, and Walton's discussion of mimetic play pays close attention to the ways in which the material affordances and limitations of props partially determine the rule-like structures of a game of make-believe. Similarly, many models are concrete objects and require that we materially manipulate them in simulation. Notably, physical scale models, like the Army Corps of Engineers' scale model of the San Francisco Bay and Sacramento/San Joaquin River system (Huggins and Schultz 1967, 1973), and simulations of levies and water flow in New Orleans during Hurricane Katrina (Interagency Performance Evaluation Task Force 2006) display such concreteness. Equally concrete (though significantly more complex in terms of their similarity to their targets) are so-called "model organisms" like fruit flies, nematodes, or laboratory mice, which often serve as surrogates for humans in particular, or broad classes of other organisms in general (Weber 2007; Ankeny 2007, 2009; Hubbard 2007).

It is tempting to consider that these sorts of physical models, given that they are comprised of "pieces of the world," are different in kind than the more abstract models and simulations found in computational practices which are more mathematical and less physical in nature (Sterrett 2005b). Yet, even computational modeling, which is admittedly on the more abstract end of the abstract-concrete spectrum does not evade the model-as-concrete-material-prop description. What appears to motivate such a distinction is the impression that simulational practices are strikingly similar to experimental practices, and that we should look to something other than thestructures of those practices to differentiate between the two.

While I agree with the broad intuition that we should look to articulate a strong analogy between simulation and experiment while maintaining a distinction between the two, I disagree with approaches that suggest that simulations are somehow nonmaterial abstract objects. Mary Morgan (2002; 2003) has argued that simulations can function like experiments, but argues that computer simulations are "nonmaterial experiments" where investigators intervene on abstract mathematical entities rather than material ones. As such, insofar as both have some part of the physical, material world as their target, material experiments should carry more epistemic weight. The weakness of this position, despite its surface appeal, becomes evident if we recognize that any computer program must be instantiated upon some material hardware. Accordingly, computer simulators are intervening on a highly flexible, but nonetheless physical system, the interpretations of which are frequently dependent on the embodied perception of simulation features via physical images. For those who would claim that the material character of computer hardware is merely accidental, while the essential nature of a computer simulation is abstract and mathematical, I offer Wendy Parker's response,

> First, the advocate of such a view should explain how intervention on an abstract system takes place during a computer simulation study and…[s]econd, if one or more abstract mathematical systems can be intervened on during a computer simulation study, presumably they will be ones that are realized by the computing system; but if it turns out that any abstract mathematical system realized by the computing system can be intervened on during a computer simulation study, then it seems that many traditional experiments can also be experiments on abstract mathematical systems, since their experimental systems also realize various mathematical systems. (Parker 2008)

The position I have mentioned above (that computational simulations are experiments without material interventions) has a more tentative and somewhat harder to shake version, namely that we may acknowledge the material character of simulations, even computer simulations, but still hold that what distinguishes simulations from experiments is that experiments are carried out with the *same kind* of material as their targets, while simulations are carried out with a different kind of material. Notably, while the first position (that simulations are abstract and nonmaterial) seems to focus on computer simulations, this slightly weaker position may cover many scale models and other concrete surrogate systems. Thus, proponents of this position might say, following Francesco Guala (2002), that the similarity between an experiment and its target is a "deep material" one, while the similarity between a simulation and its target is merely formal.

This point, while seemingly accurate and thus difficult to dismantle, can be parried by questioning the relevance of "deep material" similarities. The issue of whether an investigator is justified in taking a virtual system to be an adequate/non-fictional representation of its "real" target relies not on perfect correspondence, but rather only relevant correspondence (Grim *et. al*. forthcoming). In both experimental and simulational systems, there are features of the experiment or simulation that do not correspond to the natural world and this is usually the point of doing an experiment or simulation rather than directly observing a natural phenomenon. The features that do not correspond are often instrumental in rendering a phenomenon intelligible and analyzable, and are typically (in non-fictional contexts) intended to be inessential to the phenomenon in question. That the *streptococcus* is growing on a Petri dish and not on someone's tonsils is considered unimportant for determining its response to an antibiotic (and using the Petri dish avoids the discomfort of some prospective subject whose tonsils would otherwise be used). The torroidal shape of the field in a spatialized iterated prisoner's dilemma silmulation, or the square shape of the agents, or even the fact that the agents are not conscious subjects is not supposed to be of any consequence for showing how cooperative behavior can emerge from a very simple set of pre-programmed "preferences." These are differences between the target phenomenon and the representation used to study it, but they are differences in features that are intended to play no crucial role either in the real target phenomenon or in the representation. If there is a lapse of relevant correpondence, whether one of commission (adding a feature to the experiment or simulation that is absent from nature and that plays an active role in the production of results) or one of omission (leaving a feature out of an experiment or simulation that is present in nature and plays an active role in the natural process), then we would have cause for concern (and arguably, cause for regarding the system as a fictional prop); only then could we begin to make a good case for the critique that while our results tell us about our virtual representation of the world, they do not necessarily have any bearing on the world itself. However, so long as this relevant correspondence is met, any additional points of similarity are superfluous.[40]

Therefore any distinction that we might make between experiments and simulations on the basis of a degree of similarity (or a particular but arbitrary domain of similarity, like the kind of material used) is an inessential one if both have met the requirement of relevant correspondence. Pointing out that the experiment is more similar to its target than a simulation is irrelevant if the difference in similarity concerns only irrelevant features of the target phenomenon. If there may be good reason to suppose that both a simulation and an experiment fall short of the required relevant correspondence, there is still no basis for a strong distinction between the two, since the difference in

[40] This is well illustrated in an anecdotal experience, involving my ten-year-old cousin, who was building a terrarium for her science fair project to "study the effects of global warming" on a variety of plants. I advised her to make sure that her terrarium was enough like the "real" world to justify the generalization of conclusions about the terrarium to a larger natural ecosystem. I was surprised to find, on my next visit, that she had added toy cars and cardboard renderings of factories with smokestacks in an effort to make her terrarium "more like the real world."

similarity is a difference of degree, not of kind. Finally, an argument for a strong and meaningful distinction that pits examples of simulations that fail to meet a standard of relevant correspondence against examples of experiments that succeed in meeting that standard is a bad faith argument in that it compares “failed” simulations to “successful” experiments. It appears that in principle, simulations are capable of achieving relevant correspondence to their targets, and experiments are capable of failing to achieve it. Thus, absent any argument for why simulations on the whole are incapable of possessing the relevant features of their targets, there is no strong and meaningful distinction to be made between simulations and experiments based on the idea that experiments are more similar to the real world.

Guala’s emphasis on the “deepness” of the material similarity seems to suggest that this sort of similarity is somehow more important than a merely formal one, or put differently, the particular material of a phenomenon is an essential and relevant feature for its study. This is a bold claim, and highly contentious. While engaging this issue goes beyond the scope of this investigation, it is worth pointing out that many phenomena are considered to be “multiply realizable” and that while such claims are the source of deep debate (in, for instance, philosophy of mind and strong artificial intelligence theories), it is relatively less contentious to claim that a vortex in one fluid can be studied via a vortex in a different fluid with a similar viscosity, or that the flocking behavior of ducks is still the same flocking behavior if it is performed by geese.[41]

[41] This is not to say that the behaviors are indeed identical, merely that duck material and goose material alone do not constitute a difference significant enough to change the flocking phenomenon. One might even go so far as to consider the possibility that there is a deeper coordinated movement phenomenon shared by birds, insects, and fish that is only conventionally called flocking, swarming, and schooling respectively, but a justifiable distinction between such behaviors needs to be based on something more than the kind of material involved.

While the above constitutes a negative argument that challenges the assumption that "deep material" similarity is, *a priori*, a relevant type of similarity between a simulational/experimental system and its target, Susan Sterrett (2001; 2005a) makes a compelling positive argument for constructing successful (i.e. relevantly similar) simulations out of different materials. Sterrett focuses on the use of dimensionless constants to ensure that relevant similarity is maintained between a model and its target, specifically in cases where the relevant features of a phenomenon are already fairly well understood. For example, in fluid dynamics, a dimensionless constant called the Reynolds number (Re), which is a function of several parameters including fluid density, viscosity, relative velocity, and a linear spatial dimension, allow modelers to change any number of features of a fluid system, even substituting one fluid for another, and still ensure that the relationships between these features that give rise to observable phenomena are maintained so long as Re is the same for both systems. These dimensionless constants provide what Sterrett terms a "criterion for similarity."

Sterrett's point regarding dimensionless constants also offers a reiteration and partial elaboration of the point made in the previous chapter that a model need not be a fiction. Though we are casting this discussion now in terms of relevant similarity rather than the various standards for narrow/wide concepts of fiction discussed previously, it is intuitively absurd to suppose that a model that meets the requirements for relevant similarity to its target should be considered a fiction. Nonetheless, there is still an element of imaginative translation and prop-like character to concrete models that differ from their targets in scale or the kind of material they are instantiated in, even if they do maintain an abstract similarity in the relationships between relevant features. Here, I must emphasize that a fictional model must differ from its target in such a way that its distortions do not lend themselves to de-fictionalization in a straightforward way, and further, it should also be stipulated that the distorted/fictionalized features in question must be relevant features of simulation and/or the target system.

But while Sterrett makes a sound and helpful point here, it should be noted that she is describing phenomena which are already well known enough to have criteria for similarity already worked out and clearly applicable to real phenomena. Dimensionless constants frequently apply literally only to idealized systems. For example, the ideal gas constant, R, only accurately describes the relationship between the pressure, volume, and temperature of a gas if a variety of assumptions; that the gas in question is comprised of molecules which are volumeless, undergo perfectly elastic collisions, and experience no inter-molecular attractive or repulsive forces. While we may note that many real gas samples under "normal" conditions are close enough to this ideal for it to be approximately true in most contexts, this requires not only that we have already worked out this criteria for similarity, but also that we already know the contexts in which it can be reliably applied as even approximately true. In less well-known and significantly more complicated systems, such as those involving model organisms, identifying the relevant features for a criterion of similarity is a far messier and tentative affair. The neutral analogies (viz. those which we do not yet know whether they are positive or negative, relevant or irrelevant) between a genetically standardized population of fruit flies and humans, or any other "wild" organisms are significant enough to present barriers to graceful de-fictionalization of the fruit fly model to the human target. And even if the use of a model organism begins from an hypothesis that there is relevant similarity between the model and target, such hypotheses are constantly challenged. Yet we continue to use such model organisms, even after challenges to such hypotheses. As Creager, Lunbeck, and Wise (2007, 7) say, the process of using a model organism as a surrogate for other target organisms "is best described in terms of an unpredictable relevance … for genetic homologies that can underlie divergent physiological properties. The molecular biology of organisms such as the worm began as unabashedly reductionistic in spirit, yet the specificity and particular features of the model systems have remained experimental resources rather than mere complications."

Distinguishing between simulations and experiments

The preceding arguments appear to blur the distinction between experimental practices and simulational ones. If we cannot say that simulations are non-material, or that their lack of a "deep" material similarity to their targets marks a significant distinction between simulations and experiments, then the elimination of these alternatives strengthens the case for the distinction made at the beginning of this chapter; that experimental technologies mediate relatively directly between humans and the world while models/props that do not. We may reformulate this distinction in more practice oriented terms, stipulating that experimental practices start with their target phenomena and "cut down" by removing and controlling confounding variables, while simulational practices start from scratch and "build up" [42] according to theoretical principles. Both

[42] It should be pointed out here that "cutting down" and "building up" may imply a misleading oversimplification of experimental and simulational practices. Features are not simply added and subtracted, but also manipulated and distorted. I mean to imply something looser here, where "cutting down" includes any manipulations and distortions that make an experimental system less similar to its "wild" target, and "building up" includes manipulations and distortions that make a simulational system more similar to its target.

produce distorted versions of their targets. Experimental scientific technologies mediate by manipulating the world so that we can experience a manipulated version of it. It is plausible to assert that experimental practice employs such technologies to produce a reduced version of phenomena that instantiate a world where prospective theories are true. Simulational props also produce scenarios where theoretical rules play out seamlessly, though they do so through a process of building a scenario from surrogates composed of "pieces of the world." (Sterrett 2001)

Earlier, I mentioned that it is unfair to try to differentiate between simulations and experiments by comparing "successful" experiments to "failed" simulations, where success and failure is based on the achievement of relevant correspondence between a system and its target.[43] Prudence demands that we address the fact that this ideal form of success is uncommon in active scientific inquiries. Eric Winsberg makes this point in the course of a critique of Morgan and Guala's positions on the materiality of simulations similar to those I have laid out above (Winsberg 2009). The trouble with treating simulation and experiment as "success-terms," as Winsberg sees it, is that if simulation and experiment are defined narrowly as successful simulations and experiments only, we can only use these terms to refer to simulations and experiments where the relevant features of the target phenomena are adequately represented and no additional non-corresponding features that play a crucial role in simulational or experimental results are present in the simulation or experiment.[44]

[43] It should be noted that this is an excessively narrow way of defining success and failure. A more adequate definition should take into account precisely what the aim of an experiment or simulation is. This is taken up in CH 5.

[44] While Winsberg goes on from this point to articulate an epistemological distinction (Winsberg 2009) between simulation and experiment, this position is convincingly challenged by Frigg and Reiss in "The Philosophy of Simulation: Not New Issues or Same Old Stew?" (2011) My tack, in this chapter, is to offer a praxical distinction, rather than an epistemological one, between simulation and experiment.

It becomes apparent why, notwithstanding certain specious attempts at distinguishing between simulations and experiments (e.g. on the basis of materiality), we might have trouble distinguishing between simulation and experiment, since their conditions for success appear to be identical. Additionally, since determining what the relevant features of a phenomenon are is precisely what is on the line in virtually all interesting cases of open and progressing programs of scientific inquiry, scientists are, for the most part usually working with "un-successful" simulations and/or experiments. The utility of "experiment" and "simulation" as "success terms" is, therefore, problematically limited.

Instead I follow Winsberg and suggest that we look to *promising* simulations and experiments, that is to say ones that are not quite yet successes, but aren't so flawed as to be disregarded altogether. Though both reach "success" in satisfying some criterion of similarity to their targets, it is important to note that experiments tend to begin with the target phenomenon more or less in tact, and "cut" away, while simulations tend to begin from nothing and "add" features piecemeal, building virtual phenomena that afford an increasing fit their targets as they develop.

This description raises the question of why an experimenter would intentionally remove features from the target, since that could only lessen the chances that their experiment would correspond to it in the relevant ways. The answer to this question is that the satisfaction of similarity criteria is not the only desideratum for either experimental or simulational systems. If it were, then the “wild” natural world would suffice our scientific observations and experiments and interventions would be impediments to the accuracy of our scientific theories. We also want these systems to provide some kind of novel epistemic access or insight to their target phenomena. Experiments control variables to render phenomena simpler, more intelligible, more consistently replicable, and more easily manipulated. Ideally, an experimental laboratory system does more than meet criteria for relevant similarity to phenomena in the “wild.” It produces a manipulated version of the phenomenon that eliminates extraneous and irrelevant features, reducing the experimental system to only those features that make a difference. Promising experimental systems approach this ideal.

The promising simulation, while generally less similar to its target than the promising experiment, offers something in exchange for sacrificing similarity. It is more intelligible, more consistently replicable, and more easily manipulated than its experimental counterpart. Though both simulations and experiments seek the same sort of mean, balancing similarity and epistemic access to their target phenomena, they approach this mean from opposite directions. It should not be surprising, then, if a research program built on promising simulations and a research program built on promising experiments, might paint very different pictures of the same target phenomena. This reveals a praxical and “paradigmatic” distinction between experimental and simulational practices, stemming from the fact that simulational practices tend to “build up” while experimental practices “cut down,” and resulting in markedly different ideas about what “promising” means as well affecting potentially different representations of target phenomena in this promising phase.

Laboratory Fictions and Material Manipulation

The preceding section articulates a means for distinguishing between simulations and experiments where simulations can be, but are not necessarily fictional performances, employing models/props which are significantly dissimilar to their targets. However, given that experimental practices also introduce systematic distortions in the natural phenomena they investigate, it becomes plausible to consider them as candidates for fictions as well. This much is suggested by Joseph Rouse (2009), who argues from his conception of the laboratory as a "micro-world" that laboratory experiments have a distinctive fictional character. At the heart of his claim for parallels between the laboratory and fictions is the systematic, autonomous, self-enclosure exemplified by both, and the fact that the representational character of the laboratory only emerges in discursive practices that are distinct from the technics of laboratory work itself.[45]

Rouse's suggestion is a fairly bold one, and I'm not particularly inclined to follow him in it, mostly because he is playing a bit loose both with the concept of "laboratory experiment" and with the concept of "fiction." It is fairly telling that he leans heavily on examples of model organisms in his argument, which, following the distinctions between simulational and experimental practices that I have laid out in this chapter, belong more to the realm of simulation than to the realm of experiment. Furthermore, self-enclosure,[46] while a distinctive aspect of fictional constructions, is not enough to qualify under the

[45] *c.f.* also Rouse 2002.
[46] This aspect of fictions, games and play is taken up in Ch 6.

narrow definition of fiction that I have articulated earlier. Nonetheless, his thesis is a provocative one, and raises some significant issues regarding the material aspects of simulational practice.

First, and foremost, he offers a more nuanced and complicated account of the relatively simplistic building-up and cutting-down model I have just suggested. "Experimental work does not simply strip away confounding complexities to reveal underlying nomic simplicity; it creates new complex arrangements as indispensable background to any foregrounded simplicity" (Rouse 2009, 40). Features are not simply added or subtracted in materially instantiated systems; rather any change is inevitably accompanied by a host of additional changes. This is apparent even in very simple examples of "laboratories." A salt-water fish tank may, on first pass, appear to simply be a reduced version of the ocean, containing the same fish, water, and plants that are found in the wild, but in a more controlled and accessible environment. It is tempting to say that all we have done with the fish tank is subtracted the rest of the ocean, leaving a smaller and more manageable portion of it. However, to realize this subtraction we must also add the tank. And we aren't done there; we must also add a pump to circulate and aerate the water, a filter to ensure the pump doesn't clog and break down, and we must add chemicals to ensure that the pH and other chemical equilibria are maintained. "Subtraction" always requires "addition," and "cutting down" is always accompanied by some "building up."

This point is not restricted to experimental systems, but applies to models as well. Scale modeling via the use of dimensionless constants requires that altering one parameter of a model system be accompanied by the alteration of others if we wish to maintain the abstract relationships between relevant features of the model that are present in the target. Often, in scale modeling, it is *necessary* to use alternative materials in order to account for the density, viscosity, or some other property that would complicate the scaling process if one were to use the same material found in the target system. Early attempts at aviation were plagued with such issues, as engineers found that small models that performed well performed much poorly when scaled up (Sterrett 2005c). Making models more similar to their targets is not as simple and straightforward an affair as nudging individual parameters of the model to match the target.

This not only presents a complication of the notion that simulations "build up" while experiments "cut down," but also introduces a vexing complication for the notion that "de-fictionalizing" a physical model can ever be done in a straightforward, continuous, and stepwise manner. It is easy enough, in the abstract, to change a single parameter without introducing a host of further complications, but the work of material intervention is far more complex and, to borrow a term from Pickering, "mangled."

Hence, we might draw an ironic conclusion from these insights; namely that under the stipulation that fictions, narrowly defined, are not amenable to a straightforward de-fictionalization process, materially instantiated, concrete models and simulations tend to be more fictional than abstract ones. This is strange, given initial intuitions (expressed by folks like Morgan and Guala) that the materiality of experiments makes them *less* likely candidates for scientific fictions than so-called non-material, abstract, computational models and simulations. Systems like concrete scale models or model organisms may be more materially similar to their targets than their computational counterparts, but the complexity of manipulating, controlling, and intervening on those systems drives them into a self-enclosed laboratory "micro-world." Unraveling the mangle of distortions, additions, and subtractions to "de-fictionalize" a laboratory system to square with the "wild" is difficult, and far from straightforward.

By comparison, computer simulations allow simulators to abstract and idealize with relative impunity. This is not to retract the arguments made earlier regarding the unavoidable materiality of modeling *in-silico*. Decisions about how to simulate computationally are frequently constrained by material issues like memory and speed,[47] and other material artifacts, such as the discretization of continuous functions are fairly well known (even if infrequently discussed) issues stemming from the fact that computer simulations are not quite the abstract entities we sometimes imagine them to be. Software must always be coded on hardware that exists within a complex field of material and technological affordances and limitations.[48] Nonetheless, it is clear enough that the relative abstractness of a computer simulation avoids much of the complexity that accompanies intervening on relatively concrete systems. While reducing a sample of fish in water to a manageable finite system requires the addition of a tank (and many other concomitant additions as well) introducing artifactual distortions requiring a fish that reaches the edge of the tank to stop and turn around, a computational simulation of schooling agents intended to represent fish can be placed in a finite system without such additions, *e.g.* on a spherical or torroidal surface where an agent that reaches the edge of the field simply reappears on the other side of the field.

[47] Recall that Winsberg's "silogens" were a consequence of having to simulate a cracking silicon brick in three regions because using a quantum mechanical model for the whole system was too computationally expensive. Without the material limitations of the computing system, there would be no need for "handshaking" algorithms, and thus no need for fictional "silogens."
[48] See Parker 2008 for a thorough account of material artifacts in computational simulation.

But we need not abolish our intuitions about simulations being more fictional than experiments. It is noteworthy in the example just given that the trade-off for building a fish simulation without artificial boundaries is to build one with a topology that is not only experimentally impossible, but quite different from anything encountered in the "wild." The relative ease with which simulational systems are manipulated is often, and in large part, because the simulational system allows for scenarios that are mathematically possible, but not physically possible.

Furthermore, we need not abolish our intuitions about "building up" and "breaking down," though it should be noted that the real task of "adding" and "subtracting" features is quite complex in practice. As a general praxical paradigm, experimenters do aim to make complex phenomena progressively simpler and more epistemically accessible, and simulators do aim to build relatively simple systems into progressively richer and more complex ones. The broad epistemological consequences of these two approaches, that simulations offer relatively superior access but dubious similarity to their target, while experiments do the opposite, still holds.

Lastly, we need not let go of the distinction between straightforwardly de-fictionalizable and non-de-fictionalizable models, though it is clear that some adjustments are necessary. The movement from an abstract representational idiom to a material performative one has severely complicated the notion of straightforward de-idealization since, as we have seen, intervention on any material system has diverse and sometimes unanticipated corollary effects that must be accounted for and dealt with. But this pragmatic problem has a pragmatic solution. At stake in the question of de-fictionalization is whether or not one can translate the insights from a distorted model system to its target in such a way that those insights can be preserved for a range of intermediately distorted systems between the model and target. In an experimental system, even though our interventions are complex and layered, any distortion that has been made can be un-made. If we were able to manipulate a "wild" natural phenomenon into a laboratory system, we should, in both principle and practice, be able to manipulate the laboratory system back. Such an undoing of what has been done is not available to models and simulations. There is a large, uncharted gap between a computational simulation and its target, between the concrete scale model and its target, and between the laboratory mouse and the human. We may be able to articulate the territory between models and their targets, but not by going back the way we came.

In making these adjustments, we may have to account for a wide variety of "hybrid" systems, like flight simulators, wind tunnels, or psychological experiments, which place real objects of investigation in simulated situations, but this is a taxonomical issue, and arguably represents the sorts of scenarios that both simulational and experimental research programs aim to converge on – a balance between epistemic access and similarity to the target phenomenon.

We may also find that a large number of what we have generally taken to be experimental practices are, in fact, simulational ones that study surrogate systems. Research on model organisms, for example, has historically been taken to be experimental in nature. But this may only be the result of beginning with relatively complicated surrogate systems that must be controlled to be effectively understood, and thus follow an experimental trajectory in the short term. On a wider scale, the study of model organisms as surrogates for target organisms follows a more simulational trajectory; moving from relatively simple and dissimilar models, to more complex and similar ones.

