**UNIVERSITY OF CALGARY**

# Experiments, Thought Experiments and Virtual Experiments

**by**

Kimberley Brumble

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

DEPARTMENT OF PHILOSOPHY

CALGARY, ALBERTA

APRIL, 2011

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Experiments, Thought Experiments and Virtual Experiments" submitted by Kimberly Brumble in partial fulfillment of the requirements for the degree of Master of Arts.

Supervisor, Dr. John A. Baker, Department of Philosophy

Dr. Marc Ereshefsky, Department of Philosophy

Dr. Henderikus Stam, Department of Psychology

April 18th 2011

*Date*

Acknowledgements

I owe thanks in many places for the successful completion of this thesis.

First, thank you Professor John Baker for all your help with this project and for your faith in its author. It has been a delight to work with you. Thanks also to Professor Marc Ereshefsky for reading and discussing bits of this work and for suggesting helpful examples. Thanks to Denise Retzlaff for many hours of assistance throughout my program and an abundance of moral support. Thank you Makmiller Pedroso for your insights into this project and your inestimable friendship.

My parents, Kathy and Stan Smith- thank you for all kinds of support, love, and unwavering faith in my success. Thank you also Joseph Fischer- you are a great friend and wonderfully supportive.

**Abstract**

In the past there has been a lot of skepticism about a variety of experiments which utilize virtual materials, rather than the materials of the phenomena under investigation, or the *target materials*. I will argue this in Chapter 1, and also that this skepticism has been of the same ilk as the skepticism about whether thought experiments are genuinely experimental. Additionally, the same variety of skepticism exists as to whether models, simulations, and experiments with theory-laden apparatuses constitute robust experimentation. In this thesis I will describe the existence of a spectrum of virtual experiments, upon which all of the above forms of experimentation exist, and in which they differ only in degree, rather than in kind, from one another in terms of success conditions. I will argue that these success conditions exist relative to the quality of the idealization involved in experimental design. In Chapter 2 I will argue that thought experiments, being the most virtual of virtual experiments, have much to teach us about robust idealization, and that the key to understanding thought experiments as virtual experiments is through their commitment to underlying "conceivability to possibility" inferences (where "possibility" can refer to a set of distinct types of possibility). In chapter 3 I will argue that as a result of these inferences, different virtual experiments often achieve different levels of probative force, and that as a distinction between heuristic and probative force in virtual experiments is necessary. Finally, in Chapter 4 I will use the current debate over the robustness of climate models (particularly General Circulation Models) and climate modeling as an example of a robust, young science which necessarily utilizes virtual experiments in order to produce robust inferences. I

will then conclude by arguing the virtual experimentation is an irreducible part of modern scientific methodology.

**Table of Contents:**

## Chapter 1: The Spectrum of Virtual Experiments

### *Introduction:*

There has been skepticism in the past about experimental practices which utilize surrogate materials rather than being experiments on the actual materials in which a particular phenomenon appears "in the wild," so to speak, the "target" phenomena. Much of this skepticism derives from the concern that the use of surrogate materials in an experiment necessarily involves doing an experiment which may and of course does involve manipulation of materials very different from the 'target' materials and thus can or maybe *will* involve substantial abstraction and "idealization." In the following, I will address this skepticism by demonstrating that there exists a class of virtual experiments (that is, experiments done with surrogate materials and thus high degrees of idealization) which differ only in degree, rather than in kind, from experiments done in the materials of their target phenomena.[1] I will argue that this class of virtual experiments forms a spectrum and that recognizing the existence of this spectrum suggests the success conditions for such experiments; I will argue further that this approach permits the development of success conditions even for virtual experiments that are as highly "idealized" as thought experiments and experiments involving simulations, models, and theory-dependent apparatuses; and that thinking about such experimental procedures as a class allows the lessons of the literature on each to be applied usefully to the rest.

To do a virtual experiment is do an experiment as part of the investigation of some theory or some phenomenon, an experiment, however, in which the phenomenon or

---

[1] I owe much of this early discussion, and indeed some of the wording, to collaboration with my supervisor, Professor John Baker.

theory under investigation is examined not by manipulating the materials which constitute that field, but by manipulating *surrogate* materials, i.e., materials which *stand instead* of the materials which constitute the field under investigation. Thus, for example, imagine that one was investigating a question in the field of particle physics and in particular in that part of the theories governing the spinning of smaller particles around larger particles. Then one *might* (if one had the resources and the technology) investigate questions in this field by using electron microscopes and interfering with the behavior of electrons around nuclei to see what happens. But if one lacks the resources or the technology one might instead interfere with the spinning of billiard balls attached to string around a central anchor off to see what happens. In this case the surrogate materials would be the billiard ball, the string and the anchor. Alternatively one might sit in one's chair and think about what happens when an electron spins around a nucleus and think about what *would* happen if one interfered with the spinning. Or one might sit in one's chair and think what would happen when a billiard ball spinning around a central anchor is interfered with in some way. In the first case, in doing one's experiment one does not manipulate the materials which constitute the field under investigation (electrons and nuclei) but instead investigates surrogate materials in a surrogate field (billiard balls spinning round an anchor) of one another. And similarly in the second case too in doing one's experiment one does not manipulate the materials which constitute the field under investigation but instead one manipulates (in imagination) the contents of an imagined process and sees what in the imagined process happens. The billiard ball experiment is a *virtual* experiment on an issue in particle physics in that surrogate materials are manipulated instead of particles. Similarly too in the second case one is doing a *virtual* experiment in that surrogate materials are manipulated instead of particles. The second case is a paradigm thought experiment. Both experiments are virtual experiments. Thought experiments and other experiments in which surrogate materials are used differ from "real-world" experiments in that they are virtual experiments; they use virtual materials and interventions as opposed to the actual materials of a particular imagined real-world scenario.[2]