Accordingly, it is understandable if researchers working on such projects do not take themselves to be working with surrogates at all, but rather simply studying fruit flies, worms, or mice, etc. But the motivations for such projects are teleologically oriented at a wider application, and the aims of the larger research program do exert a significant influence (Rouse 2009, 46; Ankeny 2007, 46-47). When they do not, which is to say, when a community of researchers forget that their research is simulational rather than experimental, an entire research program may find itself needing to reverse course and unravel years of work.

Along these lines, we may also have to acknowledge that simulational practices and experimental practices do not stand on their own, but instead are part of a complex set of practices that are sometimes simulational in character, sometimes experimental, and sometimes ambiguously between the two.

A Case Study: Wake dynamics for cylindrical obstructions

As an illustration of these points, the following case-study[49] is instructive. It follows a decades-long debate in fluid dynamics regarding how to interpret experimental data in a system where a fluid flows past a cylindrical obstruction.

The debate in question begins with the 1954 publication of a dissertation by Anatol Roshko concerning the transition from a stable phase marked by regular vortex evolution to turbulent flow in a system where a fluid flows past a cylindrical obstruction.

[49] This case was brought to my attention by Isabel Peschard (*c.f.* Peschard 2009), and my exposition draws heavily from her treatment.

Roshko identifies this stable range to occur between Re=40-150,[50] and a range characterized by turbulent flow with periodic vortex production between Re=150-300. Below the stable range (Re<40) the flow is laminar, with no vortex evolution, and above the quasi-stable range (Re>300) the flow is completely turbulent.

Of key importance for the subsequent debate is Roshko's account of the stable range (Re=40-150), where a phenomenon known as a "von Karman vortex street" occurs. In this stable range, the vortices produced alternate between eddies that rotate in opposite directions. Roshko describes a linear relationship in the stable range (Re=40-150) between vortex evolution frequency and the Reynold's number.

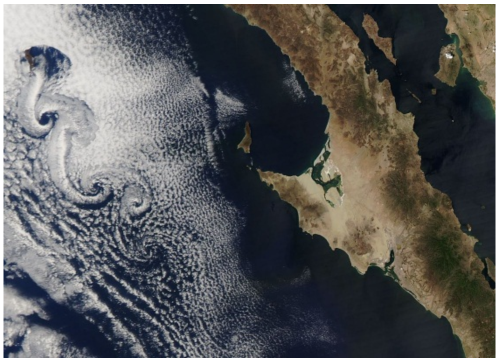

Fig 1. A naturally occurring von Karman vortex street caused by an island obstructing a weather system off the coast of Baja California, mexico. (source, http://envsci.rutgers.edu/~lintner/teaching.html)

These results were not particularly surprising and were corroborated by other subsequent experimental findings until 1959, when D.J. Tritton published a paper reporting a linear *dis*continuity in the middle of the stable range, occurring around Re=70-90. The departure from Roshko's linear formula was fairly subtle, but significant in that it posed a potential complication of an otherwise very tidy account of the experimental system.

[50] Re is the Reynolds number, a dimensionless constant described earlier in this chapter. In this system, since the only variable parameter of the Reynolds number is the upstream velocity of the fluid, it may be helpful to simply consider this value to represent the upstream velocity.

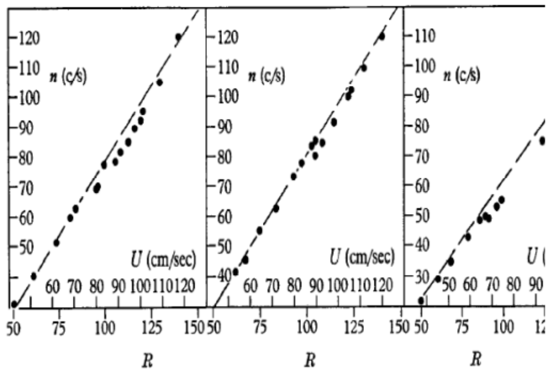

Fig. 2 – A plot of vortex shedding frequency against Re for three runs of Tritton’s experiment. The data points represent Tritton’s data and the dashed line represents Roshko’s linear formula.(source Peschard 2009)

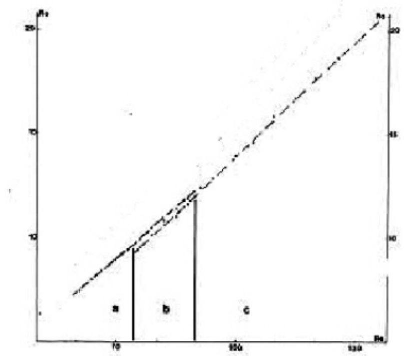

Fig. 3 – A “corrected” regression of experimental data highlighting Tritton’s reported discontinuity between Re=70-90. (source Peschard 2009)

The discontinuity was also accompanied by an observable effect in the vortex “sheets,” causing them to be slightly oblique to the cylinder, rather than parallel. Such 3-dimensional effects were considered to be uniquely characteristic of the quasi-stable range with turbulent flow and periodic vortex shedding, and were previously considered to be absent in the stable range being described by Roshko.

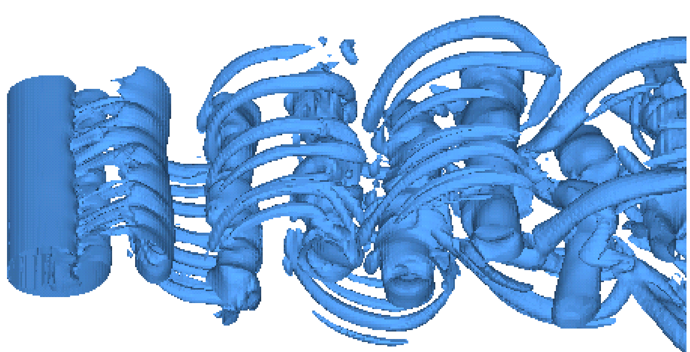

Fig. 4 – A digital rendering of parallel vortex sheet production.

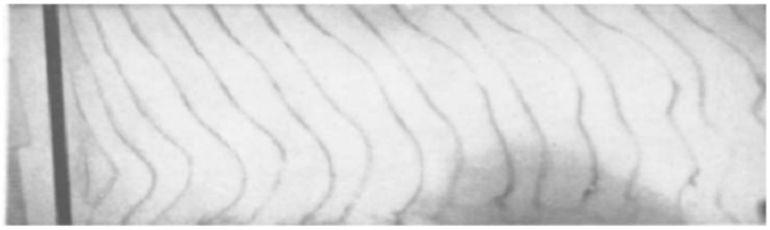

Fig. 5 – A photograph (looking down at the experimental apparatus, cylinder on the left) showing oblique vortex sheet production over the range where Tritton reported a discontinuity in the relationship between vortex frequency and Re. (souce, Peschard 2009)

Debate ensued concerning whether or not the discontinuity identified by Tritton was "real," or the product of experimental artifacts. The following 30 years saw many attempts to settle this issue by experimental means:

- Gaster (1969) – Shows that more discontinuities can be introduced by introducing irregularities in the cylinder and upstream fluid velocity. Also, linear discontinuity and oblique shedding could be reduced (but not eliminated) by reducing the length of the cylinder (which ostensibly serves as a means for reducing the relative impact of irregularities in cylinder diameter).
- Tritton (1971) – Discontinuity and oblique shedding persist even under more carefully controlled experimental conditions. Furthermore the range in Re where they occur is regular and predictable.
- von Atta and Garib (1987) – correlation between discontinuities and presence of microvibrations in the cylinder. (introducing another potential variable in addition to irregularities of the cylinder into the mix)

It is worth noting that the ability to increase and decrease these effects with the introduction and reduction of irregularities were neither successful in eliminating the effects in question altogether nor in explaining why they reliably occurred at the ranges they did, both points that Tritton was quite vocal about. But, the fact that observed discontinuitites cannot be eliminated altogether can be explained by noting that any experimental apparatus will always contain some irregularities. "Real" cylinders have discrepancies in their diameters along the cylinder axis, and really producing the upstream flow for real fluids is unavoidably irregular.

Also notable is that we are not afforded an explanation of the apparent regularity of the irregularity simply by asserting that it is an effect that is intrinsic to the phenomenon and not a product of experimental artifacts. It is certainly curious that the discontinuity occurs in a predictable range, but neither camp in this debate has an explanation for why, and it remains unclear whether this phenomenon is even something that demands explanation. In short, the experimental results over the course of 33 years were inconclusive with respect to the debate over whether the data discontinuities and oblique wake shedding were intrinsic to the phenomenon or were experimental artifacts. At this point, the debate had stalled and appeared to be unresolvable with the available experimental methods.

Then, in 1989, George Karniadakis and Michael Triantafyllou published a study called “Frequency selection and asymptotic states in laminar wakes” in which they report the results of a computer simulation of the experimental system in question with no irregularities or vibrations. In these simulations, vortex shedding frequency vs. Re is linear, with no discontinuities. The simulated system used was a two-dimensional system, a modeling decision that was motivated by computational feasibility but was also consistent with the Gaster hypothesis that shortening the cylinder results in decreased irregularities in its diameter, and also seemingly consistent with an ideal cylinder of finite length where any perpendicular cross section is identical to any other. It appeared that this simulation had settled the matter.

However, as Isabel Peschard (2009) points out, it is not obvious that the simulation has adequately represented the experimental system. If it is only instantiating an idealized version of real systems, one which carefully controlled experiments can approach but may never achieve, then we may regard it as relatively non-fictional (i.e. fictional in only the wide sense of the term). If there is something else going on, some crucial sense in which the simulation has failed to represent its target, then the absence of discontinuities in it need not tell us anything at all about "real" experimental or "wild" fluid dynamic systems. In particular, the 2-D character of the simulation is problematically ambiguous. Is a 2-D "cylinder" a simple abstraction or an approachable idealization, or does it involve some deeper distortion?

In fact, there is a potentially important way in which a "2-D cylinder," one which assumes that any point along the cylinder axis is identical to any other, differs from actual cylinders. Even in a perfect "real" cylinder, one with no irregularities or vibrations, points along the cylinder's axis are not interchangeable because they have different distances from the *ends* of the cylinder. However we choose to interpret the 2-D simulation, whether as infinitesimally short, finitely long, or even infinitely long, we are operating under the assumption that all points along the cylinder's axis are interchangeable and that the experimental system it exists in a virtual universe that is exactly as "long" as the cylinder. The simulated cylinder is one with no ends.

As Peschard says "taking the simulation as relevant is, consequently, taking the ends of the cylinder as not being relevant" to the target phenomenon. (2009, p 15) If the ends are relevant, then any real cylinder (having ends) will differ in a significant way from the simulation (with the possible exception of cylinders that approach infinite length, which is notably different than the Gaster hypothesis which asserts that discontinuities will vanish as the cylinder becomes shorter). C. H. K. Williamson recognized this, and in a 1989 study attempted to explore the influence of the cylinder ends, which up until then had not been taken to be a significant feature of the experimental/simulational systems, on wake dynamics.

Here it must be mentioned that the artifice of the laboratory system for these experiments traditionally involves the necessity of another hitherto unmentioned feature of the experimental apparatus. The experimental work described above is always done in tanks where the presence of the tank creates non-uniformities in fluid velocity. Specifically, a moving fluid in a tank or pipe slows at the edges of the container. To guard against the influence of this on the experimental system, it is common practice to place endplates on the cylinder to shield it from the effects of the tank edge.

Williamson found that by angling the endplates inward on the upstream side, he was able to reproduce the simulational results of no discontinuities in the vortex shedding frequency and parallel (as opposed to oblique) vortex "sheets."

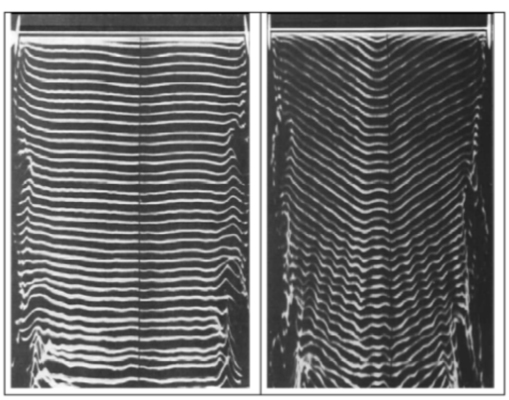

Fig. 6 – Williamson's experimental apparatus with angled endplates on the left, and perpendicular endplates on the right. (source, Peschard 2009)

At this point we have reached a temporary stopping point where the debate triggered by the discrepancies between Roshko's linear formula and Tritton's identification of a discontinuity in the relationship between Re and vortex frequency stabilizes. It is clearly no exaggeration to say that the simulations run by Karniadakis and Triantyfallou were crucial in the evolution and eventual stabilization of this debate. As Peschard argues, the model/simulation is not simply an articulation of theory but influences the construction of the theory by revealing a previously overlooked feature of the system as salient. This shows a discovery (rather than a mere demonstration, elaborating the consequences of previously known theoretical principles) emerging out of a model (rather than out of experimental data).

Peschard also draws our attention to the ambiguity of what counts as an "essential feature" in this experimental system. Are the end-effects "essential?" Clearly they are not, if by essential we mean necessary, since they can be made to disappear. But they are clearly significant, that is to say, they make a difference. Potentially, we can argue over which, if any, of the possible versions of the experimental aparatus is the "real" phenomenon, but this is not a particularly productive argument. Anyway, there are more interesting insights to be taken from this case study.

Interpreting the case study

Of particular interest for this essay is the possibly fictional character of the simulational system. Is the simulation a fiction? Initially it is intended to represent a 2-dimensional cylinder. Vaihinger's notion of a fiction in the narrow sense is that it is a self-contradictory concept. As cylinders are, by definition, 3-dimensional figures, but this cylinder is 2-dimensional, it would seem that the cylinder qualifies. Similarly, Suarez stipulates that a fictional object must be non-existent, and since 2-dimensional cylinders are, in addition to being logically incoherent, not to be found anywhere in the world, again the simulation qualifies. However, the 2-dimensional cylinder is also an abstracted cross section of a 3-dimensional cylinder, or, alternatively, a real cylinder may be shortened arbitrarily to approach an infinitesimally small axial length; so the 2-D simulation is an approachable ideal limit of potentially real cylinders. This is the interpretation of the simulation that is consistent with the Gaster hypothesis that shortening the cylinder length decreases the effects of various unavoidable experimental imperfections. Following this line of reasoning, the 2-dimensional cylinder is also interpretable as a 3-dimensional cylinder with no imperfections, and again, such cylinders are unlikely to be found in the world, but are approachable idealizations. While some of these interpretations could be argued to fit Vaihinger's or Suarez's narrow definitions of fiction, they do not fit the one settled upon in chapter three since they are all idealizations that can be approached by real experimental systems.

However, the turning point in the debate hinges on a novel interpretation of the simulation as representing a cylinder with no ends. This is not only logically incoherent and thus not available in the real world, it is not even approachable. This interpretation is a fiction by any of the narrow definitions we have considered. Finally, the simulation is reinterpretable one last time, as a cylinder of finite length, and with ends, but the effects of those ends on the wake is absent or negligible. Not only is this achievable in principle, but Williamson realizes it in practice by angling the endplates. Of key significance in the series of interpretations is the fact that the simulation sustains multiple (but not indefinite) interpretations. It is what Don Ihde has referred to as a multi-stable representation (see Ihde 2012, 45-54). The hermeneutic principles that guide the available interpretations, what Walton calls principles of generation, are influenced not only by the simulational model itself, but by a host of geometric principles as well as principles derived from the trajectory of the experimental research program. The available stable interpretations are not infinite, but surprisingly rich, and the break in the debate relies upon a novel, and thoroughly fictional re-interpretation of the simulation as a finite cylinder with no ends.

Also significant is a marked inversion that occurs with this novel, fictional interpretation. The initial interpretive possibilities of the simulation all situate the simulation as a representation of an idealized version of the experimental apparatus. Upon revelation of the simulation as potentially representing an utterly fictional, and unapproachable scenario, the simulation temporarily ceases to be a representation of anything real and is self-sufficient. The next step in the evolution of the debate is affected by a reversal of the representational relationship between the simulation and experiment, one where Williamson's apparatus is an attempt to instantiate the functional equivalent of a cylinder with no ends. When he successfully does so, the simulation is no longer a fiction, but retrospectively, a blue-print for a new variation on the experimental system.

Furthermore, we see the convergence of the experimental and the simulational systems, the so-called "wet" lab and "dry" lab. This convergence is produced by a dialogue between the two where the simulation is repeatedly re-interpreted until it reveals a novel and prospectively salient feature of the phenomenon (namely, the effects of the ends of the cylinder), at which point it becomes the pattern on which a new experimental apparatus is based. The fictional system represented by the simulation becomes the "center of gravity" for both the theoretical and experimental programs in play. The "representation *as*" structure of the simulation, as mentioned previously, not only informs how we see the simulated cylinder, but reveals a new way to see "real" target cylinders.

We may ask just how "real" the phenomenon that the simulation and experiment converge on is. If this question amounts to the previously mentioned haggling over which version of the experimental apparatus is more "real" then this is a pointless question. They are all equally real. But what is produced is a successful model, an exemplary case of obstructed fluid flow. It is an instantiation of a system where the theoretical principles are literally true. Perhaps more importantly, it is also a system from which any case of obstructed flow found in nature can be regarded as a variation.

Conclusion

I opened this chapter with the question of whether a performative approach to scientific practice can accommodate the peculiarities of fictional modeling and simulation. The subsequent discussion indicates that it can. A prop-theory of models suggests a materiality to simulational practices, and all forms of simulation, even relatively abstract computational ones, involve such a material character. Thus, efforts to distinguish simulations from experiments by supposing that the former are immaterial are wrong, and suggestions that this distinction can be made on the basis of the "kind" of material are of dubious epistemological consequence. We can discriminate between simulational and experimental practices by reference to a general trend where the former "build up" and the latter "cut down," but this is significantly complicated by the complexity of material intervention. We cannot simply "add" or "subtract" features without accommodating corollary effects of these manipulations. Still, through a dialectic between the "wet" lab and “dry” lab, both converge on a successful model. I illustrated such a convergence by way of a case study where a stalled scientific debate was stabilized by the introduction of a simulation.

While this case study did not afford an opportunity to examine the materiality of interventions on a simulational system, it did illustrate the “multistability” of the simulation. Perhaps more importantly, it showed that this “multistability” renders the putative fictional status of a simulation ambiguous, depending not only on the way the simulation is interpreted, but also on the way we are able to produce new manipulations of laboratory objects. As such, following Hacking’s brand of instrumental realism, fictional models can facilitate real, non-fictional discoveries when we reinterpret them as something we can possibly build in the “wet” lab. It may be implausible to suggest that if we can virtually “spray” a fictional object, it is real; but we can certainly use fictional constructions to find new ways to “spray” real objects.

Chapter 5
The Functions of Scientific Fictions

The preceding chapters have argued for an emphasis on the technological character of scientific models and simulations, paying close attention to articulating a space for fictional simulational practices within the broader realm of scientific practice. That such a space exists, and that fictional simulational practices can make positive contributions to scientific research is clear enough. However, a clearer account of just what the aims of such practices are, as well as an articulation of how the fictional elements of significantly distorted models and simulations fulfill these aims, remains to be given.

There is an unavoidable vagueness inherent in the question of determining the aim of any class of technology, or even some particular technology. All technologies offer some sort of material mediation[51] between human agents and the world for the sake of accomplishing some goal, and it is this goal which guides the design and use of a technology as well as judgments about whether or not the technologies in question perform their functions adequately. But these goals are virtually limitless, and attempts to reduce technology to a single functional essence are notorious for their selective attention to convenient examples. Additionally, technologies are notorious for their

[51] A case can be made for dropping this "material" stipulation, one which focuses upon the "*technê*" (*viz.* practical instrumental know-how) (*c.f.* Kline 1985) or on networks of relationships between human and non-human "actors" (*c.f.* Latour 2007) rather than on the "*technê onta*" (*viz.* things whose essence is inextricably tied up in the practical projects of some other agent) of technologies. This approach would include organizational/political structures, language, and even ideas as potentially open to technological analyses. While I have no principled opposition to this tack, it widens the already wide scope of technology to a point that one may struggle to find any human activity that isn't technological. I make the admittedly arbitrary stipulation that technologies be material and deployed by some user toward some end for the sake of delimiting a more manageable extension for the term. Given previous arguments about the irreducibly material character of even the most abstract computer simulations, I am confident that little will be lost in the application of a material sense of technology to the topic of scientific modeling and simulation.

ability to be “hacked” and for the functions that they were designed to fulfill to be subverted in unintended applications (Ihde 2008). It seems possible that an effort to analyze the function of an entire class of technologies might be doomed to fail before it even begins.

Some effort towards specifying a function of scientific models and simulations has already been made. Broadly, they can be said to afford “epistemic access” to some intended target. While this function, too, is both vague as well as potentially inclusive of every technology, it has additionally been specified that models and simulations afford this epistemic access by serving as representations (or better yet, *representatives*) of their targets. And, lest we run the risk of losing sight of their (material) technological character and slipping into an exclusively representationalist idiom when addressing scientific models and simulations, a prop-theory of models was introduced. These specifications serve to narrow the scope of the potential functions of scientific models and simulations, both fictional and non-fictional varieties, to a more manageable range of possibilities.

Still, we should be wary of a tendency to be overly reductive when talking about the technological function(s) of a model or simulation. This is particularly prudent when it comes to fictional models and simulations, as they embody an especially unstable sort of representational relationship to their targets. The function of even some particular model/simulation may shift throughout the evolution of a research program. Accordingly, in this chapter I will discuss several functions which may overlap but are not necessarily consistent with one another: pragmatic functions, predictive functions, heuristic functions, and explanatory functions. I will discuss these functions as they are satisfied in fictional and non-fictional contexts, and I will indicate that these functions not only shift over time, but also represent a plurality of potentially mutually exclusive goals that effective scientific practice struggles to coordinate. These functions are not intended to be exhaustive, but they are intended to give a reasonably thorough depiction of the complex texture of the aims of simulational practice.

Pragmatic Functions

Perhaps the most obvious and least controversial function of modeling and simulation practice is a pragmatic one. Why would anyone ever choose to study a simulated target system rather than the real thing? More often than not, the answer is because access to the real thing poses some sort of practical difficulty. The universe is full of somewhere on the order of 10^{22} stars at a wide variety of stages of the stellar life cycle, but these objects are far away and there are limitations on the amount and sort of data we can collect from them. The size and energy scales of stellar phenomena place severe limitations on the kinds of experimental systems that can be created and studied. Additionally, real stellar systems evolve relatively slowly in comparison to the sorts of time scales required for human research, and certain phenomena, like the “birth” or “death,” of the universe are temporally inaccessible. In response to these pragmatic restrictions, we must find or construct surrogate systems (*e.g.* high energy particle accelerators, or computer simulations) which exhibit varying degrees of fit to their target.

Included amongst simulations that perform pragmatic functions, I would also count those that offer ways around moral and political restrictions. Creating close approximations of genuine stellar phenomena in the laboratory is not only practically difficult, but the associated risk of destruction of property and loss of life is morally problematic as well. We use crash-test-dummies when performing safety tests on automobiles because using human passengers raises legitimate moral objections. Additionally, while some sorts of research on human subjects are morally objectionable, and the use of surrogate organisms is apparently less objectionable,[52] both the moral and economic cost of using surrogate organisms that are more similar to human targets can motivate decisions to use cheaper, less similar alternatives.

Pragmatic difficulties can compound after the decision to simulate has been made. A recent series of articles published in *Slate* magazine by Daniel Engber notes how the evolution of model organism monocultures (mice are used in 59% of studies on animals) carries with it subsequent pragmatic practices such as *ad libitum* feeding practices (allowing mice to eat whenever they want) or lack of exercise that are pragmatically appealing, but introduce new distortions to a system that is already a surrogate system. Using mice instead of humans presents certain advantages, but also certain new issues concerning how those mice will be kept, fed, bred, etc. (Engber 2011). This was noted in the previous chapter as a characteristic of the complexity of interventions on material systems and how it is impossible to "subtract" without also "adding." It is worth mentioning again here to underscore a potential "snowball" effect that introducing distortions into a model/simulation brings with it.

To point out that this constitutes an "intrusion" of non-epistemic values into the practice of scientific research is fair enough, but even if we grant critiques of a distinction between "epistemic" and "non-epistemic" values (*c.f.* Longino 1990; 1996; Rooney

[52] There is, of course, ample room for critique of the assumption that research on surrogate organisms is morally different in kind than research on human subjects. My point here does not rely upon such assumptions.

1992) and arguments that so-called non-epistemic values can play an essentially epistemic role in science (*c.f.* Dewey 1929; Rudner 1953; Douglas 2000), the fact that simulational practices perform pragmatic functions doesn't offer a compelling argument for regarding the practical utility of simulations as an essentially epistemic function. Even when simulation is used in the service of affording "epistemic access," even casual examination of concrete cases suggests that simulation is only preferable to experiment or observation in the wild when these options are blocked by practical issues.

The recurring question for simulations is not whether they provide epistemic access, but what they provide epistemic access to, and insofar as simulational practice represents a departure from experimentation and observation, it merely permits access to something like the target phenomenon, not the target phenomenon itself. Absent some other more essentially epistemic functions of simulations, it would seem that, *ceteris paribus,* we should always prefer the real thing to a virtual stand-in. Strictly taken on the basis of their pragmatic utility, if simulations are acceptable stand-ins for their targets, it is despite the ways they deviate from those targets, not because of them. Simulation is merely epistemologically tolerated for the sake of other values, but there is little indication on the basis of the pragmatic function of simulations, that they exhibit any epistemological values of their own.

Following this line of argument, we can also see that there is nothing special about fictional simulation when it comes to the satisfaction of a pragmatic function. If epistemic access to a "real" target is blocked for practical reasons, it makes no difference whether simulated access is fictional or non-fictional provided that the simulation finds a way around the pragmatic barriers. If the fictional status of the simulation *does* make a difference, then so much the worse for fictional simulations. *Ceteris paribus*, we should prefer minor deviations to major ones when substituting simulated phenomena for their targets, but we may settle for a fiction if non-fictions come with practical difficulties. If a fictional simulation works, then this is despite its fictional character, not because of it.

Predictive Functions

The preceding section represents what I take to be a fairly common attitude toward simulation in general and fictional simulation in particular. In short, the decision to simulate (or more specifically to simulate with a fiction) is motivated and sanctioned by pragmatic concerns, but this satisfaction of pragmatic concerns is an epistemological liability that we may merely tolerate. There is no epistemological advantage to simulation when the real thing is available. Setting aside admittedly legitimate arguments regarding the distinction of epistemic values from non-epistemic values, it would seem that simulation offers no unique epistemic advantage because their pragmatic function is (debatably) non-epistemic. It is precisely this attitude that this chapter[53] aims to challenge. Instead, I would argue that the ways in which simulations deviate from their targets, even, and perhaps especially when those deviations are severe enough to qualify a simulation as fictional, offer advantages that are decisively epistemic. Beginning this challenge is easy enough, for it is fairly uncontroversial to point out that simulations do not merely satisfy pragmatic functions. In particular, most philosophers who have turned their attention to scientific simulation gravitate toward one particular and apparently unambiguously epistemic function: namely prediction.

[53] perhaps more accurately, this dissertation as a whole

To call prediction an unambiguously epistemic function, meaning that it is a source of novel knowledge, is plausible but not wholly uncontentious. The waters can be muddied somewhat if we regard simulations as nothing more than extensions of theory. If a simulation is simply a concrete instantiation of a world where a theory is true, and built entirely from theoretical principles, then the predictions they yield are merely analytic extrapolations from that which is already known. The challenge to regarding the predictions of simulations as new information rests on the *"merely analytic"* aspect of this claim. It would hardly constitute an acquisition of new knowledge to proceed from knowing that "John is an unmarried male" to knowing that "John is a bachelor." At best this provides new knowledge about the identity of the concepts "bachelor" and "unmarried male," but indicates no genuine new knowledge about "John." Computer simulations, since they use analytic computational algorithms to produce predictions, are particularly susceptible to this critique. A stronger sense of an epistemic function seems to demand the production of synthetic knowledge.

Grim *et. al.* (2011), offer a radically simplified general structure for simulations (building off of Eason *et. al.* 2007) that clarifies this issue a bit.[54] This structure identifies three parts to any simulation: *input*, *mechanism*, and *output*. The purpose (or function) of a simulation is associated with which of these three structural parts is the locus of "new information." For example, a simulation designed to predict whether it will snow tomorrow may take known meteorological data from the past few days and use this information as input for a mechanism based on previously worked out and understood law-like or principles to produce an output predicting the next day's weather.

[54] The virtue of this radical simplification is that it draws a clear relationship between the structure of simulations and the intentional goals of the simulator. See Winsberg (2010, 10-12) for a more complicated structure (that he admits is still a radically simplified version of actual simulational practices).

The locus of "new information," or what is unknown prior to running the simulation (in this case the "output") is indicative of the simulation's purpose (in this case, prediction).[55]

That the output of a predictive simulation is initially unknown and represents an instance of new information is neither trivial nor is it intended in a merely figurative sense. The mere fact that a simulation may employ an analytic computational mechanism to convert known input into "novel" output does not necessarily confer the vacuity associated with deductive inferences that are "merely analytic." What is of primary concern in the question of whether predictive simulations satisfy a genuinely epistemic function is whether they allow us to proceed from what is known to what was previously unknown, not the nature of the inference this movement employs.

The fact is that most simulations instantiate causal mechanisms that are complex enough that even analytic relationships between input and output are not readily evident to the human mind. Any technology that makes inferences that we cannot make by ourselves, allowing us to know things that we otherwise wouldn't, qualifies as fulfilling an epistemic function by any but the most nit-picky definitions. Particularly obstinate nit-pickers might point out that simulation isn't necessary for all such complex inferences, and that they can be made with pencil and paper.[56] But, oftentimes, simulational methods are used precisely where such standard analytic methods fail. Many philosophers (*c.f.* Winsberg 2003, Humphreys 1990, 2004) have focused on just

[55] The other purposes discussed by Grim et. al. are discussed later in this chapter.
[56] I am not sure what difference it makes that the work done by a simulation could, in principle, be done by other means, but I am hesitant to concede that working things out manually rather than through digital computation somehow takes us out of the realm of simulation. It strikes me as plausible to suggest that one who performs their analytic inferences with pencil and paper is simply simulating on paper, rather than in a computer.

such uses of simulation to deal with analytically intractable mathematical models. Even relatively "simple" models, like those represented by the classical three-body problem, pose significant difficulties for ordinary analytic solutions.

While the preceding makes a convincing case for regarding predictive simulations as fulfilling an epistemic (as opposed to pragmatic) function, it remains unclear whether or how fictional simulation offers some particular advantage in making predictions. In the tripartite structural model described by Grim *et. al.* a predictive simulation is one where the input and mechanism are already known. Therefore there appears to be no compelling reason for why one would deviate from a faithful representation of the target system, as distorting the input or mechanism would invite the risk of distorted output. In fact, Grim *et. al.* go so far as to say that significant failures of correspondence between the simulation and its target at any of its three moments constitute grounds for regarding the simulation as a failure.

But considerations of success and failure are symptomatic of applying "simulation" as a success term, and ignore the fact that the vast majority of scientific simulations deal with situations where the available input data is limited and/or potentially erroneous, and the mechanism is not necessarily well understood. If reasonably accurate prediction is our goal, then overcoming uncertainties in the input and mechanism may demand that elements be added to a simulation that deviate from our best empirical and theoretical hypotheses about inputs and mechanisms for the sake of making reliable predictions. This is certainly the case for the prevailing approach to the construction of models and simulation in climate science, where the systems under investigation are so complex, theory so incomplete, and data so squishy, that modelers are unable to reliably attribute their successes and failures in prediction to any particular features of their models. Consequently, their strategy, insofar as they are primarily concerned with the predictive function of their simulations, is to "kludge" and "parameter tweak" their way to more successful predictions.[57] These adjustments are made without any expectation that they represent the input or mechanism of the target phenomena accurately, and in fact it is difficult to say with any reasonable amount of certainty what (if any) features of the target phenomena they correspond to.

This indicates the existence of legitimate cases where fictional elements are introduced into simulations in response to uncertain empirical and/or theoretical knowledge. In such cases, fictional simulation offers an advantage, rather than a liability, for the satisfaction of predictive goals. Additionally, fictional distortions may be introduced for the sake of reliable prediction in a direct response to distortions independently introduced to satisfy pragmatic goals. Previously we examined the case of the inclusion of fictional "silogen" atoms that enabled "handshaking" algorithms to communicate between regions in a simulation of crack propagation in a silicon brick (Winsberg 2009). Recall that the reason that handshaking algorithms were necessary in the first place was that it was too computationally expensive to simulate the entire system with the quantum mechanical model that adequately characterized the target system at tip of the crack. The need to simulate the target system in three regions was driven by a pragmatic limitation, but this, in turn, lead to a secondary problem of how to adjust the simulation in light of this distortion to preserve the predictive function. The use of "silogens" was an instance of "getting it wrong locally in order to get it right globally."

[57] This is a shameful oversimplification of the incredible sophistication and complexity with which climate models are adjusted and evaluated, but it serves for the purposes of this discussion. For a more thorough analysis, see Winsberg 2010, 93-119.

Similarly, in the use of early versions of computational simulation by Von Neuman in the development of the atomic bomb, a project regarded by many as the first significant effort at scientific computer simulation (Galison 1996; Fox-Keller 2003; Winsberg 2010), the pragmatic concerns that initiated the effort to computationally simulate fission reactions gave rise to the need for a fictional solution in order to maintain predictive adequacy in the form of Monte Carlo methods. Simulating the genuine randomness of nuclear phenomena in a computational context demanded a pseudo-random mechanism that was locally wrong but globally adequate. What began as a pragmatic function, a means to the end of solving an intractable mathematical problem, became something more complex and more vivid in its virtuality. As Peter Galison remarks;

> [P]hysicists and engineers soon elevated the Monte Carlo above the lowly status of a mere numerical calculation scheme, it came to constitute an alternative reality – in some cases a preferred one – on which 'experimentation' could be conducted. Proven on what at the time was the most complex problem ever undertaken in the history of science – the design of the first hydrogen bomb – the Monte Carlo ushered physics into a place paradoxically dislocated from the traditional reality that borrowed from both experimental and theoretical domains, bound these borrowings together, and used the resulting bricolage to create a marginalized netherland that was at once nowhere and everywhere on the usual methodological map. (Galison 1996, 119-20)

We may well treat Galison's evaluation of the virtual vividness of Monte Carlo simulations with some skeptical caution. Following Evelyn Fox-Keller (2003) I would maintain that while Galison's intuition about the metaphysical and epistemological novelty of simulations seems to be correct, his example is a weak one. The alternative reality constructed by the punch-card computers used by the Manahattan Project was a rather thin one, and we might be justifiably suspicious of Galison's hyperbolic attribution of Monte Carlo games as an ‘alternative reality’ for the scientists, as well as the scare quotes he places around ‘experimentation.’ The computers themselves did not instantiate anything remotely resembling the ‘alternate reality’ of the Monte Carlo game, and consequently we may have serious doubts about the extent to which this simulation resembled experimental practice and all the embodied interactions that are associated with it. Add to this the fact that these punch-card computers were not being operated by the scientists themselves (Schlombs 2010), and there is sufficient distance from these computer ‘experiments’ to cast serious doubts on the extent to which their ‘alternative reality’ was real for Von Neuman *et. al.*

Since these early versions of computational simulation, however, the development of a richer computer interface has developed significantly with more sophisticated visual displays as well as the regular incorporation of interaction with such displays into the everyday lives of their users. We can take the phrase ‘alternative reality’ a bit more seriously when attributed to a computer simulation now then we could in the 1940’s. Accordingly, the verisimilitude of scientific computer simulations today represents the realization of Galison’s hyperbolic interpretation of early Monte Carlo methods.

Several things are evident from the preceding remarks. First, prediction, even when reduced to the deduction of analytic consequences of theoretical principles and empirical observation, is a praxis, and accordingly carries with it the sorts of pragmatic complications that are conflatable with so-called “non-epistemic” pragmatic functions of simulations. When equations become analytically intractable in principle, or even simply prohibitively difficult, simulation provides scientists with a means to an end, and sometimes this end is making predictions. Because they are part of a complex material praxis, simulations do not necessarily function transparently as means to these ends. A pragmatic solution to a practical problem introduces new practical problems that demand new pragmatic solutions, and, once again, innocuous deviations from a relatively faithful simulation quickly compound into more serious ones.

Second, despite the fact that the praxical character of prediction allows for the potential conflation of the predictive function and non-epistemic pragmatic functions of simulations, the respective goals of these functions are potentially conflicting. This was already hinted at when I remarked that simulating for pragmatic reasons inevitably introduces deviations from the target system, deviations which, *ceteris paribus*, we would prefer to minimize. These deviations may manifest themselves in undesirable ways by corrupting the reliability of the predictions made by a simulation. But interestingly, this problem is frequently solved not by correcting these initial deviations, but by introducing new ones. Thus we begin to see how the introduction of fictions into simulational systems is not merely a liability, but an advantage. Fictions allow the practitioners of simulational methods to partially coordinate a multiplicity of conflicting goals.

Third, there exists a risk that simulations may be reduced to an exclusively technological function if held only to the fulfillment of pragmatic and predictive goals. If Fox-Keller is correct that early Monte Carlo methods did not really succeed in creating a virtual world, then the extent to which they constitute a *simulation* of the randomness of nuclear reactions is suspicious. At the very least it suggests a dual sense of simulation, where one sense indicates an alternative means to a functionalist end (e.g. solving difficult equations), and the other holds simulations responsible for a deeper representational relationship to their intended targets. I have argued, in previous chapters, for a two-stage "representation-as" structure for models and simulations, but a "simulation" that is used exclusively or even primarily as an analytic tool exhibits this representational character weakly or not at all. If these tools are "simulations," then they are so only in a sense that is derivative of and weaker than those simulations that facilitate the experience of the simulation as a virtual version of the target.

The same risk is posed by simulations that "tweak" and "kludge" their way to reliable prediction. If their mechanisms become sufficiently opaque and "black-boxed," then they essentially become prediction-making machines. To say that a prediction-making machine does not serve an epistemic function would seem unfounded for reasons already discussed; but there does seem to be something rather unscientific about a "simulation" whose "black-boxed" mechanisms we have no intention of ever opening. Whether examples like climate modeling practices represent these sorts of permanently "black-boxed" mechanisms is unclear. There does, at the very least, appear to be some anxiety regarding the attributive impenatrability of climate models, and an interest in opening up their "black-boxes," even if the prospect of success looks rather bleak.

Compare this to the language parsing software described in chapter two, which is a much clearer example of a "simulation" that has been fully reduced to a technology, and no longer simulates in any strong or scientific sense. The mere fact that the software parses language in a very different way than natural language users does is not what makes it a poor example of a simulation, nor does the fact that we do not care that the software operates differently than the mind of a natural language user (for this is characteristic of fictional simulations). What makes this "simulation" different is the fact that it has no discernable representational relationship whatsoever to the natural language use it ostensibly “simulates.” It would be difficult to learn anything about natural langue use by studying this software, but this is not its purpose. Its purpose is entirely instrumental— to get a computer to perform a useful task.[58] Consequently, there is little prospect that its “fictional” deviations from what it “simulates” will ever bear the hallmark of a scientific fiction: that they will eventually be jettisoned for a more accurate, non-fictional representation of natural language use.

It is notable that the line between what I would call genuinely scientific simulations and those which have been reduced entirely to a technological function is a fine one and subject to revision over time. Furthermore, in keeping with a preference for focusing on the ’promising’ rather than the ’successful,’ the question of which side of this line a particular simulation falls upon depends largely on the intentions of the designers and users, since it is difficult to say in advance whether the fictional elements of a “promising” simulation can ultimately be phased out. What is drawn to our attention here is the difference between whether we don’t care *at all*, or whether we don’t care *for now* when a “simulation” performs a function similar to some target but does so in a markedly different way

A Breif Discussion about “Multiple Models”

In chapter two, I discussed two kinds of idealization (“Galilean” and “minimalist”), taking many cues from Michael Weisberg’s essay, *Three Kinds of*

[58] It should also be noted that in the case of language parsing software, the utility of the “predictions” made is so weak as to raise a serious question about whether a grammatically parsed sentence produced by the “simulation” is actually intended as a *prediction* of how natural language users would parse the sentence. That it functionalistically mimics the parsing of natural language users is plausible enough, but if we had genuine questions about how a given sentence would be parsed, we would more than likely simply ask any native speaker rather than turning to the simulation. It seems more accurate to say that while such a technology is a stand-in for a natural language user, it offers no epistemic access to what it is standing-in for.

Idealization (Weisberg 2006). I skipped over a discussion of the third kind, namely those employed in aggregations composed of multiple models, because it did not particularly fit the discussion at the time. I believe the context of the discussion in this chapter fits the topic better.

there are at least three interesting aspects of multiple model scenarios. 1) Whether the employment of multiple, incompatible models automatically implies the use of fictions. 2) Whether the attempt to reconcile the mutual incompatibility of multiple models presents a problem that can be solved by the use of fictions. And 3) whether the use of multiple models to represent target phenomena is limited to specific sorts of scientific aims.

The first has to do with the question of what the use of multiple distinct models says about the fictional/non-fictional status of those models. We may be tempted to conclude that, given the qualitative differences between several models, they can't all be "correct," and thus they cannot all be non-fictional. If we take an example from Winsberg, the total silicon brick system is modeled with a quantum mechanical region, a molecular dynamic region, a classical continuum mechanical region, and hand-shaking regions at the boundaries between the others. Without even beginning to adjudicate which model is the best, or most correct, it seems to be the case that the discrepancies between the various models indicates that they cannot all be equally correct, and thus that there are some fictions to be sniffed out.

However, we may resist this temptation to conclude that any model which is, itself, comprised of multiple distinct or even mutually exclusive models, is automatically fictitious. In the case of the silicon brick modeled in three regions with three distinct models, each model is adequate so long as it is not extended beyond its appropriate context. Winsberg, himself, takes this position, stressing that each region, on its own, is non-fictional despite their mutual incompatibility. While he offers very little justification for this position, it is compatible with a perspectival realism, such as the one described earlier and attributed to Giere. From a perspective that is concerned with the meso- and macro-scale properties of a silicon brick, a classical model works just fine. It is as good of a scientific description of the brick as quantum model. From a perspective that is concerned with the ultra-microscale properties of the brick at the point where bonds are breaking, the classical model is no longer adequate, but this does not discount its adequacy in other contexts. To do so would be to privilege the ultra-microscale perspective over others, and there seems to be little basis for this privileging.

I must say that I am hesitant to commit to this position altogether. One need not privilege one perspective over another to say that the quantum model is to be preferred as "better" than the classical model, for it works equally well in all contexts. This is debatable if we include computational expense as part of our rubric for how well a model works, and, indeed, it is precisely this disadvantage that precludes scientists in Winsberg's example from modeling the entire system quantum mechanically. Furthermore, it is not my intention to argue for a broad, reductionistic position by pointing out that in this particular model, we might privilege the quantum mechanical depiction by virtue of its adequacy for all the regions of interest in the target phenomenon. The question of whether a perspectivalist appeal represents an unassailable answer to the suggestion that the use of multiple, incompatible models automatically indicates the presence of fictions is a difficult one and demands further attention, but I believe it can be bracketed within the scope of this project.

Regardless of whether we are entitled to claim that multiple incompatible models can, in principle, all be non-fictional, there is a second point that can be addressed. While the three models corresponding to the three regions of the silicon brick model could be plausibly regarded as non-fictional, the "silogen" atoms in the handshaking region between the quantum mechanical and molecular dynamical regions are clearly fictional. In the attempt to combine multiple incompatible models into a single model, the inclusion of uncontroversially fictional elements may need to be included. Thus, the perspectivalist appeal does not get us entirely clear of the possibility that multiple model scenarios are of particular interest for the philosopher who is interested in fictional modeling. In fact, it seems to steer directly into a distinctive role for scientific fictions. A given model may be fictional in the wide or even the narrow sense in order to exemplify a particular feature of the target phenomenon and render it epistemically accessible, but there does not appear to be any principled reason why a fictional representation could not eventually be dispensed with or give way to a non-fictional one as the phenomenon becomes better understood. When we are attempting to combine multiple incompatible models in order to represent phenomena that span multiple perspectives, fictions play a slightly different role, and may be significantly more intransigent.

This introduces a third point, regarding the construction and use of multiple models as a distinctive style of scientific practice. (See Weisberg 2006 and Godfrey-Smith 2005) In many fields, particularly those dealing with very complex phenomena, scientists work with collections of models without any expectation that a single best model will be generated. One notable example of this practice can be found in meteorological and climatological modeling. Not only are these models tuned with an eye toward predictive accuracy in a way that renders them attributively opaque (that is to say that scientists become unable to determine how what, if any, features of the target correspond to key features of the model), but predictions are generated from collections of independently constructed models. In the face of unsuccessful attempts to generate a single ideal model, The U. S. National Weather Service has determined that the best route to reliable predictions is to employ three primary models (as well as many more secondary ones) (USNWS). Other notable examples can be found in economics (*e.g.* Li and Dorfman 1996; Chipman, George and McCulloch 2001; Norgard 2004) and population biology (*c.f.* Levins 1966) Whereas Winsberg's "silogen" example shows the utility of fictional elements in knitting together a single model from multiple distinct ones, these scenarios are not looking to smooth out the incompatibilities between the multiple models they employ. Nonetheless, the strategy they employ lends itself well to fictional modeling, as the use of multiple models means that particular models can be tuned for reliability with respect to certain features while ignoring concomitant unreliability with respect to others. Thus, robustly reliable meteorological predictions can be generated by a collection of models, even though none of the particular models are themselves robustly realistic. One model may be reliable in predicting precipitation, but unreliable in predicting temperature while another may give reliable predictions of temperature and unreliable predictions of precipitation. Neither of them are realistic models of the weather, nor can we say that a collection of such models is realistic, though they are, as a group, predictively accurate. So long as we are only concerned with their collective predictive accuracy, the particular models can be developed without regard to their global accuracy, drifting progressively further from any realistic depiction of the global target.

The preceding addresses the use of multiple models for predictive aims where individual models are only expected to be predictively reliable for a subset of the features of a complex, global phenomenon. The same principle applies for the use of multiple models to satisfy a variety of potentially incompatible epistemic functions. In chapter five, I claim that fictions may be employed to adjudicate between pragmatic, predictive, heuristic, and explanatory functions. To develop a model/simulation that satisfies all (or even more than one) of these functions, it is frequently necessary to fictionalize that model. Alternatively, we may permit a model to unrealistically focus on one of these functions to the detriment of its ability to satisfy the others, taking a multiple models approach to satisfying them all. As Richard Levins argues,

> The multiplicity of models is imposed by the contradictory demands of a complex, heterogeneous nature and a mind that can only cope with few variables at a time; by the contradictory desiderata of generality, realism, and precision; by the need to understand and also to control; even by the opposing estetic standards which emphasize the stark simplicity and power of a general theorem as against the richness and the diversity of living nature. These conflicts are irreconcilable. Therefore, the alternative approaches even of contending schools are part of a larger mixed strategy. But the conflict is about method, not nature, for the individual models, while they are essential for understanding reality, should not be confused with that reality itself. (Levins 1966, 431)

The takeaway here is that regardless of whether we are pursuing a single model strategy that attempts to accommodate multiple contexts, perspectives, or goals, or we are instead taking a multiple models approach to such an accommodation, there is a tendency to fictionalize our models.

Heuristic Functions

The preceding account of the predictive function of simulation establishes a non-controversially epistemic function for simulations, one in which the prospective advantage of simulating fictionally is evident. However, simulations that are only responsive to pragmatic and predictive goals seem to paint a rather thin picture of the role of simulation in scientific practice. If this is all that there is to say about simulation, then it is easy to sympathize with those who express anxieties that simulations cannot yield genuine *discoveries* about the world or challenges to pre-existing theories. A depiction of simulations as sites of genuine scientific inquiry and discovery demands that they enable us to do more than make predictions, but also achieve some understanding of the real mechanisms by which the predicted states are produced in target phenomena. This goal is fulfilled by simulations that have a heuristic function.

The term, 'heuristic' is sometimes meant to denote a rough and ready, or "fast and frugal" (Gigerenzer and Todd 2000) rule of thumb or algorithmic strategy for reaching some desired goal. Mental heuristics like "when you hear hoofbeats, think horses" are sanctioned by their tendency to produce reliable and effective action (despite the knowledge that other non-horse animals may also make hoofbeats). Used in this sense, a heuristic function for simulations can easily be assimilated into the pragmatic and predictive functions already discussed.

However, I use "heuristic" here in a slightly more precise sense, one with a closer link to its etymological derivation from the Greek, *euriskein*, meaning to "discover" or "find." While arguing that predictions are not "discoveries" is about as pedantic and silly as arguing that they do not constitute "novel knowledge," there does seem to be a very real sense in which a project that aims only at making a specific sort of prediction may foreclose the possibility of making discoveries. This much is indicated by the attributive impenetrability of complex simulations that have been tweaked and kludged for the sake of predictive reliability. In order for a simulation to facilitate an open ended pursuit of discoveries that influence the very theoretical principles they are (partially) constructed from, the functional significance of its own parametric and structural features, as well as the correspondence between these features and those of the target system, must be made intelligible.

In chapter four, I made a provisional claim that simulational practices make up for their issues with correspondence to their targets by being more intelligible due to the fact that they are built up from scratch. That this is not necessarily the case should be apparent by now in light of the way that predictive functions are potentially in tension with the sort of intelligibility that a heuristic function requires. In order to make a simulation predictively accurate, it may be necessary to make distortions so complex that the simulation becomes attributively opaque. As an additional gesture toward confusing a provisional distinction previously made, I concede that adjudicating whether a particular moment in a scientific research program is experimental or simulational is more difficult than I have suggested. "Data modeling" practices, since they concern the manipulation of empirically collected data, would ostensibly be categorized as "experimental" rather than "simulational." Yet, the style and extent to which this data is manipulated can compromise the ease with which that data can be causally traced back to the phenomenon itself, thus compromising the very thing that distinguishes experimental practices from simulational ones. This hybrid/intermediate territory is exemplified by the aptly named (in light of the focus of this section) Eureqa system. (CCML)

[Eureqa] works by stringing together simple mathematical expressions to create large banks of equations. Each equation is tested to see how well it fits experimental data. The majority of these equations are sheer nonsense, but by chance some fit the data a little better than others. In an analogue of sexual reproduction, the software saves these equations for 'breeding', combining one half of a 'father' equation with one half of a 'mother' equation. Sometimes, it alters a term in the equation, to mimic random genetic mutation. Over thousands of generations, equations emerge that fit the data quite well (DaSilva 2011).

While Eureqa is ambiguously simulational and experimental, it is a fairly clear example of the fulfillment of what I am calling a heuristic function. The genetic algorithm it uses does not simply produce a function that matches the data, allowing us to make predictions of future data, it identifies unacknowledged relationships between variables that are suggestive of the potential salience of certain variables over others. Of course, the equations produced must be interpreted. Furthermore, they only represent suggestions and there are numerous equations that may fit the data with relative elegance and simplicity, so a program like Eureqa functions as a tool that, at best, guides subsequent experimental inquiry in light of past data.

What Eureqa does is not so different in kind from standard curve fitting techniques that have long been part of experimental practices of data interpretation. In this respect, Eureqa is a poor example for examining how simulations work. However, it is an excellent example for framing a discussion of heuristic functions as facilitating discoveries. The equations produced by Eureqa and other similar systems provide a relatively tidy and simplified version of something messy and complex. They are mathematical models that hone in on certain features of a system as potentially significant, and others as insignificant, and they guide subsequent empirical research.

In this respect, simulations that are simplified versions of their targets are well positioned to play a heuristic role in scientific inquiry. Such is the case for many "model organism" systems, which are chosen not just for the non-epistemic pragmatic advantages they offer in relation to direct study of more complex target systems, but also because they are simpler and more easily standardized and manipulated than their targets. The expectation is that this relative simplicity facilitates the discovery of salient features of both the model and target system. Jane Hubbard (2007) describes a particular instance of this sort of discovery through the study of *C. elegans* earthworms, specifically worms with an abnormal cell lineage called *lin-12* that is associated with abnormal vulva development and the inability to lay eggs properly. Despite the unlikelihood of application to reproductive disorders in more complex organisms, the radical simplicity of *C. elegans* allowed for the study of the genetic mechanism by which this abnormal cell lineage is elevated or reduced as well as its impact on cell-cell communication and binary cell fate "decisions." However, it was noticed that T lymphoblastic leukemia disrupts a certain gene in humans that is similar to the *C. elegans lin-12* gene (as well as to the *Notch* gene in *Drosophila* fruit flies). Researchers were able to engineer forms of worm and fly genes that were even more similar to the disrupted human gene to reveal the mechanism by which the leukemia-associated mutation increased the activity of cell-surface receptor proteins in worms, fly, and human cells alike. Quoting Hubbard, "This result would have been very difficult to obtain using human tissue culture or other means. Since then, three additional *lin-12/Notch*-like genes have been found in humans, two of which have also been implicated in cancer. *Lin-12* mutations do not cause cancer in worms, and yet they are providing insights into the mechanisms of human cancer." Subsequent research has let to the reliable manipulation of *lin-12/Notch*-like genes in mammals, as well as the identification of additional, distinct gene groups implicated in vulval development in *C. elegans* that also have analogs in human forms of cancer (Hubbard 2007, 65-67).

I have argued previously that the use of model organisms has a compelling simulational character (despite their tendency to be associated with straightforward experimental practices). Yet, the extent to which the discovery of the mechanism of T lymphoblastic leukemia's disruption of the proper functioning of cell-surface receptor proteins constitutes an example of a discovery made in a fictional simulation is a bit uncertain. The stark differences between worms, flies, and humans constitute precisely the sorts of non-de-idealizable distortions that are a hallmark of the narrow conception of fictions I have articulated in chapter three; but over time, the evolution of our understanding of *lin-12* mutations from signifying reproductive abnormalities in worms to signifying abnormal cell-surface receptor activity reframed the interpretation of the gene to allow for a far more straightforward analogy with human biology.

We saw the same sort of re-interpretive evolution in the fluid-dynamics case study examined in chapter four. Karniadakis and Triantyfallou's 2-D simulation was first understood as an idealized version of the experimental apparatus that could be approached (even if not reached). Then it was understood as an impossible and unapproachable object, namely a cylinder with no ends. Finally it was understood as a cylinder with negligible end-effects, which was ultimately instantiated in the laboratory by angling the cylinder end-caps. On the basis of imaginative re-interpretation, the simulation went from fiction in only the wide sense, to fiction in the narrow sense, to non-fiction. This is typical of the way that fictional simulation fulfills a heuristic function. In order for a fiction to be scientific, it must exhibit some practical utility, serving as a means to an end. If that end is the fulfillment of a heuristic function, that is to say facilitating discoveries about the real target phenomenon, the fiction must point the way to a non-fictional understanding of what was initially represented fictionally. A transition from the fictional to the non-fictional must occur, and it is in this heuristic context that we can see how a fiction can facilitate its own disappearance by facilitating a discovery.[59]

In the discussion of the predictive functions of simulations, we saw how the practice of making predictions opens up a space where the virtues of fictionalizing distortions becomes clear. Making discoveries is also a practice, one which is facilitated by heuristic models and simulations that are relatively simple and standardizable/ reproduceable, thus opening up a space for potentially fictionalizing distortions. But in order for these discoveries to be genuine discoveries about simulational targets, the fulfillment of a heuristic function must not only involve the inclusion of fictional distortions, but their removal as well. In large part, the de-fictionalization of a fictional model is affected hermeneutically by imaginative re-interpretation of what the features of the model signify.

The fact that this also involves an intimate relationship with experimental manipulation of the target phenomenon cannot be overlooked or underappreciated. Simulations that fulfill a heuristic function make predictions, but this function is not

---

[59] See chapter three of this dissertation, pp 83-85.

simply reducible to a predictive function. They do not simply fill in a blank, they reveal new blanks to be filled in, and they do this by revealing unanticipated ways of manipulating and observing their target phenomena. Hence the discoveries that they facilitate are not complete until the scientific community transitions from the reliable manipulation of the simulation to the reliable manipulation of the target. While this may seem to indicate that discovery occurs in experiment, rather than simulation, discovery is a process.[60] It is an event that is spread out over time and cannot be pinpointed to a specific experimental (or simulational) result. It may begin in simulation (even fictional simulation) and conclude in experimentation.

The consonance here with an instrumental realism of the sort propounded by Ian Hacking (1983, discussed in chapter one) should be fairly evident. Simulations that aim primarily at prediction are susceptible to being reduced to their instrumental utility, as prediction making machines or functional substitutes for their "targets." We can regard such systems as "real," but only if we take them on their own terms, as fully autonomous objects. To do so severs any intention of robust correspondence between the "simulation" and target, and raises serious questions about the extent to which these objects "model" or "simulate" a real target outside of themselves at all. On the other hand, so long as an intended robust correspondence remains, that is to say so long as there is an expectation that the structural and causal "joints" of the simulation also indicate "joints" where the target can be cut as well, there is an additional heuristic dimension to the simulation.

By speaking of "cutting nature at its joints" it is not my intent to take a metaphysical position about "natural kinds" or invite arguments about the existence of

[60] This is well discussed by Kuhn (1962)

objective essences. The “joints” of real systems, whether in the wild or in the laboratory, need only be understood as those features by which those systems can be reliably manipulated. They are determined by what we can *do* in the world (*c.f.* Hacking 1991). When a simulation reveals these “joints,” it goes from merely supporting discoveries about itself and it’s own virtual world, to affecting discoveries about the real world.

It is noteworthy that a lot of weight is placed on the intention or expectation of robust correspondence between the structure and mechanism of a simulation and its target. This echoes a point made about predictive functions for simulations, that the line between genuinely predictive and merely instrumental “simulations” is thin, vague, and dependent on the intentions of the designers and users of simulational systems, particularly when we focus on the domain of promising, rather than already successful simulations. Simulations that are promising with respect to a heuristic function may only possess a loose, and even fictional correspondence to their targets, but the possibility of articulating a clearer and non-fictional relationship must be kept alive. The realization of this possibility can be partially affected by imaginative reinterpretation of the simulation, and the suggestion of potentially salient and previously overlooked features of the target phenomenon, but without the prospect of exploring these potentially salient features in the target, and reliably manipulating them, the promise of a simulational discovery remains unfulfilled.

This ‘heuristic’ function, through which a model or simulation facilitates discoveries of novel, salient features of the target phenomenon, is thus framed equally well as a hermeneutic function. In the discovery of features, what is at stake is the interpretation of phenomena, and the significance of the parts of both the model and target in relation to the whole. Furthermore, the shifting of the relative priority and autonomy of simulation and experimental apparatus (such as was seen in the case study in Ch 4) is characteristic of what is frequently called the "hermeneutic circle." Both model and target are successively interpreted and re-interpreted in terms of one another, such that neither has an absolute, foundational priority over the other. Instead, there is a cyclical interpretive and material dialectic in which an evolving understanding of the model and target co-inform and co-constitute one another.[61]

Explanatory Functions

As an illustration of a scenario where the complete fulfillment of a heuristic function is blocked, consider simulational studies of quantum phenomena that employ "hidden variables" in order to model quantum systems with classical deterministic mechanisms. "Hidden variable" interpretations of quantum mechanics go back to the earliest days of quantum theory and Albert Einstein's insistence that "God does not play dice with the universe." Along with Boris Podolsky and Nathan Rosen, Einstein argued that the apparent indeterminacy of quantum phenomena must be a symptom of an incomplete theory in order to avoid the paradoxical conclusion that the "spooky action at a distance" between entangled particles would require faster-than-light communication of information (Einstein, Podolsky, and Rosen 1935). The hypothesis that this demanded a "local" hidden variable for a complete quantum theory was argued by John Bell (1964) to be testable experimentally, and subsequent experimental evidence has been taken by the

[61] See Crease (1993, 113-120) for an account of the hermeneutic circle at work in experimentation. This basic structure is also at work in the case of fictional models/simulations that manage to converge with experimental practices through successive manipulation and reinterpretation, though the gap between the two is much more pronounced. What the fictional context adds to the material hermeneutics of experimental practice is an element of play and make-believe. This is taken up in chapter six.

vast majority of physicists to indicate that no such local hidden variables exist. However, some physicists have maintained that this does not rule out "non-local" hidden variables of the sort described by Louis deBroglie's "pilot wave" theory or David Bohm's causal interpretation of quantum mechanics (Bohm 1957).

Measurement scenarios of quantum phenomena determined by non-local "hidden" variables have been simulated computationally (Brassard, Cleve, and Tapp 1999; Steiner 1999; Dakic *et. al.* 2008), specifying the number and character of hidden variable states required to model a given quantum state of affairs. But absent an available experimental test, these deterministic quasi-classical models and probabilistic quantum-mechanical models are empirically indistinguishable, which is also to say that these simulations do not currently point toward any novel ways of manipulating and observing real (as opposed to simulated) quantum phenomena. So long as "hidden variables" remain beyond the reach of what we can experimentally influence and observe, it is difficult to say that their simulation counts as part of their discovery. Such efforts amount only to a demonstration that a hidden variable theory is not ruled out.[62]

The modal qualification of these sorts of simulations, that empirically observable features of quantum phenomena *could possibly* be a result of a deterministic mechanism by way of some set of non-local hidden variables specified in simulation, is characteristic of simulations that fulfill an explanatory function. Grim *et. al*. (forthcoming) identify this as a consequence of the structure of explanatory simulations, which they define as

[62] What *is* made possible by these types of simulations is attempts to model quantum phenomena classically (say in the classically behaving circuitry of an ordinary computer), though this sort of project (short of some future breakthrough) is a thoroughly instrumental one. As in the case of black-boxed climate predicting machines, or language parsing softwares, it is not only the case that we do not care whether the targets of these simulations act according to the same mechanism as the simulation, we also have no foreseeable intention of ever articulating the relationship between simulation and target so long as it yields reliable, useable output.

providing new information at the mechanism of a simulation's tripartite structure. Recall that a predictive simulation provides new information at the output stage of this tripartite structure. If the input and mechanism are known with certainty (and the simulation produces output *via* a deterministic mechanism that translates input to output), then the output can be regarded with the apodictic certainty of any sound deductive inference. If, on the other hand, the predictions made by the output of a simulation fail to square with what actually obtains, then this counts as sufficient evidence that something is wrong with the input or mechanism of the simulation. The testability of simulational predictions against empirical observation derives from the fact that the output of a simulation is a necessary consequence of its input and mechanism. An explanatory simulation, however, if taken to denote a simulation where the input and output are known to agree with the target but the mechanism is not, merely establishes the simulational mechanism as *sufficient* for producing the output from the input. It suggests one possible way that the input conditions could lead to the output conditions, but there may be others and there is no guarantee that the target phenomenon follows the same causal mechanism as the simulation.

Grim *et. al*. also point out that this same modal qualification exists for "retrodictive" simulations, i.e. simulations where the input is the locus of the new information. Showing that certain hypothetical input conditions, say, some particular mass and angle of impact for the Chicxulub asteroid thought to have caused the extinction event that killed off the dinosaurs, would follow known causal mechanisms to produce output conditions that match our current observations, is only to demonstrate the sufficiency of these input conditions. Thus it is possible that past events were as a simulation retro-dicts, but only possible. That "retro-dictive" and "explanatory" simulations share this trait despite the structural difference articulated in Grim *et. al.* is unsurprising if we consider that both can be seen as offering explanations in a broader sense. For instance, under a "deductive nomological" (DN) model of scientific explanation, an explanans is comprised of a collection of sentences, some of which are empirical statements and at least one of which has a law-like structure, and taken together these sentences constitute a sound argument for the necessity of the explanandum (*c.f.* Hempel and Oppenheim 1948). Grim *et. al.*'s tripartite structure can be mapped onto the DN structure of explanation such that the "input" corresponds to empirical claims, the "mechanism" corresponds to law-like statements, and the "output" corresponds to the explanandum. If any part of the explanans (that is to say either the input/empirical claims or the mechanism/law-like statements) is the locus of new information, then a simulation can be said to be offering an explanation of the known phenomenon that corresponds to the simulation's output.

The historical uncertainty regarding explanation's place in philosophy of science (particularly in the early 20th century's positivistic climate) and the fact that there is broad dissatisfaction with the DN model of explanation was noted in chapter one. But there seems little doubt today of the fact that explanation plays a substantial, if somewhat cryptic role in scientific inquiry, and there have been significant efforts to address difficulties associated with Hempel's DN model and its inductive and statistical variations. Such efforts, including those addressing issues of relevance (Salmon 1971), the relationship between causality and explanation (Salmon 1984), and to a lesser extent, the ways that good explanations unify a broad range of related phenomena (Kitcher

1989)[63], all seem to buy in to a central dogma about scientific explanation, namely that good explanations are sound arguments for their explananda.

This presents a difficulty for fictional simulations and their suitability for explanatory functions, for soundness demands the truth of the premises of an argument. In the case of simulations that would putatively explain their target phenomena, this seems to preclude that they not introduce fictionalizing deviations into the input or mechanism moments of the simulational structure. A fictional simulation cannot be explanatory if explanations must be true. And unlike the case of the discoveries facilitated by the heuristic function of simulations, explanation is not an event that is spread across history, such that it may be initiated by a fictional simulation and completed in non-fictional laboratory practices and the reliable manipulation of target systems or their experimental variations. While there is a deep relationship between the way that simulations execute a heuristic function by articulating the salient "joints" of target phenomena, and explaining those phenomena, if we adhere to the dogma that good scientific explanations must be true, there would appear to be this crucial distinction between heuristic and explanatory functions: heuristic functions can be fulfilled fictionally, but explanatory ones cannot.

But this dogma concerning the truth of explanations is not universally adhered to, and notably challenged by Nancy Cartwright (see Cartwright 1983, 44-74). Not only does Cartwright permit that a good scientific explanation may fail to be true, she argues

[63] Kitcher's commitment to the soundness of a scientific explanans is somewhat weaker than folks like Salmon. There is a clear consonance between Kitcher's "unificationist" account of explanation and some like Cartwright's comments about grouping diverse phenomena together; yet Kitcher seems to stop short of fully abandoning the notion that an explanation must be a deductively sound argument for the explanandum. The stringency of the patterns used to describe diverse phenomena may be traded off in favor of some other concerns, but Kitcher nonetheless maintains that a proper explanation be as accurate as possible while fitting as many phenomena as possible.

that the *ceteris paribus* conditions that constitute the explanatory efficacy of fundamental theoretical laws over more accurate but impossibly complex and messy phenomenological laws are typically (if not necessarily) in conflict with the empirical adequacy of these laws. Thus, since "the laws of physics lie" and "the truth doesn't explain much," our best examples of scientific explanations routinely make appeals to theoretical laws that are not even approximately true.[64]

I am not particularly inclined to accept Cartwright's strong position that theoretical laws are not approximately true. In part, this stems from a desire to use the term "fiction" conservatively, and the narrow conception of fiction that I have developed in chapter three excludes the sorts of abstractions and idealizations that ground much of Cartwright's argument. But there is something intuitively appealing about the spirit of her position, that the effort to render phenomena mind-sized and group diverse phenomena together under a common model runs counter to efforts to make our accounts of phenomena empirically accurate. It is this sense of explanation that fictional simulations can aim at.

This is not to say that there is a safety net for all simulations that seek to fulfill some explanatory function. A retro-dictive simulation that aims to determine how the moon was formed from the collision of an asteroid with the Earth should yield knowledge about when this collision occurred, the size of the asteroid, and angle of impact that are at least approximately true. An explanatory simulation that aims to determine the mechanisms by which amino-acid sequences fold and combine to yield the quaternary structure of complex enzymes should yield knowledge of a mechanism that is at least approximately true. But these are scenarios where the point of new knowledge is fairly

[64] See Ch 1 for a somewhat fuller account of Cartwright's position.

precise and the parts of the simulation that are known are known with a high degree of confidence. Again, shifting from a context of successful simulations to promising ones dramatically changes the complexion of questions about whether and how simulations fulfill their various functions.

Grim *et. al.* make a gesture toward this more complex and epistemically impoverished simulational context by identifying simulations where the modal qualifications are particularly strong and the correspondence between simulation and target is particularly vague and ambiguous. These simulations are labeled as providing "emergence explanations" and are characterized by more than one part of their tri-partite structure that is unknown and/or poorly understood.

Noteworthy amongst simulations that provide "emergence explanations" are those that demonstrate how complex outputs may emerge from relatively simple inputs and mechanisms. Simulations of the emergence of "coorperative behavior" by agents in iterated prisoners dilemma games (Axelrod and Hamilton 1981), complex "life-like" patterns as in Conway's "Game of Life" (Conway *et. al.* 1982), flocking/schooling/swarming behaviors as in "Boids" (Reynolds 1987), or geographic segregation patterns (Schelling 1978) all make the suggestion that complex phenomena need not have complex explanations. Additionally, they are all open to a common criticism, that their inputs and mechanisms are markedly different from the corresponding features of their target phenomena. Such a criticism of Schelling's segregation simulation (which claimed to demonstrate how communities of agents may become segregated as a result of relatively mild preferences to not be minorities in their own neighborhoods) is made by Bruch and Mare (2006), showing that assumptions made by

Schelling's model not only fail to square with available empirical data, but that these inaccurate assumptions are crucial to the way the simulation functions and produces its output.

But I would argue that an undue focus on this sort of criticism, while it is both well motivated and productive, risks missing a remarkable quality of "emergence explanation" simulations. While it is entirely possible that such simulations may be deployed in an effort to bootstrap our way to a heuristic function, one in which the ultimate goal is a closing of the gap between a fictional simulation and a non-fictional articulation of the "joints" of target phenomena, they do something else as well. If part of the goal of an effective scientific explanation is to make phenomena "mind-sized," then complex emergent phenomena are particularly resistant to explanation. Following Cartwright, explanations that proceed by relatively accurate phenomenological laws may be true, but they are unwieldy and cognitively opaque. Radically simplified explanations, on the other hand, are less true (even downright fictional) but comprehensible. Part of the challenge of explaining complex emergent phenomena is helping us to see across the complexity barrier, to understand how complex phenomena can be explained in terms simple enough for humans to understand them. The resultant correspondence between a simulation that responds to this challenge and its target may be loose and weak, but it does not evaporate all together. If we demand that our explanations be sound, then these are deficient explanations. But if we also demand that our explanations be "mind-sized," then fictional "emergence explanation" simulations may be our only option.

Some, such as David Weinberger (2012), have suggested that complex phenomena and the age of "big data" science have ushered in a new way of doing science where we can no longer hope to understand the things we study, and instead outsource the "understanding" to machines. Such a shift is certainly happening, and factors significantly into the ways that simulational technologies satisfy pragmatic, predictive, and even heuristic functions. But this is not the only way to deal with complex phenomena. We can also simulate them in a way that is graspable, and this may require that our simulations be fictional.

Conclusion

Simulations can aim at and fulfill a variety of functions. The four functions discussed here, pragmatic, predictive, heuristic, and explanatory, are not intended to be exhaustive, but they do reveal a few things. First, they show a role for fictionalizing deviation from a simulation's target in each of the functions that have been addressed. Even when these functions have an unambiguously epistemic aim, substituting a surrogate with significant and non-negligible differences can prove effective, and even necessary. Thus fictional modeling and simulation practices are not merely pragmatically justified and epistemologically tolerated, but the ways in which a simulation is different from its target can be instrumental in achieving its epistemic purposes. Additionally, while the plurality of simulative functions have regions of overlap, they are quite distinct and sometimes in tension with one another. The need to satisfy non-epistemic pragmatic ends makes effective prediction difficult, simulations that only need to predict may do so without regard to whether their mechanism becomes so obscured as to forclose the possibility of genuine discovery, and the need to keep explanations "mind-sized" and accessible may demand that a simulation have only a very vague correspondence to its target, one where the “fit” between model and target is significantly strained. Furthermore, it is rare that simulations only aim at a single function. More often, a simulation study must adjudicate between a variety of aims, and over the course of a real research program these aims may shift from attending primarily to one function to focusing on another. Here again, a role for fictions is revealed, since the negotiation of a plurality of goals can result in the sedimentation of many layers of distortion that are introduced, compounded, excavated, reinterpreted, and sometimes removed.

Lastly, the use of fictions in the satisfaction and negotiation of these various functions walks a very fine line between simulations that attempt to represent their targets, and those that do not. This is the ambiguous domain of fictional representation, and whether or not a simulation actually *simulates* something in a way that permits epistemic access to it, or is instead merely a transparent means to some end or a functional substitute for a target is frequently uncertain and subject to change. It may depend on nothing more than the intentional attitude of the investigator who designs or uses a simulation.

Chapter 6
Make-belief and Belief

Introduction

Play is a relatively underappreciated and under-theorized dimension of human experience. Within the handful of canonical treatments of the subject, play occupies some fairly central role in human experience and the fact that vexing epistemological issues constitute recurring themes in said treatments speaks to our intuitions that play has something to do with knowledge/belief. Piaget identified the importance of play in childhood development, not only for honing motor skills but also (and even especially) for epistemological development (Piaget 1962; Ginsburg and Opper 1965). Johan Huizinga (1955) stresses the central role of play in religious ritual and belief. Hans Georg Gadamer's discussion of play is an engagement of predecessors like Kant and Schiller, all of whom place play centrally in understanding the sorts of knowledge produced in aesthetic experience.

But it may not be surprising (even if it is regrettable) that such discussions of religious and aesthetic experience and childhood development don't find much overlap with much contemporary work in epistemology. Contemporary epistemology is, in many circles, synonymous with philosophy of natural science[65] and "serious" epistemologists may be inclined to dismiss play as properly within the purview the sorts of knowledge and beliefs held by small children or the religious, or whatever sort of knowledge is produced in art. Even where an irreducible creative element in scientific practice is acknowledged, play is taken to be related only tangentially (or metaphorically) to the

[65] Gadamer's own treatment of play in *Truth and Method* is explicitly part of an effort to rethink the human sciences in a way that is distinct from natural science's fixation on method. (Gadamer 2004, xxi-xxv)

rigorous empirical study that "serious" scientists engage in. Nothing could be further from the truth.

Throughout the preceding chapters I have touched on various points that make promissory notes regarding a characterization of the intentional stance involved in fictional scientific modeling and simulation. Thus far it is unclear what it means to comport oneself toward a fictional model and take it as a surrogate for some non-fictional target. How do we do this, and what, if any relationship does that intentional attitude have to other attitudes like belief, acceptance, and commitment that we typically associate with a scientific attitude?

In chapter two, I addressed this in terms of "representing-as," following Elgin and Goodman, articulating a two-stage representational structure, the first of which represents something that may or may not exist and the second, which represents some real referent. This structure helps to elucidate that what is produced in fictional modeling is neither the representandum, nor the representans, but some third thing, the representans-as-representandum. In chapter three, I added to this by aligning modeling and simulation practices with prop-based make-belief, following Walton, indicating that the scientist's attitude toward the fictions that models and simulations instantiate is "something like belief." I also established a narrow view of scientific fiction that indicates a discontinuity between fictional and non-fictional that can be articulated in terms of whether and how a model can be de-fictionalized. In chapter four, I underscored the materiality of modeling and simulation practices and addressed the close but problematic relationship between simulation and experimentation. This analysis yielded a picture of complex research programs that are sometimes "building up," sometimes "cutting down," sometimes both, and where the status of simulations as representations of ideal experimental apparatus, autonomous constructions, and templates from which new experimental apparatus can be designed is fluid and reinterpretable. The discontinuity between non-de-fictionalizable models and their targets is echoed in the potential discontinuity between simulational studies that "build up" and experimental studies that "cut down." In chapter five, I followed up on the autonomy of simulational practice and the way that fictionalizing distortions can compound and sediment. I also touched on the intentional relativity with respect to the sort of relationship that a model/simulation has to its target; one where the prospective promise of a fictional model/simulation depends largely (though not entirely) on its functional purpose and on whether we intend to ultimately open up the "black-boxes" of its structures and mechanisms to more fully articulate the correspondences to the structures and mechanisms of the target.

All of these point to questions regarding the intentional attitude proper to fictional simulation, and further demand not only that these questions be explored from the perspective of the builder/user of fictional models *while those models are in use*, but also that we attend to the fragility of this attitude as these objects drift back and forth between being characterizable as fictional models, independent objects of inquiry, independent technologies, and non-fictional models. Furthermore, our attention is drawn to a need to account for the way that practices under this attitude exhibit a discontinuity with practices under a more "serious," non-fictional attitude, but can nonetheless become continuous through creative reinterpretations of models and discoveries of novel features of their targets.

We can begin by identifying the intentional attitude in question as "make-belief" and the activity of taking up models that are fictional props as an essentially playful one. The notion of scientific practice as sharing striking similarities to play is not new, though it is frequently noted in passing or with substantial qualification. Few get closer to taking the playfulness of scientific practice seriously than Robert P. Crease. In *The Play of Nature* (1993), Crease lays out an extended analogy between theatrical performances and scientific experimentation. The fact that this is "merely" an analogy should not be taken to indicate a half-hearted commitment, however. Crease goes to great lengths to emphasize the rigor of what he calls "argumentative analogy," and the way that he uses his theatrical analogy to scientific practice affects a reshaping of the way we understand both theatrical performance and experiment.[66]

Since it has already been argued that simulational practices share a great deal with experimental practices, it should not be surprising that a theatrical analogy finds ready application in simulation as well. For instance, Crease notes that experimentation is not merely a *praxis*, but also a *poeisis*, a staging of the events that bring phenomena into being (Crease 1993, 82), and my efforts to distance modeling and simulation practices from conceptions that would have us understand them as little more than extensions of theory make the same point about models as well.[67] But additionally, it is precisely where this analogy begins to break down for experiment — namely, the role of fiction and make-belief that are such key components of theatrical performances — that it finds

[66] An interesting aspect of this text is that the philosophical argumentative strategy Crease employs mirrors the hermeneutic circle by which theoretical "scripts" and experimental performances co-inform and reshape one another when scientists look to apply theoretical resources in novel contexts.
[67] Crease gestures towards this in "From Workbench to Cyberstage" (Crease 2006) but as mentioned previously, without attending to the ways in which the virtual staging of virtual phenomena admit of fictionalization.

an even stronger application in the context of fictional simulations. It was noted earlier that the advantage of performative accounts of scientific practice (that they find a way around the quagmire of anti-realist anxieties) makes its application to fictional contexts difficult. Ironically, Crease's approach, though "merely" an analogy for experimental practice, is surprisingly accurate for simulational practice. Thus, in this chapter, I am upping the ante on the analogy between scientific practice (at least in the context of fictional modeling and simulation) and play. Fictional simulational practices are not only "like" acts of mimetic play in certain respects (and unlike in others); they *are* acts of mimetic play. This requires a relatively careful consideration of what is meant when I say that a scientist engages in play. Accordingly, for the remainder of this chapter, I will take up the topic of play directly, only occasionally checking back in to scientific modeling and simulation to articulate the applicability of my conclusions about play and make-believe. Readers who would like more examples directly from the context of scientific modeling and simulation are encouraged to apply these conclusions to the examples given in previous chapters (or any other examples of fictional models and simulations that may be handy).

So what is this peculiar sort of intentionality that is involved in play, and which I am calling "make-belief?" When I qualify some statement or action with the disclaimer "I'm only playing," I signal that my words and actions should not be taken as they normally would. When we are "only playing," we aren't serious, and while this may be enough to qualify as peculiar, different than our normal comportment towards, actions in, and statements about the world, this hardly presents a philosophical puzzle. For many, a search for a definition or "essence" of a playful attitude may appear satisfied by a conception of play as the privation or negation of seriousness.

This conception is not quite satisfactory, however, in part because there is good reason to be skeptical of the suggestion that playfulness can be fully characterized by a privation of seriousness. Taking up a negative definition commits us to supposing an absence of something in play that is present in seriousness, but this seems difficult to accept when the effort required to adopt a playful attitude over a serious one seems to suggest just the opposite and when playfulness seems to have its own kind of "seriousness." These issues come to the surface in a particularly provocative claim made by Hans Georg Gadamer (2004, 101-110); that the distinction between play and seriousness is problematically ambiguous for the person immersed in play. Play, for the player at play, *is quite* serious, and, in fact, the recognition that one is merely playing seems to prohibit the special credulity that fully immersing oneself in play demands.

We would be right to distinguish this "special credulity" that is involved in make-belief from genuine belief, but in order to do so, we should be able to clearly articulate the basis on which such a distinction could be maintained, and herein lies the puzzle. To be reflexively aware that one is *currently* playing at something, say, for instance pretending that one is an airplane, is essentially to say "I am an airplane, but I don't believe it." Yet, Gadamer's claim about the "seriousness" of play seems to preclude the possibility of the reflexive awareness that would allow someone to make such a statement. The parallel here to G.E. Moore's paradox (frequently articulated in the statement "It is raining, but I don't believe it.") is instructive, as the problematic nature of a Moore sentence stems from the implication that one believe what she asserts. If it isn't belief, what sort of intentional attitude is at work when one asserts "I am an airplane" such that it does not permit that one may simultaneously assert "I don't believe it"?

If, indeed, one cannot authentically affect the playful credulity of make-belief while simultaneously holding a belief explicitly contrary to it, then it seems that being at play involves an intentionality that is not wholly distinct from belief. Likewise, if, as I am claiming, the use of fictional models and simulations involves make-belief, then our attitudes toward these models are also not wholly distinct from belief in them. If any distinction between playful credulity and belief can be made, it stems from the fact that play involves a fragile sort of "quasi-belief" that is similar enough to straightforward belief that one cannot simultaneously hold this quasi-belief and be explicitly aware that it is not a genuine belief. It is not any character of this quasi-belief when it is active that distinguishes it from genuine belief, it is the way that it breaks, precipitating the cessation of play.

This is not to advocate for a full-blown scientific anti-realism; that is to say, I am not arguing that science on the whole is based on an elaborate fantasy. I am arguing that insofar as fictional modeling has a place in scientific practice, there are elements of make-believe involved in this practice, and this element is more easily confused with genuine belief than we typically give it credit for. Thus, both engaging in playful fantasy and working with fictional models pose the risk of vicious ways of make-believing, and a virtuous mean that moderates these vices. We must be able to take fictional props up with enough "seriousness" that our engagement with them can fully run its course, and for them to take on an autonomous character. If we are too circumspect, fictional props cannot function properly and fictional models cannot fulfill their scientific functions. But we cannot be so credulous that our make-believing is not capable of breaking, allowing us to distinguish between models and targets. Thus this chapter concludes with the recommendation that we begin to take make-belief seriously as a scientific virtue, one that can be corrupted into vicious extremes.

Specifying the sense of "play" that is at stake

A distinction can be drawn between relatively unstructured and more deliberate and structured play. Roger Callois (2001, 27-35) labels these distinct sorts of play *paidia* and *ludus* respectively. A paradigmatic example of *paidia* is the child who spins in circles to the point of dizziness. In more "serious" contexts, this sort of structureless play can be found in a sort of blind and aimless tinkering with things to see how they behave without being committed to any particular practical projects. Even slightly more purposive tinkering —of the sort that is at work in simulations that are tweaked and kludged toward accurate prediction, or that folks like Andy Pickering seem to be talking about when they speak of "tuning" in the "mangle" of experimental practice (Pickering 1995)— has something of the flavor of Callois' *paidia* to it. Regardless of whether this tinkering is quasi-goal directed or not, there is a lack of structure in *paidia*, and what is particularly notable about this is that it is accompanied by the inability to express some proposition that would be true in the context of play, but not outside of it.

In contrast, *ludus* has structure; we might even say that it has rules (even if those rules only aim to elicit the spirit of a game without fully capturing it).[68] While we may

[68] Crease might refer to this as a "script," but in his theatrical analogy, "scripts" correspond roughly to theoretical principles that loosely guide but underdetermine what is produced in an actual experimental performance. In the context of fictions and make-believe, ludic structure also includes those principles that assert fictional deviations from reality. Hence we would do well to think of ludic structure more in terms

play or tinker without any particular direction to our activity, only *ludic* play can rightly be called a *game*.[69] The sort of play that involves fantastic make-believe is *ludic*, for there is a rule-like structure to the way that fantasies are asserted and unfold. Additionally, "mimetic" play, in which props or surrogates are used as stand-ins for something else require rule-like structure in the form of what Kendall Walton calls "principles of generation" in order to designate what these props represent and determine how they are to behave as play proceeds. In all of these cases of *ludic* play, it is possible to articulate some aspect of this structure propositionally. More to the point, we can clearly articulate aspects of this structure that *only* apply in the context of the game, not outside of it. Hence, while I may be able to say "Don't jump into the lava, you'll get burned!" and believe this, both while at play and in the seriousness of real-life, it is only in the context of some sort of *ludic* play that I can designate the floor as "lava" and say "If you fall off the sofa, you'll get burned."

Thus it is in *ludic* play, where the possibility of fictional rule-like structures that somehow run contrary to ordinary activity exist, that the possibility of asserting propositions which we wouldn't seriously assent to emerges. Statements like "I am an airplane," "This [pointing to a stick] is a sword," "Zombies are compelled to feast on the brains of the living," or even "The pendulum swings on a massless string about a frictionless pivot in a uniform field of gravity," "There are 'silogen' atoms at the boundary between the quantum mechanical and molecular dynamical region of a fracturing silicon brick," or "The flow of a fluid is obstructed by a two-dimensional

---

of what Kendall Walton "principles of generation," those loose, prescriptive principles that instruct how a prop is to be interpreted and used.

[69] This distinction between a general sense of 'play' and 'game' is particularly vivid in English, as opposed to German or French where the same word is used for both (*spiel* and *jeu* respectively).

cylinder" may all be followed with the phrase "but I don't *really* believe this." In such compound statements, speakers seem to be attempting to capture something from a playful context and a serious one simultaneously. These are the sorts of play scenarios I am concerned with in this essay.

A Word on Pragmatic Conceptions of 'Belief'

It may be objected that in taking beliefs that are made explicit, and represented propositionally, I am misrepresenting the sort of pragmatic approach to belief that characterizes much of scientific belief. **"**When scientists "believe" something, it's not like believing in a creed. It's more like the way we express belief in gravity by the way we walk."[70]

I agree with this sort of pragmatic approach to belief (and even, perhaps, with a Peirce-ian maxim that two beliefs with identical practical consequences are, in fact, not distinct beliefs at all). However, I resist the assumption that the concept of belief can be reduced to "habit of action." but wish to maintain that it entails a mental disposition as well. Habits of action may indicate the content of beliefs, and belief may entail a commitment to action, but this does not mean that habits of action are identical to beliefs.

This should not be taken to be a position of epistemological internalism. I do not require that a scientist, or anyone else for that matter, be explicitly aware of their own belief in order for them to believe it. Thus, I agree with a qualified version of the objection; *many* beliefs that scientists hold are not like beliefs in a creed, but more like the way we express belief in gravity by the way we walk. However, while this accounts

[70] The precise wording of this objection comes from Dr. Robert P. Crease. I thank him for drawing it to my attention.

for a huge domain of scientific beliefs performed in experimental practice, a domain that has been historically ignored, this does not mean that scientists never become explicitly aware of their beliefs and express them propositionally. In fact, this is a very important (and, I would argue, necessary) aspect of scientific inquiry that makes it a special kind of technical practice, one where we attempt to formulate our habits of action into explicit theoretical propositions.

Much of my discussion of belief and make-belief in terms of propositional content (rather than habits of action) is for the sake of analytic convenience. I would not claim that only those habits of action that can be expressed propositionally count as beliefs, only that beliefs that are propositionally expressible are preferable for the sort of analysis I am carrying out, if for no other reason than we can have a relatively clear sense of what is believed in those cases. Those propositions may be short-hand for something more complex, including a commitment to a habit of acting, but it ought to suffice to refer to some belief 'P' as expressible in the propositional form "That P" even if the believer never entertains such a proposition in her mind or utters it aloud. Furthermore, it is, in principle, possible to become explicitly aware of one's beliefs, and such awareness is also propositionally representable (*e.g.* "I believe 'that P'").

If our focus is on the sort of intentional attitude proper to fictions, namely make-belief, then it is plausible (again, for analytic convenience) to set aside cases of tinkering and structureless play, such as those that Roger Callois labels *paidia*. Not only are these cases awkward for the reasons expressed in the previous paragraph (what, precisely, is believed/make-believed when we tinker or dance?), but even if we decided that we must be tolerant of ambiguity in such cases (perhaps even noting that this is one of the main points of taking a pragmatic approach to the concept of belief), there are no discernible habits of action that are unique to a playful attitude and absent from a "serious" one. Note that I am not saying that nothing is believed or make-believed when one tinkers or dances, only that it is difficult to pin down precisely what is believed/make-believed and how these beliefs/make-beliefs are related to genuine beliefs. Accordingly, since I am interested in exploring what makes make-belief different from belief, I focus on cases where there is something that is make-believed that we would never genuinely believe, where that something is expressible in terms of the rule-like structure of a game. This is not to suggest that we are scrupulous rule-followers when we play, or that the propositional expression of rules fully captures the habits of action of players, just that the propositional expression of rule-like structures is one way of capturing the spirit of whatever it is that we make-believe when we play at a game, and this way is adequate for my analysis. If this is not an acceptable position, then I am at a loss as to how we might talk about beliefs and make-beliefs with any analytic rigor.

Play and Seriousness

For Johan Huizinga, whose seminal work *Homo Ludens* is a touchstone for all contemporary philosophical reflections on the topic of play, the distinction between the contexts of play and seriousness is a fairly sharp one. Play is, he maintains, marked off from seriousness by its superfluousness and the freedom with which one enters into it. "Play can be deferred or suspended at any time. It is never imposed by physical necessity or moral duty. It is never a task" (Huizinga 1950, 8). Furthermore, "play is not 'ordinary' or 'real' life. It is rather a stepping out of 'real' life into a temporary sphere of activity with a disposition all of its own" [*ibid,* 8], and this distinction from 'ordinary' life functions with respect to "locality and duration." "It is 'played out' within certain limits of time and place." [*ibid*, 9] Thus Huizinga gives two criteria for play that distinguish it from seriousness, one concerning its freedom from necessity and/or duty, and another concerning its spatio-temporal boundedness.

The fundamental premise of this chapter is that fictional scientific models involve acts of playful make-believe, but objections may be raised against this premise according to each of the two criteria Huizinga asserts. Frequently, we do model out of pragmatic necessity, and it might even be said that such pragmatic concerns are the primary motivating factor of decision to model rather than investigating a target of inquiry directly. While fictional deviations in scientific models and simulations may arise out of pragmatic, moral, or political concerns (suggesting a tension with Huizinga's claim that play is never imposed by physical necessity or moral duty), it should be noted that fictional models that are employed solely for the sake of a pragmatic function and reducible to their functional, technological utility were previously pegged as having an especially strained relationship with their targets. It is debatable whether and to what extent such "models" actually model. The far more compelling cases of fictional modeling are ones where there is an active (if vague) promise of an unarticulated robust correspondence between the structure and mechanism, one where heuristic, hermeneutic, and explanatory aims are sought in order to understand the target. While it is well established that science is inextricably tied up with technology (both in instrumental praxes as well as the practical problems that motivate research to explore particular projects) there is an unmistakable dimension of scientific research that is not reducible to practical necessity. That we desire epistemic access to phenomena and seek to understand them is something that we do for its own sake, not because it is forced by any necessity. Insofar as there is this dimension of scientific inquiry that is not reducible to technology, we engage in it simply because, as Aristotle said, we "desire to know," not out of physical necessity or moral duty.[71]

While Huizinga admits a certain fluidity in the contrast between play and seriousness (*ibid*, 8), the limiting spatio-temporal boundaries of the playful context is a hallmark of his account. We know when and where the playful context begins and ends by reference to boundaries that delimit the game; the stage after the lights go down and before they come back up, between the opening and closing whistles and within the chalked off boundary lines, on the "consecrated playground" until the streetlights come on. These boundary lines set the limits of the playful context in a way that play begins and ends when we cross them. Within these boundaries the rule-like structures of play are in effect and real-world interests are checked at the door, and here skeptics about the playfulness of scientific modeling might object again, since scientific inquiry does

[71] We may equally note the response of Robert Wilson, then director of Fermilab, to questions during a congressional hearing about the national security applications of Fermilab's particle accelerator. Wilson said it had no such applications at all. "It only has to do with the respect with which we regard one another, the dignity of men, our love of culture. It has to do with those things. It has to do with, are we good painters, good sculptors, great poets? I mean all the things that we really venerate and honor in our country and are patriotic about. It has nothing to do directly with defending our country except to help make it worth defending." While it is plausible to claim that scientific inquiry frequently aims toward technological practice, it is notable that even for folks like Ian Hacking, who takes a particularly strong position in this respect, these practices are often characterized by the skillful manipulation of entities *in scientific inquiry*. What is on the line here is whether the seemingly playful modeling and simulation activities that scientists engage in can be legitimately considered to be instances of play, and an affirmative answer to this question must answer a pragmatist objection that serious inquiry like science is tied through-and-through to problems that present themselves in the physical necessity of practical life. But this thesis is unconvincing if taken too strongly to indicate that scientific inquiry never takes its eye off of "serious" practical goals. This is not to suggest that we should buy in to the hubris of modernist depictions of "pure" theoretical science as having a privileged epistemological priority over technological applications. It is merely to say that scientific inquiry does distinguish itself from practical problems in moments of "pure" inquiry, and in these moments it is pursued for its own sake.

produce real-world applications that would cross a supposed boundary between a playful context and a serious one. In fact, a progressivist account of fictional modeling would seem to require that this boundary be undermined.

But even in clear cases of play, these boundaries are not always so sharp and impermeable. Callois (2001, 3-7) provides an excellent example of a case where they are not. Gambling, or so-called “games of chance” confounds Huizinga’s contention that serious, real-world interests are excluded from the play-ground. We may place friendly wager on any game, and all professional game-players have economic interests tied up in the outcome of play, but one might easily dismiss these as accidental to the essential aims of the game itself. For gamblers, however, real-world consequences are essential to the game they play. As anyone who has played poker with chips that have no real monetary significance attached to them knows, the game ceases to function in its proper way when no real risk is involved and a player may bet without impunity. It is part of the essence of such games of chance that they not be divorced from real-world interest.

Callois seems less inclined than I am to take this as a cue to be wholly suspicious of Huizinga’s claim regarding the closedness of the boundary between play and seriousness.[72] Nonetheless, the integrity of the boundary is clearly compromised in the case of gambling, and it is clear that the knowledge that the outcome of the game has consequences outside of its own context can be an essential factor in maintaining the spirit of the game as such. The same may be said in response to those who would resist

[72] Callois attempts to rescue Huizinga’s point by maintaining that while material goods change hands in play, none are produced, which as a means for distinguishing play from non-play is neither necessary nor sufficient, and anyway, hardly seems to strike at the essential difference I am after which concerns varying sorts of credulity.

the idea of the playfulness of some scientific practices on the basis of their essential orientation toward real-world consequences.

Furthermore, an excessively naïve reading of the claim that play is entered into and exited freely runs afoul of counterexamples as well. Huizinga himself points out the phenomenon of the spoil-sport, the problematic player who only need to refuse to play, take his ball home, or even simply point out to everyone that the game is *only* a game in order to disrupt play. When the spoil-sport ends the game, she does so against the will of the other players. We might also add the zealot to our taxonomy of problematic players, who takes the game too seriously and seems to confuse it with real life. In light of such familiar examples of problematic players, it is difficult to maintain that players may commence and cease play freely. Players may be at the mercy of the other players' willingness to "play along," some may be carried off by it and require external intervention to stop (even after the "whistle" has been blown), and sometimes, play may sneak in without announcing itself, permitting players to recognize only in retrospect that a game has commenced and even then without any certainty of exactly when this happened. These sorts of spontaneous starts, stops, and continuations of play have little to do with the free will of all the players or the spatio-temporal boundaries. It is not the case that spatio-temporal boundaries of the play-ground signal the switch from a playful attitude to a serious one or *vice versa*. Rather the switching of this attitude *is* the boundary of play, and this boundary is vague, porous, and fragile.

Thus, in response to Huizinga's emphasis on the delimiting boundaries of play, we should be inclined to recognize that play and games are ambiguous (Sutton-Smith 2001), amphibolous (Spariosu 1989), liminal/liminoid – i.e. on the boundary between reality and unreality (Turner 1969), and/or that it simultaneously is and is not what it appears to be (Bateson 1956). Such descriptions are more poetic and obscure than we might hope for, but serve as a corrective for an understanding of play as unproblematically opposed to seriousness, and set off from it by way of a sharp and closed spatio-temporal boundary. Furthermore, the resonances of this conception of play with the way that models shift ambiguously in their relationship to their targets over the course of their use in larger research programs should be evident.

It is this problematic ambiguity that Gadamer invokes in his account of play. His own professed goal is to “free the concept from the subjective meaning that it has in Kant and Schiller” and which we have seen echoed by Huizinga. Rather than understanding play as something executed by a free subject, who can commence and cease the game at will, play is something that possesses the player, and enlists her in a performance through which the game reveals itself. Even the most rigidly structured *ludic* games overflow their formal rules and the intentions of the players to give rise to a novel creative product. In short, the game plays the players as much as (or more than) the players play the game.

It is in this pull, the way that play draws the player into the game, that the paradoxical ambiguity of play lies.

> The player himself knows that play is only play and that it exists in a world determined by the seriousness of purposes. But he does not know this in such a way that, as a player, he actually *intends* this relation to seriousness. Play fulfills its purpose only if the player loses himself in play. Seriousness is not merely something that calls us away from play; rather, seriousness in playing is necessary to make the play wholly play. […] The player knows very well what play is, and that what he is doing is “only a game”; but he does not know what exactly he “knows” in knowing that. (Gadamer 2004, 102)

Hence, there is a strange double consciousness in play,[73] such that in order to be fully immersed in play, one must be serious about the play itself. To become explicitly aware of the fact that the play is "mere play," that "it is only a game," is to lose this seriousness. When the game reveals itself as not serious, it evaporates.

For the player absorbed in play, which is to say for the player who is *really* playing, the game is serious but not serious. Perhaps there is a sly equivocation here, for the "seriousness" of play is not the seriousness against which play distinguishes itself. Nonetheless, Gadamer insists that these two seriousnesses are similar enough as to not permit both conjuncts of "The game is 'serious,' but it is not serious." to be held explicitly and simultaneously.

The paradox of play, that it is "serious" and "not serious," can be expressed even more concisely as the simple reflexive assertion "I am playing right now." To be able to assert this, one must know that the current activity is not serious, that it is *mere* play, and this makes genuine play impossible. Interestingly, non-reflexive formulations of the same propositional content are not problematic. I can say tomorrow that "I was playing yesterday," or someone else may say of me "He is playing now," without implying a contradiction. But, following Gadamer, for me to express this content in the first person present indicative is to utter something absurd. If I know that I am playing right now enough to be able to assert as much, then I cannot, in fact, be playing right now. "Serious" play will have been revealed as mere play, and (at least temporarily) been

[73] Gadamer credits Huizinga with the identification of this double consciousness, quoting as passage about the "savage's" inability to distinguish between play and non-play (Gadamer 2004, p104). Whether or not Huizinga has a more subtle touch than I have given him credit for, it is fairly apparent that Gadamer intends this ambiguity between play and seriousness as a problem that is not peculiar to the "savage" but to all players. Furthermore, Huizinga's emphasis on the freedom of the playing subject to initiate and end play, as well as his emphasis on the spatio-temporal boundaries of play have been shown to be contrary in spirit to the key point I take Gadamer to be making.

destroyed. Thus the statement, “I am playing right now” is false whenever uttered (provided that we understand “play” in the strict sense of being absorbed in play).

There is a similarity between this sort of statement and the Cartesian assertion “I am thinking,” with the obvious difference that this latter statement *is true* whenever it is thought or uttered. Both are reflexive, referring us to the speaker at the time they are speaking. The difference in truth-value between them is due to an obscured syntactical difference. Were we to negate the predicate of the Cartesian assertion and change it to “I am *not* thinking (right now),” then we have a statement whose propositional content is not impossible, but is false whenever uttered. Similarly, a speaker who says “I cannot speak right now,” expresses a state of affairs that is, in principle, possible, but is a performative contradiction. All are of a family of statements that are not only reflexive, but contain some sort of negation. The unspoken negation in “I am playing right now” is buried in the ambiguity of play’s non-seriousness.

We can also see this hidden negation captured in the specific content of what is “make-believed” in play. In play, one may assert or act as if they believe something that they do not believe in any strict sense. As mentioned earlier, the player may say “This [stick] is a sword,” but the expectation is that insofar as they are only playing, they may also add “…but I don’t *really* believe this.” This form helps to demystify Gadamer’s statement “The player knows very well what play is, and that what he is doing is “only a game”; but he does not know what exactly he “knows” in knowing that.” The player knows the stick is a sword and she knows that she doesn’t really believe this, but something blocks her from being explicitly aware of both simultaneously. When I say, “this stick is a sword,” there is the tacit awareness that the stick is a stick and also that within the confines of the game it is to be taken as a sword. But a player cannot be too aware of this, cannot add "…but I don't *really* believe this," without breaking the illusion that being absorbed in play requires, that is to say, without ceasing to *really* play.

The same goes, *mutatis mutandis*, for working with fictional models. In order for a model system to function as a representation of the target-system-as-model-system, in order for it to play out its hermeneutic role, it must temporarily draw its user into a virtual world populated with 'silogen' atoms, idealized pendula, two-dimensional cylinders, or whatever fictions may be instantiated by the model system. We may know that the model system is just, for example, the hardware and software of the computer, or that the real target is not idealized and fictionalized in all the ways the model represents it as being, but we cannot be explicitly aware of these things *while the model is at work* (or perhaps more accurately, when it is *in play*). It is as easy to say "I am fictionally modeling right now" as it is to say "I am playing right now," but I maintain that both are equally false whenever they are uttered.

Moore's Paradox and Belief

As noted earlier, the recognition or articulation of the structural elements of play as propositions that attempt to straddle a playful and a serious attitude, e.g. in the statement "This [stick] is a sword, but I don't *really* believe this" (or "the cylinder has no ends, but I don't *really* believe this") bears a striking resemblance to Moore's paradox. Moore's paradox is usually represented with a statement of the form "X, but I don't believe X" (e.g. "It is raining, but I don't believe it). While the "paradoxical" nature (i.e. the implication of allegedly contradictory propositions) of such a statement is not readily apparent, anyone listening to a speaker who utters such a statement is bound to get the impression that something strange has happened (Moore 1993, 211).[74]

The reason that contradictoriness is not apparent is that the two conjuncts of a Moore sentence don't refer to the same thing. The first conjunct, "It is raining," refers to a state of affairs in the world, while the second, "I don't believe it," refers to a state of affairs in the speakers own mind. Hence, it is difficult to demonstrate how the conjuncts are in tension with each other. As was the case for the play paradox discussed above, this becomes clearer if we formulate the propositional content differently (e.g. if someone says of me "It is raining, but he doesn't believe it." or if tomorrow I say "It was raining yesterday, but I didn't believe it"). Such statements assert the same content as "It is raining (now), but I don't believe it (now)," without the same apparent tension as the original Moore sentence. As before, there is a deeper peculiarity in Moore's paradox lying in the fact that statements bearing the explicit or tacit reflexivity of first person and present tense indexicals are problematic, while other formulations, even those with the same propositional content, are not. (Moore 1993, 208-209)

There are ways to avoid the strangeness of a Moore sentence. One may split the conjuncts with a pause between the first and second conjuncts to imply that the speaker has changed their mind or otherwise intends to rescind the assertion that "X," or by switching audiences mid-sentence as if to say to one person that "X," and then to another that this was not spoken in earnest, implying some attempt to deceive. Splitting the conjuncts in these or other ways avoids the problem, but these are unusual sorts of

[74] Again, one might object to the use of propositionally explicit beliefs/make-beliefs as a model for our analysis. But I stress here that the central problem addressed in this chapter concerns what happens when the content of beliefs/make-beliefs and the fact that we believe/make-believe them becomes relatively explicit. I do not think that anything said in the following arguments runs contrary to a pragmatic approach to belief, provided that we can allow that beliefs can, and frequently do become explicit to the believer.

utterances that don't address the problematic case where the sentence expresses a single idea; namely the *conjunction* of the assertions that "X" and that "I don't believe X."

Moore points to a tacit norm in assertoric implicature, namely that we believe what we assert, in order to explain why these sorts of statements seem strange. (Moore 1993, 211) Thus I may speak of someone else not believing X while implying that I do, or that yesterday I didn't believe X while implying that now I do, without such implicatures suggesting a contradiction. But if I state that *I currently* don't believe X while implying that *I currently* do believe X, it is understandable that this is a confusing or nonsensical utterance. If something seems to have gone wrong when someone utters a Moore sentence, it is because this norm has been violated. We may violate norms of language without necessarily violating logical *laws*, thus putatively sparing us from outright contradiction. This is all well and good for the sake of explaining the paradox away if we understand the crux of it to concern why two sentences can have the same propositional content while one is problematic and the other is not. But it's not clear that this explanation is complete or even entirely true.

First, the suggestion that the offending sentences reveal a norm of language use does not tell the whole story. Moore states, 'I don't see that there's any sense in which you can be said to be using language improperly by saying something assertively when you don't believe it. When you are lying, it doesn't follow that you are using language improperly." (Moore 1993, 211) This seems to suggest a distinction for Moore between the violation of norms in a linguistic community and "using language improperly" where one may violate these norms without necessarily risking outright contradiction. Perhaps this is correct, and there is nothing inherently contradictory about lying. What is left unspecified, however, is whether one can switch in mid-sentence between violating this norm and adhering to it. Can I say that "It is raining, but I don't believe it," where I am lying about it raining but not about my not believing it without either "using language improperly" or splitting the conjuncts and essentially uttering two distinct statements under the guise of one? It would seem that a conversational norm like "believe what you assert" or "assume that others believe what they assert" may be contextual, that is to say that it may be broken without operating outside the rules of language itself, but is nonetheless not the sort of thing that can be violated at will by a competent language user.[75] To express the unified proposition of a conjunction, it seems problematic to follow a particular assertoric norm for one conjunct, but not for the other.

Second, one might be able to adhere to the norms of an assertoric context without necessarily believing what they assert if the standard for assertion, even honest assertion, is lower than the standard for genuine belief. There are various terms which are employed by philosophers to indicate intentional attitudes less than belief but more than disbelief, including 'acceptance,' 'commitment,' and a spectrum of probabilities/degrees of belief, all with idiosyncratic technical extensions. This suggests the possibility that one may be inclined to assert something without also being inclined to assert that they believe it. It is plausible to suppose that the credulity involved in playful make-believe, which I am holding to also be the credulity involved in fictional modeling, can be assimilated into one of these alternative intentional attitudes.

Belief and its derivatives – 'Acceptance' and 'Commitment'

[75] It's not clear that Moore would disagree with this, but the point is left vague in his own treatment of the paradox.

Moore's paradox, if indeed it is a paradox, derives its character from the fact that the first conjunct implies something that contradicts the second, namely that one believes what one has asserted. The various forms of our paradox concerning play share some structural similarity with Moore's paradox, where the explicit awareness of disbelief precludes the credulous absorption required for genuine play. But we may hesitate to concede that this similarity means that the credulity implied by playful actions or assertions is the same intentional attitude as belief. Even if we do take the structural similarities between our play paradox and Moore's paradox seriously, it has been suggested that a possible way around Moore's paradox would be to offer a standard for earnest assertion that is weaker than belief, such that one may assert "X" without believing it. In this section I consider whether such strategies for diffusing Moore's paradox can be applied to the play paradox, and by extension to fictional modeling practices. If they can, they would suggest ways in which the intentional attitude of playful credulity could be something other than belief.

One way out of Moore's paradox is via Jonathan Cohen's distinction between belief and *acceptance*. (Cohen 1992) While acceptance shares quite a bit with belief, there are key differences. Chief among these is Cohen's definition of belief as a passive while acceptance is active. This invokes a somewhat Humean view where belief is "a disposition normally to feel that proposition is true" in conjunction with a somewhat Cartesian view that we are completely free to accept or reject any of these passive psychological states in action or speech. (Engel 2000, 8-11) Thus one may accept something contrary to belief, as van Fraassen (1980) suggests that scientists do when using scientific models, or as Bratman (1999) describes when a person believes that a ladder is stable, but double or triple checks anyway, since prudence demands that they not necessarily accept this belief.

This should raise red flags for application to our previous characterization of play. If playful credulity is nothing more than a form of acceptance rather than belief, then it must be freely and actively taken up in contrast to the passivity of belief. It is precisely this freedom that Gadamer challenges and complicates in his effort to wrest play from its subjectivist history. A distinction that maintains that playful credulity is distinct from belief on the basis of activity and passivity glosses over an undeniable phenomenological truth about game-play; namely that being immersed in play is, in large part, a passive/reactive experience. It is this passive-reactivity of being absorbed in play that is broken upon the recognition that one is “merely” playing, and the resumption of authentic play requires the forgetting of that recognition.

There certainly are elements of willfulness in the practices of play, fantasy, and make-believe, particularly in the form of what is sometimes referred to as “willful suspension of disbelief.” To begin to play at something, to make-believe that a prop is something other than it is, it is true that such a willful suspension is necessary. But the phenomenological point that Gadamer makes is that this is not consistent with being absorbed in play (therefore not yet *genuine* play), and I am saying the same about the use of scientific models. Some willful suspension of the knowledge that the squares on the computer screen are not people and that real people don’t behave the ways that agents in a simulation behave is required to begin to model. But in order for a model to fully play itself out, we must go beyond willfulness of ‘acceptance,’ and be swept up in the model’s virtuality.

Alternatively, we may use the term "commitment" to capture much of the distinction from genuine belief that "acceptance" offers, but without the problematic voluntarism that Cohen's conception of acceptance requires. "Commitment" may indicate an inclination to act or speak in a certain way in an appropriate context, whereas belief is not so contextually constrained. I may not voluntarily *accept* this contextual commitment, instead passively/reactively responding in accordance with it when the situation presents itself. Thus I may acknowledge this commitment without assenting to its validity in all contexts, as I might when performatively acknowledging a commitment to always stop at red lights when driving, without assenting to the general belief that red lights *always* mean 'stop.' There is an appealing deflation of the puzzle here, as it seems that the apparent confusion stems from a failure to specify the specificity/generality of a statement like "A red light means ' stop,'" "All the tree stumps are bears," or "the ether rotates counter-clockwise around the axes of the lines of magnetic force." Such statements may function well enough if limited to certain contexts, but not if extended beyond those limits.

However, application of this solution to play and modeling is still problematic. First, while we have avoided the voluntarism of "acceptance," we have done so at the expense of making the distinction between belief and commitment a function of context. This requires that the context be logically prior to the shift in attitude from disbelief to commitment. It also requires a clear distinction between the attitude of commitment and the attitude of belief to depend upon a clear distinction of the boundaries that delimit the context in question. We have already seen that the boundary between "play" and "seriousness" is not so clear, distinct, or closed. Furthermore, it was also shown that the shift from a serious to a playful attitude is not a function of crossing some supposed boundary (however vague it might be), but rather the boundary is a function of the shift in attitude.

Likewise, the fluidity of the representational relationship between model and target suggests problems with harnessing a distinction between the credulity we bring to models and genuine belief to the concept of "commitment." It is true that certain scientific models are best applied in specific contexts, but these contexts are neither clearly delimited nor are they prior to our engagement with and constitutive of the intentionality we bring to those models. Rigidly contextualized models are predominantly found where the model has been instrumentalized for some predictive or functional purpose. For the hermeneutic contexts that models with heuristic and explanatory functions work within, it is part of the function of the model to bring the boundaries that delimit the model's context into question and attempt to expand them. Thus, these sorts of models, particularly when they involve fictional elements, are more properly examples of playful make-believe than context-dependant "commitment."

Where 'commitment' is concerned, there is nothing preventing us from simultaneously being aware that a certain action is appropriate for the context we are currently in, and also that this action is not necessarily appropriate outside of this context. Yet we have argued that such simultaneous awareness is impossible for genuine playful make-believe. This is the source of the paradox that Gadamer presents us with, and why we seem to be unable to be absorbed in genuine play while being aware that we are merely playing. The "player" who is merely committed to the rules and structures of a game, in the sense that they can be consciously following a rule like "Acts as if the stick is a sword while the game is in session, but not otherwise," is simply going through the motions.

Authors like Bas van Fraassen, who argue that our engagement with scientific models is characterized by some form of acceptance or commitment (Van Fraassen 1980), do so in order to distinguish the intentional attitude of modelers from belief. But it seems clear that if there is a substantial element of make-believe in the hermeneutic relationship between models that fictionally represent their targets *as-if* they were other than they actually are, then the above characterizations of acceptance and commitment won't do. Given the way that genuine absorption in make-belief precludes the simultaneous recognition of disbelief, it appears that this intentional attitude is strikingly similar to belief, yet somehow not identical to belief. We still lack some clear way of characterizing make-belief that articulates its difference from belief.

Quasi-Belief

Since we do not enter into or exit from play with complete freedom, nor do we do so in response to clear boundaries that delimit the playful context, it would be incorrect to identify the attitude that one brings to scenarios involving play, fantasy, or make-belief with willful "acceptance" or contextually delimited "commitment." I am claiming that the same goes for the attitude that one brings to fictional scientific models as well, specifically insofar as they serve a hermeneutic function by representing their targets as-if they were other than they really are. I am also claiming that for ordinary forms of play and scientific modeling both, there is a necessary absorption in the game/fiction that results in make-belief looking an awful lot like genuine belief. This absorption is characterized by an inability to be explicitly aware that one does not really believe what is implicitly or explicitly assented to in make-belief.

But it would be epistemologically problematic if we were unable to make a clear distinction between scientific make-belief and scientific belief. We would no longer be talking about scientific fictions that are sometimes utilized as a tool for developing scientific non-fictions, but rather edging toward a full-blown scientific fictionalism. In order to retain a space for scientific fictions within a moderately realistic conception of science, we must identify and characterize the sort of intentionality that is involved in make-belief, such that it is distinguishable from belief. This can be done through what François Recanati (2000) refers to as "quasi-belief."

Recanti's concept of "quasi-belief" is characterized by propositions that introduce some sort of tension, where there are reasons for credulity as well as reasons for skeptical caution. He develops this concept in response to Dan Sperber's account of how the hermeneutic attitude gives rise to a peculiar sort of "belief" that is different than that produced under an ordinary descriptive attitude. At the most basic level, Sperber is concerned with what it means to "believe" something that one does not understand. He offers, as a paradigmatic case of this, the belief put forward by Lacanians that "the unconscious is structured like a language." What makes this "belief" peculiar, Sperber argues, is that it is not clear just what this means, even to the Lacanian who professes to "believe" it.[76]

We can distinguish this "believing-without-understanding" from other, more normal ways of arriving at belief by contrasting it with a descriptive attitude. Under the

[76] This is not intended (by me at least) to be a slight against Lacanians, nor am I particularly committed to the suggestion that it is true that Lacanians don't know what it means when they say "the unconscious is structured like a language." I use it as an illustrative example of Sperber's take on quasi-belief.

descriptive attitude, we first determine what propositional content is expressed in a sentence, and then we determine whether or not this is true. Sperber's Lacanian however, first takes the statement as true, and then aims to interpret what it represents. They are able to do so because they believe a validating meta-proposition, namely that "Lacan says, ' The unconscious is structured like a language.'" By couching the statement "the unconscious is structured like a language" as a quotation in a meta-belief, it is temporarily insulated from the rest of our genuine beliefs.

We frequently "emancipate" propositions from this quotational insulation. A great deal of education seems to work this way. We are told "X" on the authority of someone else, even if we're not sure what it means. After some dialectical struggle to interpret "X" and experientially confirm its truth, there is no need to keep it in the meta-propositional form "S says 'X.'" However, sometimes this emancipation is blocked. Specifically, Sperber points out, if a statement remains semantically indeterminate, we are left with a "belief" that we are attached to because of some deference to the authority that validates it, but which we do not understand. (Sperber, 1975; 1985)

While Sperber is content to count this as belief (albeit a peculiar sort that is without understanding) (Sperber 2000), it is certainly a type of "belief" whose epistemological warrant is suspicious, perhaps even irresponsible. At the very least, a species of belief whose warrant rests on deference to some authority ought to strike us as distinctly unscientific.[77] I am resistant to Sperber's decision to regard "believing without

---

[77] *cf* the distinction between the method of "authority" and the scientific method made by Charles Sanders Peirce in "The Fixation of Belief." (Perice 1877) The notion that scientists do not fix their beliefs via the method of authority is significantly challenged by many today (notably Latour (1979), who points out that "creating a laboratory" to test and verify the claims made by fellow researchers can be prohibitively difficult, and thus that much scientific credibility is derived predominantly by authority). Nonetheless, as argued in CH 1, the unavoidable presence of social dimensions in scientific practice is hardly enough to

understanding" as a species of ordinary belief. I'm more inclined to follow Recanati here, who is keen to insist on a distinction between "disquotationally un-emancipated" and "emancipated," "beliefs," calling the former "quasi-beliefs." Furthermore, Recanati expands quasi-belief such that the scenarios described by Sperber are one example of a broader class of quasi-belief described as "pretending to believe" or "simulating belief" (Recanti 2000).

The "quasi-" dimension of quasi-belief, the part that prevents pretending to believe from being the same as believing, rides on a tension between some sort of credulity and the risk of inconsistency with other beliefs. We have reason to accept or commit to a quasi-belief (even in a sense that doesn't demand the deliberate willfulness of Cohen's concept of "acceptance," or the rigid contextualism of "commitment"), but something prevents it from becoming a full-fledged genuine belief. For Sperber, this credulity is derived from deference to authority, and the risk of inconsistency stems from the semantic indeterminacy of propositions we don't understand. There are other ways to arrive at this tension. There may not only be the vague possibility of inconsistency with other genuine beliefs, there may be an actual inconsistency, and even a partial interpretation of a quasi-belief may reveal this inconsistency. One may resist these inconsistencies by tenaciously clinging to a quasi-belief, or by appealing to justified trustworthiness of the authority from which they are handed down (though again, such strategies ought to strike us as epistemologically irresponsible or even unscientific). But another resource for resistance to the explicit awareness of potential and real

suggest a reduction of scientific practice to the social, and I maintain that the central role that experimental evidence plays marks the defining characteristic of a scientific method of "fixing belief."

inconsistencies between what one is “pretending to believe” and what one really believes can be found in the internal coherence and structure of a make-belief performance.

Gadamer’s treatment of play occurs in the broader context of the hermeneutic dimensions of the human sciences, where the paradigmatic context of interpretation is understanding a text. But while this context is consistent with the hermeneutic context Sperber is concerned with, his analysis is very different. The sort of “fusion of horizons” described by Gadamer’s hermeneutic attitude is not simply deference to the authority of a text or its author. (Gadamer 2004, 346-379) Rather it is willingness to let the logic of a text reveal itself. Again, the willfulness of this “willingness” is only temporarily necessary. It is true that the novice reader may have to make an explicit effort to not reject a text at the first sign of inconsistency. However, to become fully absorbed in the act of interpretation is to be drawn into a text as it unfolds, and this being drawn into is an affective response to a coherence that emerges from the interaction between reader and text, not an explicitly justified deference to the reliability of the text as an authority that trumps conflicts without previously held beliefs. Inconsistencies with other beliefs may persist, but these are ignored as we attend to a perceived emerging coherence and completeness of the text as a self contained whole.

The translation of this structure to play should be evident. Mimetic play, fiction, and make-believe present players with more than semantic indeterminacy; there is an actuality of inconsistency between fictions and our genuine beliefs. In Kendall Walton’s example of a game where tree stumps are designated as bears, we cannot genuinely believe that stumps are bears, for we have a network of other beliefs that are contrary to the proposition “Stumps are bears” and the logical and practical corollaries that follow from it. Even if we grant that positing a stump-as-bear is semantically indeterminate, these inconsistencies are bound to reveal themselves as play unfolds, and a robust game of make-believe is somehow resistant to such revelations resulting in the cessation of play. True, the novice player will have to make a concerted and willful effort to take a stump as a bear, but eventually this may give way to a passive reactivity where this resistance is not a function of deference to the authority of the game in any sense of explicit epistemological justification. Rather, it is an affective response to the game revealing itself, the enjoyment of play as such, and the complete absorption in the "seriousness" of play.

It is this structure, one that poses a tension between the internal coherence of a fiction and its inconsistency with a network of genuine beliefs about the world, which I am claiming extends to the intentional attitude we bring to fictional scientific models as well. Recall (from chapter three) that fictional models are not only in disagreement with what we believe about their targets, but additionally, at least while those fictions are at work, we don't care about their truth or falsity. Models have an internal coherence that makes them relatively autonomous, so that their users can disregard all of the ways that they deviate from and distort their targets. They cannot be completely autonomous, which is to say that they are not taken at face value, altogether independent of their representational relation to their targets; such completely autonomous models and simulations only "model" or "simulate" in a derivative sense (like the grammar parsing software that is not intended to reveal anything at all about the mechanism of the way that real natural language users parse sentences). But they represent their targets *as* something they are not, and the fit of this representational relationship is loose, and vague, and only articulated as the model plays itself out in its use, when it is taken up as a prop. We do not genuinely believe what our fictional models say about the world, but neither can we explicitly disbelieve them when they are in use. We make-believe or “quasi-believe” them.

Therefore, if the hallmark of quasi-belief is a resistance to potential or actual risk of inconsistency with genuine beliefs, then it would seem that quasi-belief captures the character of make-believe that is active in the mimetic play of working with fictional models. And if we are still wondering how it is that we are unable to make-believe while simultaneously recognizing that our attitude is *merely* playful and *not* genuine belief, it is because doing so compromises the self-contained coherence and completeness of the game. Without this, make-believe offers nothing that can resist the inevitable conflicts with real, serious life that are bound to be revealed as play runs its course. Make-believe is robust in that it allows us to ignore all of the ways in which it is not real, not genuine belief.

Of course, the inconsistency of play with our genuine, serious beliefs about the world cannot be forestalled indefinitely. Play is robust but it is also fragile, and striking the appropriate balance between the two is important in order to play in a healthy way. It was previously mentioned that the spoil-sport and the zealot are problematic players. For the former, play is too fragile, and the intrusions of serious concerns prevent play from running its course. For the latter, play is too robust, and there is the risk of taking it too seriously to enjoy it or failing to recognize when it has run its course. For the player who plays at swords with sticks, taking it too seriously may turn problematically violent. Not taking it seriously enough may mean failing to see any threat whatsoever in being “hit,” and without this, there is no play at all. For the movie-goer engaged in the make-believe of a scary film, taking it too seriously may reduce them to tears, or cause them to flee the theater. Not taking it seriously enough may mean failing to enjoy the make-believe scenario for what it is. And for the modeler, too robust a make-belief can result in confusion about whether a fictional model should be regarded as a serious hypothesis, while too fragile a make-belief can result in a premature dismissal of the hermeneutic, heuristic, and explanatory promise of a model. What drives us to reinterpret and adjust fictional models, closing the gap between them and their targets and affecting a process of discovery where fictions give rise to non-fictions, is their quasi-believability. Otherwise, when a fictional model yields surprising results, there is nothing to stop us from dismissing these results as a product of the distortions of the model.

More often than these extreme, and almost pathological play disorders, we drift in and out of play. A "sword fight" with sticks is real enough until some aspect of the real world intrudes to break the illusion. We are hit, but not hurt, and we must make a conscious, willful effort to mimic being wounded and restore the seriousness of play. The on-screen monster is real enough until the first reflexive urge to flee, at which point we realize that we are in a theater and must then make an effort to be reabsorbed into the cinematic experience. The model is quasi-real at the "bench" of the "dry lab," but not when we step away from the bench. To prevent ourselves from taking our playful fantasies too seriously, we must have a network of beliefs grounded in reality so that we do not get swept away and are able to recognize the distinctness of seriousness when it intrudes into the game. To avoid being so serious that we cannot play at all, we must possess the willingness to give the game a chance to reveal itself; but beyond this, the game itself must be compelling enough to draw us in, to be self-contained, and keep us from constantly being confronted with intrusions from the serious concerns of the real world. If, as mentioned at the outset of this essay, the use of fictional models in scientific practice and theory construction bears an instructive similarity to play and games of make-believe, then we should expect to find a similar pragmatic virtue characterized by a robust yet fragile credulity when it comes to "believing" in those fictions.

This robust yet fragile character, the way that play resists potential and/or actual conflicts with genuine beliefs but eventually breaks when the affectively compelling enjoyment of the activity runs out, is what distinguishes playful credulity from genuine belief. It is also the reason that being explicitly aware of play as *mere* play renders genuine play impossible. We may credulously assert a structural aspect of play, or we may assert that we don't really believe this to be true, either being absorbed in the playful attitude or the serious attitude respectively. To try to express both simultaneously, to say "I am playing now," requires a paradoxical attitude.

It must be conceded that the distinction between belief and make-belief is not always as sharp as I have described it. I have argued that despite their surprising similarity, and despite the inadequacy of distinguishing between them by recourse to context or the free will of the subject in suspending disbelief, make-belief differs from belief in its fragility. I explain this by identifying make-belief as a type of "quasi-belief" that is either potentially or actually in conflict with a broader, coherent network of genuine beliefs. Whether the occasion for make-belief is a game, or a fictional scientific model, it is crucial that conditions that forestall the explicit recognition of this conflict are in place so that the make-belief can play itself out. If a make-belief is not sufficiently insulated from the intrusion of the real world, it cannot build up enough momentum to engender the passive/reactive absorption of the player. When the real world does intrude in a way that makes the conflict between what is make-believed and what is genuinely believed explicit, the playful performance ruptures. This rupture that signals that shift from make-belief to belief is a cognitive event that I do not think can be adequately captured by reference to habits of action alone.

However, while I have argued that many beliefs and make-beliefs can be pinned down and represented propositionally, some may not be so sharply defined. Thus the distinction between genuine beliefs and quasi-beliefs is not so sharp, and there is a continuum of vague, unarticulated, and weakly held beliefs that are between the unambiguous cases. Conflicts involving these intermediaries are not so sharp, and consequently we must admit scenarios where the line between fantasy and reality is blurred and where shifts in attitude are less precipitous.

Nonetheless, I would resist a line of thought that devolves into a *sorites* paradox and undermines the intuition that clear cases of both genuine belief and make-belief exist. I think that these cases are interesting enough to warrant exploration without being fatally problematized by hybrids, and do my best to focus on these clear cases when examining fictional scientific models. I take it as an acceptable method of inquiry to attempt to understand such cases before proceeding to more vague, intermediary ones.

Make-Belief as a Scientific Virtue

In light of the remarks just made regarding the fragile robustness of make-believe, there are normative implications that follow from the insight that the use of fictional scientific models is characterized by an attitude of make-belief. To avoid problematic forms of make-believe, those identified in the context of game play as the “spoil-sport” and the “zealot,” the attitude of the make-believer must strike an appropriate balance between openness and resistance to intrusion. For the modeler, the virtual reality of the model must be approached in a way that is “serious but not serious,” with the right amount of credulity in the right instances and the right respects.

The ability to engage in make-believe is not only something that a scientist who uses models (especially fictional models) must be able to do, but must be able to do appropriately. In this way, make-belief is a virtue, in an Aristotelian sense of the term, for scientific practice. It concerns the appropriateness of one’s responsivity to particular situations, one which resists reduction to a general rule for when one should be credulous, and when one should be skeptical, and “appropriate” responsiveness is situated between “vicious” extremes.

The determination of what counts as virtuous or vicious make-belief, with respect to games of fantasy or in contexts of scientific modeling, is thus available for naturalistic, empirical study, with an eye to virtue epistemological philosophical analyses. Such combinations of empirical and normative study of playful fantasy have been advanced with respect to “healthy” play in children (beginning as early as Piaget 1962; Winnicott 1971). But, while I would argue that scientific make-belief shares a common genus with the play that children engage in, the broad applicability of studies of the latter to the former should be regarded with caution. Scientific make-belief and other sorts of fantasies are fairly distinct species of this common genus. Following Deanna Kuhn (1989), we would do well to note that provocative and metaphorical claims about the ways that children are intuitively scientific fall short when held to careful scrutiny, especially when it comes to the metacognitive capacities and practices of children. Children and adolescents who are not trained in scientific thinking do not exhibit much metacognitive sophistication in making distinctions between and articulating the roles of theories, models, and evidence in the formulation of their beliefs. While this metacognitive awareness is fatal to the absorbed immersion that genuine make-belief requires, it is essential for the attitude that initiates the cessation of make-belief and takes over when its characteristic absorption breaks. This skillful scientific modeler must be capable of transitioning back and forth between playful and serious attitudes (and doing so at the appropriate times), and this skill has not (to my knowledge) garnered much attention from social scientists. It is my hope that this dissertation both indicates the need for studies of this aspect of scientific practice and offers some conceptual tools for thinking about it. Ultimately, we should be strategically cultivating skillful scientific make-belief as a character trait in young scientists.

Thus there are two sides to "virtuous" modeling practices, a side viewed from the perspective of the playful context, and a side viewed from the perspective of the serious context. The latter side has been amply treated by philosophers of science and epistemologists, though I would argue that the former side has not. For example, while much ink has been spilled on the topic of how we should, in the sober light of theoretical reflection, think about models that distort their targets, little to no attention has been given to the question of what we are left with when a fictional model fulfills its scientific goal, and brings about its own obsolescence.

In chapters four and five, I touched on the way that fictional models evolve through reinterpretation, modification, and convergence with experimental systems, potentially leading to non-fictional discoveries. In light of the structure of quasi-belief, which I am claiming characterizes the attitude of the modeler when they are working with a model, a fictional model initiates a tension between its internal coherence (which allows us to become absorbed in its virtuality) and the potential and real contradictions with our network of genuine beliefs. What happens when the fictional model's evolution and reinterpretation begins to erode these contradictions? Does the "quasi-belief" become progressively more and more like a genuine belief until there is no distinction between the two? Does the context in which the model is applicable become more and more clearly delimited until the "quasi-belief" becomes a clear case of contextual commitment? Or does the model become progressively empirically adequate such that we may willfully accept it as empirically adequate but nothing more? It is likely that we can find examples that fit each of these possibilities. It is also likely that there are more ambiguous cases that depend largely on the personal metaphysical commitments, or doxological inclinations of individual scientists. The analysis given in this chapter is compatible with any of these options, but it also raises serious questions about positions, like van Fraassen's, that advocate for blanket agnosticism about the literal truth of scientific theories and models, and relegate explanatory functions of models to the realm of *mere* pragmatic virtues for attaining empirical adequacy. When virtuous scientific practices pass through an attitude of quasi-belief that is internally indistinguishable from genuine belief, such that it precludes the explicit realization that one is only pretending, only working with a *mere* model, these practices may not be compatible with the sort of agnosticism that empiricist anti-realists call for. Fictions may give way to non-fictions, and in the process, make-belief may become genuine belief.

Furthermore, the make-belief engendered by fictional models and simulations is not simply a function of the responsivity of the human users of these models. It also depends upon the model itself. In part this can be cashed out in terms of a sort of abstract "internal coherence," but the impact of supposedly innocuous or accidental design decisions about simulational systems should not be underestimated. Early examples of computer simulation, like those used in the "Manhattan Project," hardly invited users to become immersed in a virtual reality. But today, the vividness and virtuality of models is significantly more compelling. As Sherry Turkle (2009) notes, if we ask what simulations "want," the answer is that they "want, even demand, immersion" (Turkle 2009, 6). In the paradigm shift from wet-lab experimentation to dry-lab simulation she examines throughout a wide variety of science, technology, engineering and mathematical disciplines, it is clear that simulation is not only getting what it wants, but has also initiated a generational gap between the old guard of scientists who came up through a wet-lab paradigm, and younger scientists who are coming up through a dry-lab paradigm. One gets the impression that the old guard's resistance and skepticism about simulational practices is a bit overblown, and exhibits many of the familiar signs of techno-phobic romanticism. But at the same time, many of their anxieties are warranted, and the new guard might just be too credulous in their virtual laboratory practices. While immersion is a beneficial and necessary component for the sort of make-belief that modeling involves, *virtuous* modeling practices require that this immersion be appropriately fragile.

Thus, in addition to the normative implications about the sorts of character traits that should be studied and cultivated in young scientists, we should be addressing the need for virtuous scientific make-belief in the design of scientific simulations. Design decisions that facilitate and interrupt immersion are not simply aesthetic window dressing, they are intimately linked to the epistemological functions of scientific models and simulations. I hope that this dissertation has taken some steps in articulating this link, and indicates some directions that can be taken to strategically design models and simulations with an eye to these connections between the ways that they invite embodied immersion, the ways that embodied immersion is disrupted, and the ways they produce scientific knowledge.

In addition to the character traits of model users and the design of models and simulations, the technological "lifeworld" beyond the laboratory has an influence on whether and how virtuous modeling practices develop. This is, in part, related to the previously discussed factors in that the human-technology relations that get implemented in the laboratory are frequently developed outside of the laboratory. The inclination and ability of modelers to become immersed in the make-believe virtuality of models, the capacity for model/simulation design to support this immersion, as well as the ability for this immersion to be ruptured where appropriate, develop from the skillful embodied relationships between humans and technologies in everyday life (*e.g.* the use of interactive screen technologies) that are transferable to laboratory contexts.

An illustrative example, not only of this consonance between non-scientific interactive technologies and scientific ones but also of the striking success of bringing game-inspired design that facilitates immersion is the online protein folding game, "Foldit." Foldit, collaboratively designed by the University of Washington's Center for Game Science and Department of Biochemistry, "allows players to fiddle at folding proteins on their home computers in search of the best-scoring (lowest energy) configurations." (Marshall 2012) Due not only to the "crowd-sourcing" of protein structure problems, but also compelling game-play that encourages players to invest incredible time and mental energy (Garylewski 2008), Foldit players have been contributing members of the international project, Critical Assessment of Techniques for Protein Structure Prediction. In 2011, Foldit players solved the crystal structure of M-PMV monomeric retroviral protease, a key protein in an AIDS virus that affects monkeys that had stumped researchers for 15 years. Foldit players solved the enzyme structure in 10 days. (Khatib *et. al.* 2011; FOLDIT) More recently, Foldit players were able to design an enzyme that catalyzes Diels-Alder reactions. The new catalyst is 18 times more active than the most active available catalysts. Foldit players are currently looking into tackling designs for improving small protein inhibitors that would block pandemic flu strains (Stomp 2012; FOLDIT)

Lastly, and more tentatively, the distinctions between the artifice of simulations, the artifice of the laboratory, and the world beyond the laboratory are fast becoming increasingly blurred. That there is a continuum between the vague context of make-belief and the sharply delimited context of commitment, one which gets filled in as fictional models and laboratory apparatus converge, has already been mentioned. As we find new ways to manipulate real phenomena, we can create "real" engineered scenarios that are within reach of what previously fictional models can be reinterpreted and adjusted to fit (*cf* the "endless" cylinder from chapter four). The same sort of evolving continuum exists between the special circumstances that we are able to create in a carefully controlled and distorted laboratory environment and the world we encounter in our everyday lives. Cartwright's major objection to the truth of theoretical laws is that they do not apply universally, but instead are only even approximately true in a small handful of carefully engineered (and some outright fictional) scenarios. In the terms set out above, regarding the conditions that obstruct genuine belief, the narrowness of the context of applicability limits our attitude toward such theoretical laws to contextual commitment at best. But, today's exotic phenomenon, cultivated in the laboratory, is frequently tomorrow's technology, exported out of the laboratory and into the "lifeworld" of the non-scientist (or scientist away from the bench). This is not to make dramatic (or alarmist) claims about the disappearance of the natural world, or the advent of a Baudrillardian simulacrum. But if the difference between commitment and belief is that commitments are restricted to specific contexts while beliefs are not, then the distinction between the two weakens as the applicable contexts of a commitment become more and more ubiquitous, and this softens the boundaries between theoretical scientific principles that only strictly apply in select controlled cases, and those that fit most of the cases we are likely to encounter.

This has implications for the justifiability of genuine belief in scientific theories — implications that deserve exploration in their own right, but fall outside of the focus of this dissertation. But issues of justifiability aside, it would seem that we are on a trajectory that makes credulous immersion in models and simulations easier, both as the models become more experientially similar to the world we encounter in our everyday lives, and as the world we encounter in our everyday lives becomes more experientially similar to our models. And if one of the primary normative concerns about how we use scientific models and simulations is striking a virtuous balance between credulous immersion and a skeptical withholding of genuine belief, then it is important to recognize that the material, social, and existential conditions that are constitutive of the attitudes we naively bring to fictional scientific modeling are shifting underneath our feet.

Bibliography

Ackerlof, George A (1970), “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism”, *Quarterly Journal of Economics* 84: 488-500.

Ankeny, Rachel (2007) “Wormy Logic: Model Organisms as Case-Based Reasoning” in *Science Without Laws: Model systems, cases, and exemplary narratives.* Eds. Creager, A., E. Lunbeck, and M. N. Wise. Durham, Duke University Press.

Ankeny, Rachel (2009) “Model Organisms as Fictions” in *Fictions in Science: Philosophical Essays on Modeling and Idealization*. Ed. Mauricio Suarez. New York, Routledge. 193-204.

Ariew, Roger, "Pierre Duhem", *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Ed. Edward N. Zalta. http://plato.stanford.edu/archives/spr2011/entries/duhem/.(Accessed 4/16/12)

Ariew, Roger and Peter Barker, 1986, “Duhem on Maxwell: a Case-Study in the Interrelations of History of Science and Philosophy of Science,” Philosophy of Science Association, 1: 145–56.

Axelrod, R. and W. Hamilton (1981) “The Evolution of Cooperation” *Science* 211: 1390-1396.

Ayer, A. J. (1936) *Language*, *Truth and Logic*. Reprint, 1983. New York, Penguin.

Barberousse, A. and Pascal Ludwig (2009) “Models as Fictions” in *Fictions in Science: Essays on Modeling and Idealization*. Ed. Mauricio Suarez. New York, Routledge.

Bateson, Gregory. (1956) “This is Play.” in *Group Processes*, ed. B Schaffner. New York, Josiah Macy.

Batterman, Robert (2002) “Asymptotics and the Role of Minimal Models.” *British Journal for the Philosophy of Science*. LIII: 21-38.

Bell, J.S. (1964) “On the Einstein-Podolsky-Rosen paradox,” *Physics*, 1: 195–200.

Bohm, D. (1957) *Causality & Chance in Modern Physics*. Philadelphia: University of Pennsylvania Press.

Boyd, Richard (1983) “On the Current Status of the Issue of Scientific Realism” *Erkenntnis*, 19: 45–90.

Boyd, Richard (2002) "Scientific Realism" *The Stanford Encyclopedia of Philosophy* (Summer 2010 Edition) http://plato.stanford.edu/archives/sum2010/entries/scientific-realism/ (Accessed 4/16/12)

Brassard, G., R. Cleve and A. Tapp, (1999) "Cost of exactly simulating quantum entanglement with classical communication", *Phys. Rev. Lett*., 83(9): 1874–1877.

Bratman, Michael (1999), *Faces of intention.* Cambridge, Cambridge University Press.

Brown, H. (1986) "The insolubility proof of the quantum measurement problem" *Foundations of Physics*, 16, 857-70.

Bruch E., and R. D. Mare (2006) ''Neighborhood choice and neighborhood change''. Available online at: http://repositories.cdlib.org/ccpr/olwp/CCPR-013–05 (Accessed 4/16/12)

Callois, Roger. (2001) *Man Play and Games*. (originally published as *Le jeux et les homes* 1958, first English translation 1961) Urbana, University of Illinois Press.

Carnap, Rudoph (1928) *The Logical Structure of the World and Pseudoproblems in Philosohy*. Reprint, 1967. Berkeley, University of California Press.

Carnap, Rudolph (1956) "The methodological character of theoretical concepts." in *Minnesota Studies in the Philosophy of Science, vol. I: The foundations of science and the concepts of psychology and psychoanalysis*. eds H. Feigl and M. Scriven. Minneapolis, University of Minnesota Press, 38-76.

Cartwright, N. (1983) *How the Laws of Physics Lie*. Oxford, Oxford University Press.

Cartwright, Nancy (1989) *Nature's Capacities and their Measurement*. Oxford, Oxford Univeristy Press, 183-224.

CCML, Cornel Creative Machines Lab website for *Eureqa* sofware, http://creativemachines.cornell.edu/eureqa (Accessed 4/16/12)

Chipman, H. A., George, E. I. and McCulloch, R. E. (2001) "Managing Multiple Models" *Artificial Intelligence and Statistics*. Eds. Tommi Jaakkola, Thomas Richardson. 11-18. Available online at http://math.acadiau.ca/chipmanh/papers/aistat2001-paper.pdf (Accessed 5/4/12)

Cohen, Jonathan (1992) *An Essay on Belief and Acceptance*. Oxford, Oxford Univ. Press.

Cohen, M.R. (1923). "On the Logic of Fictions." *Journal of Philosophy*, 20:477-488.

Collins, Harry M. (1981) “Stages in the Empirical Programme of Relativism” Social Studies of Science. 11(1): 3-10.

Conway, J., R. Guy and E. Berlekamp (1982) *Winning Ways for Your Mathematical Plays*, Vol. 2, London:Academic Press, 927–963.

Cooper, S. *et al.* (2010) “Predicting protein structures with a multiplayer online game” *Nature* 466: 756-760

Creager, A., E. Lunbeck, and M. N. Wise (2007) “Introduction” *Science Without Laws: Model systems, cases, and exemplary narratives.* Durham, Duke University Press.

Crease, Robert P. (1993) *The Play of Nature*. Bloomington, Indiana University Press.

Crease, Robert P. (2006) “From Workbench to Cyberstage” in *Postphenomenology: A Critical Companion to Ihde*. Ed. Evan Selinger. Albany State University of New York Press.

Crease, Robert P. (2010) *The Great Equations: Breakthroughs in Science from Pythagoras to Heisenberg*. New York, Norton.

Da Costa, Newton, and Steven French (2003) *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning.* Oxford, Oxford University Press.

DaSilva,E. (2011) “Eureqa! Signs of the Singularity?” *h+ Magazine*. http://hplusmagazine.com/2011/03/25/eureqa-signs-of-the-singularity/ (accessed 1/8/12)

Douglas, Heather (2000) “Inductive Risk and Values in Science” *Philosophy of Science*. 67(4): 559-579.

Earman, J. and A. Shimony (1968) ‘A Note on Measurement’, *Nuovo Cimento*, 54B, 332-334.

Eason *et. al.* (2007) “What sort of Science is Simulation” *Journal of Experimental & Theoretical Artificial Intelligence.* 19(1): 19–28.

Einsten, A., B. Podolsky, and N. Rosen (1935) "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review*. 47: 777-780.

Elgin, Catherine (2009) “Exemplification, Idealization, and Scientific Understanding” in *Fictions in Science; Philosophical Essays on Modeling and Idealization*. Ed. Mauricio Suarez. New York; Routledge.

Engber, Daniel (2011) “The Mouse Trap.” *Slate*. November 16, 2011. Available at http://www.slate.com/articles/health_and_science/the_mouse_trap/2011/11/lab_mice_are_they_limiting_our_understanding_of_human_disease_.html (last accessed: 2/22/12)

Engel, Pascal. (2000) “Introduction: The Varieties of Belief and Acceptance” in *Believing and Accepting*. Ed. Pascal Engel. Dordrecht, Kluwer, 1-30.

Feigl, Herbert (1956) “Some Major Issues and Developments in the Philosophy of Science of Logical Empiricism.” in *Minnesota Studies in the Philosophy of Science, vol. I: The foundations of science and the concepts of psychology and psychoanalysis*. eds H. Feigl and M. Scriven. Minneapolis, University of Minnesota Press, 3-37.

Feyerabend, Paul (1962) “Explanation, Reduction and Empiricism”, in *Minnesota Studies in the Philosophy of Science, Volume III: Scientific Explanation, Space, and Time.* eds. H. Feigl and G. Maxwell. Minneapolis, University of Minneapolis Press, 28–97.

Feyerabend, Paul (1975) *Against Method*. Verso, London.

Fine, Arthur (1970) “Insolubility of the Quantum Measurement Problem” *Physical Review* D, 2, 2783-87.

Fine, Arthur (1984a) “The Natural Ontological Attitude” in *Scientific Realism*, Berkeley: University of California Press, 83-107.

Fine, Arthur (1984b) “And Not Anti-Realism Either” *Nous* 18: 51-65.

Fine, Arthur (1993). “Fictionalism.” *Midwest Studies in Philosophy*, XVIII, 1-18.

FOLDIT. available at http://fold.it/portal/ (Accessed 4/16/12)

Fox-Keller, E. (2003) “Models, Simulations, and ‘Computer Experiments’” in *The Philosophy of Scientific Experimentation* ed. Hans Radder. Pittsburgh: University of Pittsburgh Press.

Frank, Philipp. (1949) *Modern Science and its Philosophy*. Cambridge, Harvard University Press.

Frigg, Roman and Hartmann, Stephan (2012) "Models in Science", *The Stanford Encyclopedia of Philosophy* (Spring 2012 Edition), Edward N. Zalta (ed.), forthcoming http://plato.stanford.edu/archives/spr2012/entries/models-science/. (Accessed 4/16/12)

Frigg, Roman, and Julian Reiss (2009) “The philosophy of simulation: hot new issues or same old stew.” *Synthese*. 169(3): 593–613.

Gadamer, Hans Georg. (2004) *Truth And Method.* Second revised edition. (originally published as *Warheit und Methode* 1960, first English translation 1975) New York, Continuum Publishing.

Galison, P. (1996) *Image and Logic: A Material Culture of Microphysics*. Chicago: University of Chicago Press, 689-776.

Garylewski, Andrea (2008) “Foldit for fun” *The Scientist* News and Opinion. Available at http://classic.the-scientist.com/blog/display/54677/ (Accessed 4/16/12)

Gaster, M. (1969) “Vortex shedding from slender cones at low Reynolds numbers.” *J. Fluid Mech.* 38: 365.

Gibbard, Allan and Hal Varian (1978), “Economic Models”, *Journal of Philosophy* 75: 664-677.

Giere, Ronald (2006) *Scientific Perspectivism*. Chicago, University of Chicago Press.

Giere, Ronald (2009). “Why Scientific Models Should Not Be Regarded As Works of Fiction” in *Fictions in Science: Essays on Modeling and Idealization*. Ed. Mauricio Suarez. New York, Routledge.

Gigerenzer, G., and P. Todd (2000) *Simple Heuristics that Make us Smart.* Oxford, Oxford University Press.

Gilbert, Nigel and Klaus Troitzsch (1999) *Simulation for the Social Scientist.* London, Open University Press.

Ginsburg, Herbert and Sylvia Opper. (1969) *Piaget’s theory of intellectual development*: *an introduction*. Englewood Cliffs, Prentice Hall.

Godfrey-Smith, Peter (2006) “The Strategy of Model Based Science” *Biology and Philosophy*. 21: 725-740.

Goodman, Nelson (1968) *Languages of Art; an approach to a theory of symbols.* Indianapolis, Hackett.

Goodman, Nelson (1978) *Ways of World Making.* Indianapolis, Hackett Publishing.

Grim et.al. (forthcoming) “How Simulations Fail” *Synthese*. — Available online at http://www.springerlink.com/content/6746n65516j54433/ (Accessed 4/16/12)

Guala, Francesco (2002) “Models, Simulations, and Experiments,” in *Model-Based Reasoning: Science, Technology, Values*. Eds. Lorenzo Magnani and Nancy Nersessian. New York: Kluwer, 59-74.

Hacking, Ian (1983) *Representating and Intervening: Introductory topics in the Philosophy of Natural Science*. Cambridge, Cambridge University Press.

Hacking, Ian (1991) “A Tradition of Natural Kinds” *Philosophical Studies*. 61(1): 109-126.

Hanson, N.R. (1958) *Patterns of Discovery*, Cambridge, Cambridge University Press.

Hartmann, Stephen (1998) “Idealization in Quantum Field Theory” in *Idealization In Contemporary Physics*. Ed. N. Shanks. Amsterdam, Rodopi, 99-122.

Hempel C. and P. Oppenheim Hempel (1948). "Studies in the Logic of Explanation". *Philosophy of Science.* 15: 135–175.

Hesse, Mary (1966) *Models and Analogies in Science*. Notre Dame, University of Notre Dame Press.

Hoyningen-Huene, Paul. (1993) *Reconstructing Scientific Revolutions: The Philosophy of Science of Thomas S. Kuhn*. Chicago, University of Chicago Press.

Hubbard, E. J. A. (2007) “Model Organisms as Powerful Tools for Research” in *Science Without Laws: Model systems, cases, and exemplary narratives*.Eds. Creager, A., E. Lunbeck, and M. N. Wise. Durham, Duke University Press.

Huggins, E. M., & Schultz, E. A. (1967). “San Francisco Bay in a warehouse” *Journal of the Institute of Environmental Sciences and Technology*, 10(5), 9–16.

Huggins, E. M., & Schultz, E. A. (1973). “The San Francisco bay and delta model” *California Engineer*, 51(3), 11–23
http://www.spn.usace.army.mil/bmvc/bmjourney/the_model/history.html (Accessed 4/16/12)

Huizinga, Johan. (1950) *Home Ludens*. Boston, Beacon Press.

Hume, David. (1993 [1777]) *An Enquiry Concerning Human Understanding* 2nd. Ed. Indianapolis, Hackett.

Humphreys, P. (1990) ‘‘Computer simulations’’, *Philosophy of Science Association*. 2: 497–506.

Humphreys, P. (2004) *Computational Science, Empiricism, and Scientific Method.* Oxford, Oxford University Press.

Husserl, Edmund (1999) *The Essential Husserl*. Ed. Donn Welton. Bloomington, Indiana University Press.

Ihde, Don (2008) “The Designer Fallacy” in *Ironic Technics.* Automatic Press.

Ihde, Don (2012) *Experimental Phenomenology: Multistabilities*. 2nd ed. Albany, State University of New York Press.

Interagency Performance Evaluation Task Force. (2006). *Performance evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System: Draft final report of the Interagency Performance Evaluation Task Force*, volume 1. www.asce.org/files/pdf/executivesummary_v20i.pdf (Accessded 4/16/12)

Karniadakis, G. E. and G. S. Triantyfalou (1989) “Frequency Selection and Asymptotic States in Laminar Wakes.” *J. Fluid Mech.* 199:441.

Khatib, F. *et al. (*2011) “Crystal structure of a monomeric retroviral protease solved by protein folding game players” *Nature Structural & Molecular Biology*. 18: 1175–1177.

Kitcher, P. (1989) ‘Explanatory Unification and the Causal Structure of the World’, in *Scientific Explanation*, eds. P. Kitcher and W. Salmon. Minneapolis, University of Minnesota Press, 410–505.

Kline, S. (1985) “What is Technology” in *Philosophy of Technology:The technological condition*. Eds. R. Scharf and V. Dusek (2003). Blackwell Publishing, 210-212.

Knuuttila, Tarja (2006) “From Representation to Production” in *Simulation: Pragmatic Construction of Reality*. Eds. Johannes Lenhard, Gunter Kuppers, and Terry Shinn. Springer, Dordrecht.

Kuhn, Thomas (1962) *The Structure of Scientific Revolutions*. (Third ed. 1996) Chicago, University of Chicago Press.

Kuhn, Thomas (1983) 'Rationality and Theory Choice', *Journal of Philosophy*. 80: 563-570.

Kuhn, Deanna (1989) “Children and Adults as Intuitive Scientists” *Psychological Review* 96(4): 674-689.

Latour, Bruno and Steve Woolgar (1979) *Laboratory Life: The Construction of Scientific Facts*. Princeton, Princeton University Press.

Latour, Bruno (2007) *Reassembling the Social: An introduction to actor-network theory*. Oxford, Oxford University Press.

Levins, Richard (1966) "The Strategy of Model Building in Population Biology." *American Scientist*. 54: 421-31.

Li, David and Dorfman, Jeffrey (1996) "Predicting Turning Points Through Integration of Multiple Models" *Journal of Business and Economic Statistics*. 14(4): 421-428.

LLNL (1999) "Duplicating the Plasmas of Distant Stars" Lawrence Livermore National Laboratory - S&TR. April 1999. Available at https://www.llnl.gov/str/Springer.html (Accessed 4/16/12)

Lutz, Sebastian (2010) "What's Right with a Syntactic Approach to Theoreies and Models?" Available online at http://philsci-archive.pitt.edu/5264/1/whats_right_with_syntactic_view.pdf (Accessed 4/16/12)

Mäki, Uskali (1994) "Isolation, Idealization and Truth in Economics", in *Idealization VI: Idealization in Economics. Eds.* Bert Hamminga and Neil B. De Marchi. *Poznan Studies in the Philosophy of the Sciences and the Humanities*, Vol. 38: 147-168. Amsterdam: Rodopi.

Marshall, Jessica (2012) "Victory for crowdsourced biomolecule design" *Nature* News and Opinion. Available at http://www.nature.com/news/victory-for-crowdsourced-biomolecule-design-1.9872 (Accessed 4/16/12)

Maxwell, J. C. (1861) "On Lines of Physical Force" Originally published in *Philosophical Magazine and Journal of Science.* Available online at http://www.vacuum-physics.com/Maxwell/maxwell_oplf.pdf (Accessed 10/2/2011)

Maxwell, G., (1962) 'On the Ontological Status of Theoretical Entities', in *Minnesota Studies in the Philosophy of Science, Volume III: Scientific Explanation, Space, and Time*. eds. H. Feigl & G. Maxwell. Minneapolis, University of Minnesota Press.

McMullin, Ernan (1985) "Galilean Idealization" *Studies in the History and Philosophy of Science* 16: 247-73.

Moore, G. E. (1993) *G.E. Moore: Selected Writings*. Ed. Thomas Baldwin. New York, Routledge.

Morgan, Mary (2003) "Experiments Without Material Intervention: Model Experiments, Virtual Experiments and Virtually Experiments," in *The Philosophy of Scientific Experimentation*.Ed. Hans Radder. Pittsburgh, University of Pittsburgh Press, 216-235.

Morgan, Mary, and Margaret Marrisson (1999) "Introduction" & "Models as Mediating Instruments" in *Models as Mediators. Perspectives on Natural and Social Science.* Cambridge, Cambridge University Press.

Morrisson, Margaret (1999) "Models as Autonomous Agents" in *Models as Mediators: Perspectives on natural and social science.* Eds. M. Morgan and M. Morrisson. Cambridge, Cambridge University Press.

Musgrave, Alan (1981) "'Unreal Assumptions' in Economic Theory: The F-Twist Untwisted" *Kyklos* 34: 377-387.

Nagel, Ernest (1961) *The Structure of Science: Problems in the Logic of Scientific Explanation*. New York: Harcourt, Brace and World.

Norgard, Richard (2004) "Environmental Economics: An evolutionary critique and a plea for pluralism" *Journal of Environmental Economics and Management*. 12(4): 382-394.

Online Etymology Dictionary: "Heuristic" Available at http://etymonline.com/?term=heuristic (Accessed 4/16/12)

Parker, Wendy (2009) "Does Matter Really Matter? Computer Simulations, Experiments, and Materiality." *Synthese* 169 (3):483 - 496.

Passmore, John (1967). "Logical Positivism" in *The Encyclopedia of Philosophy* (Vol. 5, 52-57). ed. P. Edwards. New York: Macmillan.

Peirce, C.S. (1877) "The Fixation of Belief" in *The Essential Peirce: Vol. 1.* ed. N. Houser and C. Kloesel. Blomington, Indiana University Press, 109-123.

Peirce, C.S. (1878) "How to Make Our Ideas Clear" in *The Essential Peirce* Vol. 1. ed. N. Houser and C. Kloesel. Blomington, Indiana University Press, 124-141.

Peirce, C.S. (1894) "What is a Sign" in *The Essential Perice, vol 2 (1893-1913).* ed. The Peirce Edition Project. Indiana University Press, Bloomington.

Peschard, Isabelle (2011) "Modeling and Experimenting" in *Models, Simulations, and Representations*. Eds. Paul Humphreys & Cyrille Imbert. New York, Routledge.

Piaget, Jean. (1962) *Play, Dreams, and Imitation in Childhood*. (originally published as *La Formation du symbole chez l'Enfant* 1946, first English translation 1951) New York, Norton.

Pickering, Andrew (1995) *The Mangle of Practice: Time, Agency, & Science*. Chicago, University of Chicago Press.

Putnam, Hilary. (1975) *Mathematics, Matter and Method: Philosophical Papers*, vol. 1. London: Cambridge University Press, 73.

Putnam, Hilary (1982) "Why There Isn't a Ready-Made World." *Synthese* 51 (2): 205-228.

Quine, W.V.O. (1951) "Two Dogmas of Empiricism" *The Philosophical Review* 60: 20-43.

Recanati, François. (2000) "The Simulation of Belief" in *Believing and Accepting*. Ed. Pascal Engel. Dordrecht, Kluwer, 267-298.

Reiss, Julian (2006) "Beyond Capacities", in, *Nancy Cartwright's Philosophy of Science*. Luc Bovens and Stephan Hartmann (eds.)London, Routledge.

Reynolds, C. W. (1987) ''Flocks, herds, and schools: a distributed behavioral model'', *Comput. Graphics*, 21: 15–34.

Roshko, A. (1954) "On the development of turbulent wakes from vortex streets." Report National Advisory Committee for Aeronautics, No 1191. Available at http://authors.library.caltech.edu/428/ (Accessed 4/16/12)

Rouse, Joseph (1987) *Knowledge and Power: Toward a Political Philosophy of Science*. Ithaca, Cornell University Press.

Rouse, Joseph (2009) "Laboratory Fictions" in *Fictions in Science: Philosophical Essays on Modeling and Idealization*. Ed. Mauricio Suarez. New York, Routledge, 37-55.

Rudner, Richard (1953) "The Scientist *Qua* Scientist Makes value Judgments" *Philosophy of Science*. 20(1): 1-6.

Russell, Bertrand (1948). *Human knowledge, its scope and limits*. New York: Simon and Schuster.

Salmon, W. (1971) *Statistical Explanation and Satistical Relevance*, Pittsburgh: University of Pittsburgh Press.

Salmon, W. (1984) *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.

Sankey, H. (1993) “Kuhn's Changing Concept of Incommensurability”, *British Journal for the Philosophy of Science*, 44: 759–774.

Scheffler, Israel (1967) *Science and Subjectivity*. Indianapolis, Hackett.

Schelling, Thomas C. (1971) "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1:143-186.

Schelling, Thomas (1978) *Micromotives and Macrobehavior*, Norton, New York.

Schlick, Moritz (1932) “Positivism and Realism” in *Logical Positivism*. ed. A. J. Ayer. The Free Press, New York.

Schlombs, Corinna (2010) “A Gendered Job Carousel: Employment effects of computer automation” in *Gender Codes: Why women are leaving computing.* Ed. Thomas J. Misa.Hoboken, Wiley & Sons.

Schroedinger, Erwin (1980 [1935]) “The present situation in quantum mechanics” Transl. John D. Trimmer. *Proceedings of the American Philosophical Society*, 124, 323-38.

Smith, J. M. (1995) "Life at the Edge of Chaos?," The New York Review of Books, March 2, 1995. available at http://www.nybooks.com/articles/archives/1995/mar/02/life-at-the-edge-of-chaos/ (Accessed 4/16/12)

Spariosu, Mihai. (1989) *Dionysus Reborn*. Ithaca, Cornell Univ. Press.

Sperber, Dan. (1975) *Rethinking Symbolism*. Cambridge Cambridge Univ. Press.

Sperber, Dan (1985) “Apparently Irrational Beliefs” in *Anthropological Knowledge*. Cambridge, Cambridge Univ. Press.

Sperber, Dan (2000) “Concepts, Beliefs, and Metarepresentations” in *Believing and Accepting*. Ed. Pascal Engel. Dordrecht, Kluwer, 243-266.

Stein, H. (1997) “Maximal Extension of an Impossibility Theorem Concerning Quantum Measurement”, in *Potentiality, Entanglement and Passion at a Distance*, R. Cohen and J. Stachel, (*eds*.), Dordrecht, Kluwer Academic Publishers.

Steiner, M. (1999) “Towards quantifying non-local information transfer: finite-bit non-locality”, *Proc. of 29th Winter Colloquium on the Physics of Quantum*

Electronics, Snowbird, Utah, 1999. Available at http://arxiv.org/abs/quant-ph/9902014 (Accessed 4/16/12)

Sterrett, Susan (2001) Physical Models and Fundamental Laws: Using One Piece of the World to Tell About Another." Available at http://philsci-archive.pitt.edu/720/ (Accessed 4/16/12)

Sterrett, Susan (2005a) "Models of Machines, Models of Phenomena" available at http://philsci-archive.pitt.edu/2245/ (Accessed 4/16/12)

Sterrett, Susan (2005b) "Kinds of Models" available at http://philsci-archive.pitt.edu/2363/ (Accessed 4/16/12)

Sterrett, Susan (2005c) *Wittgenstein Flies a Kite*. New York, Pi Press.

Steuer, R.H. (1985) "Artificial Distintegration and the Cambridge-Vienna Controversy," in *Observation, Experiment, and Hypothesis in Modern Physical Science*. eds. P. Achinstein and O. Hannaway. Cambridge, MIT Press, 239–307.

Stomp, W. (2012) "Another Success for Foldit: Gamers increase enzyme activity by a factor of 18" *MedGadget*. Available at http://medgadget.com/2012/01/another-success-for-foldit-gamers-increase-enzyme-activity-by-a-factor-of-18.html (Accessed 4/16/12)

Sutton-Smith, Brian. (1997) *The Ambiguity of Play*. Cambridge, Harvard Univ. Press.

Strevens, Michael (2007) "Why Explanations Lie" unpublished draft, available online at http://www.strevens.org/research/expln/Idealization.pdf (Accessed 4/16/12)

Suárez, Mauricio (2003), "Scientific Representation: Against Similarity and Isomorphism." *International Studies in the Philosophy of Science* 17: 225-244.

Suárez, Mauricio (2004) "Quantum selection propensities and the problem of measurement" in *British Journal for the Philosophy of Science* 55 (2).

Suárez, Mauricio (2009a) "Fictions in Scientific Practice" in *Fictions in Science: Essays on Modeling and Idealization*. Ed. Mauricio Suarez. New York, Routledge.

Suárez, Mauricio (2009b) "Scientific Fictions as Rules of Inference" in *Fictions in Science: Essays on Modeling and Idealization*. Ed. Mauricio Suarez. New York, Routledge.

Suppe, Frederick. (1989), *The Semantic View of Theories and Scientific Realism*. Urbana and Chicago: University of Illinois Press.

Suppes, Patrick. (1967) “What is a scientific theory?” in *Philosophy of Science Today*, ed. Sidney Morgenbesser, New York: Basic Books Inc., 55–67.

Suppes, Patrick (2002) *Representation and Invariance of Scientific Structures*. Stanford: CSLI Publications.

Tritton, D. J. (1959) “Experiments on the flow past a circular cylinder at low Reynolds numbers.” *J. Fluid Mech.* 6: 547-567.

Tritton, D. J. (1971) “A note on vortex streets behind circular cylinders at low Reynolds numbers.” *J. Fluid Mech.* 45: 203-208.

Turkle, Sherry (2009) *Simulation and its Discontents*. Cambridge, MIT Press.

Turner, Victor. (1969) *The Ritual Process: Structure and Anti-structure*. New York, Aldine.

USNWS (US National Weather Service) http://www.meted.ucar.edu/nwp/pcu2 (last accessed 4/13/12)

Vaihinger, H. (1924) *The Philosophy of ‘As If’: A system of the theoretical, practical and religions fictions of mankind.* Transl. C.K. Ogden. London, Kegan Paul.

Van Atta, C. W. and M. Garib (1987) “Ordered and chaotic vortex streets behind circular cylinders at low Reynolds numbers.” *J. Fluid Mech.* 174: 113-133.

van Fraassen, Bas C. (1980) *The Scientific Image.* Oxford, Oxford University Press.

van Fraassen, Bas C. (2000) “The semantic approach to scientific theories.” in *The Nature of Scientific Theory*, ed. Lawrence Sklar, vol. 2 of *Philosophy of Science*. New York: Garland Publishing Inc.175–194.

Walton, Kendall (1990) *Mimesis as Make-Believe: On the foundations of the representational Arts*. Cambridge, Harvard University Press.

Warwick, A. (2003) *Masters of Theory: Cambridge and the rise of mathematical physics.* Chicago, University of Chicago Press.

Weber, Marcel (2007) “Redesigning the Fruit Fly: The molecularization of *Drosophila*” in *Science Without Laws: Model systems, cases, and exemplary narratives.* Eds. Creager, A., E. Lunbeck, and M. N. Wise. Durham, Duke University Press.

Weinberger, D. (2012) *Too Big to Know: Rethinking Knowledge Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room*. New York, Basic Books.

Williamson, C. H. K. (1989) “Oblique and parallel modes of vortex shedding in the wake of a circular cylinder at low Reynolds number.” *J. Fluid Mech.* 206: 579-627.

Winnicott, D. W. (1971) *Playing and Reality*. London, Tavistock Publications.

Winsberg, Eric (2001) Simulations, models, and theories: Complex physical systems and their representations. *Philosophy of Science*, 68, S442–S454.

Winsberg, Eric (2003) “Simulated experiments: Methodology for a virtual world.” *Philosophy of Science*, 70,105–125.Weisberg, Michael (2007) "Three Kinds of Idealization," *The Journal of Philosophy*, 104 (12): 639-59.

Winsberg, Eric (2006) “Handshaking Your Way to the Top: Simulation at the nanoscale.” *Philosophy of Science*, *73*, 582-594.

Winsberg, Eric (2009). “A Tale of Two Methods” *Synthese*, 169 (3): 575-92.

Winsberg, Eric (2009) “A Function for Fictions: Expanding the scope of science” in *Fictions in Science: Essays on Modeling and Idealization*. Ed. Mauricio Suarez. New York, Routledge.

Winsberg, Eric (2010) *Science in the Age of Computer Simulation*. Chicago, University of Chicago Press.Wigner, E. [1963]: “The Problem of Measurement”, *American Journal of Physics* 31, 6-15.

Wray, K. B. (2007) ‘A Selectionist Explanation of the Success and Failures of Science’, *Erkenntnis*, 67: 81–89.