Virtual interventions are interventions upon materials which are surrogate, such surrogate materials being chosen because *to some extent* they preserve the relevant features of the phenomena under investigation. Say that I want to investigate whether a new soap product will cause blindness, my imagined experiment involves testing how chemical compounds will interact within a new soap product, and how that particular combination will effect the eyes of humans using the soap. Consider now that, before *or instead of* doing *this* experiment, I might review what I know about those compounds, and what I know about the human eye, and in so doing I might run the experiment *in my imagination* and in that sense "virtually" and in the medium of thought and having run this "virtual" experiment to consider the outcome of the experiment: call this 'experiment A". Similarly, instead of "doing the experiment" "in my imagination" I might do it on paper: call this 'experiment B'. In experiment A the *material* in which the experiment is done is thought, in B it is marks on paper. We might refer to the materials in which an

[2] It is an interesting question whether it can be said of all thought experiments that they stand in relation to an imagined experiment, one which I do not have the space here to address in greater detail. Suffice it to say that I think it probable that a great number of thought experiments have an imagined experiment, one which is either possible to perform in this physical world or in a possible world (i.e. in the world in which one can in fact travel at the speed of light, etc.)

experiment is done the 'medium' for the experiment. We can, I suggest, imagine a spectrum of experiments, with the various experiments on the spectrum using different media – thoughts, marks on paper, little cardboard models, etc., and at the end of the spectrum we would have imagined experiments in which we would imagination combine the actual chemical compounds which we are interested in in a real dish. If, however, I am more interested in testing whether those chemicals will combine in such a way as to be damaging to the eyes of a mammal, I had better investigate what they do in the presence of mammalian optical tissues, etc. Ethical considerations may make me choose to place the compound in the eye tissue of a dead mammalian eye, or even in a similar compound of organic materials, etc.

A variety of activities found in the sciences satisfy this account of virtual experimentation, including (but perhaps not limited to) the use of models, of simulations, of analogy, and of course thought experiments:[3] I would here that, as I argue later, there is good reason even to count experiments with apparatus whose use in the experiment is heavily theory-laden can, and probably should be counted as virtual experiments on my account of that notion. These virtualized experiment types differ in the particular target properties and materials utilized, and in their closeness along a variety of axes to the target phenomena, as it occurs outside of experimentation.

The relationship between thought experiments and "real-world" experiments is, I will argue, best thought of as follows. There is an N dimensional space at one point in

[3] Hacking's discussion of the epistemic problems of theory-laden apparatus like microscopes is useful in understanding why they ought to be included. Hacking's question "Do we See through a Microscope" is the question of whether

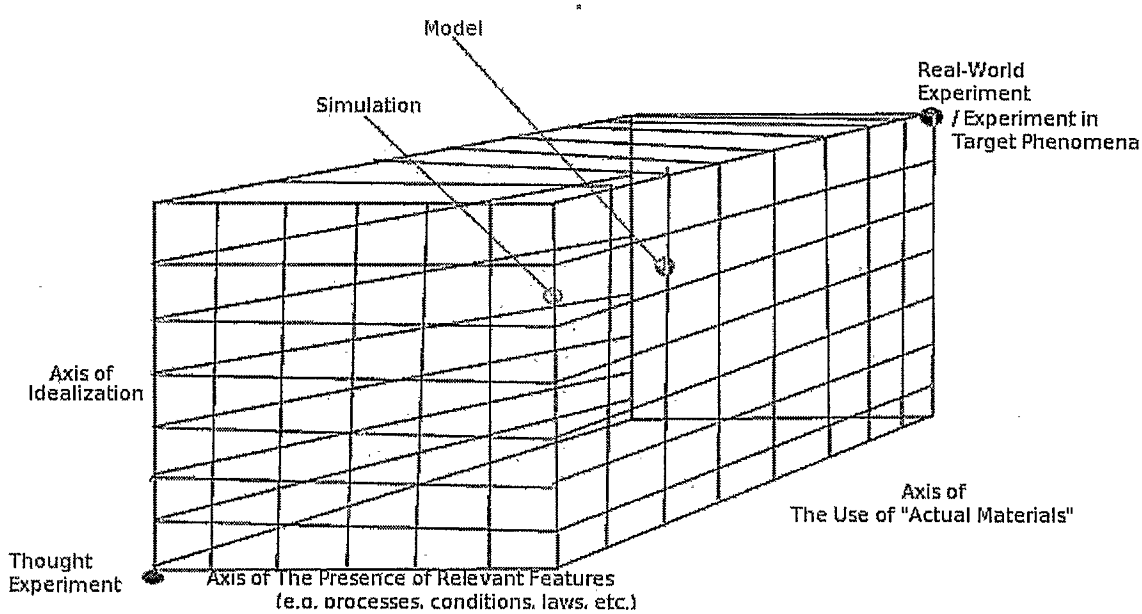
which can be located the paradigm "regular" or real-world experiments on what I will continue to call the "target phenomena", and at another point in which are thought experiments. At the end of the spectrum opposite to the paradigm thought experiment is the paradigm "regular" or real-world experiment. Very generally, an experiment in the physical (including biological) sciences can be taken to be an intervention in some process(s), sequence of events, or state(s) in the physical world. This is an experiment in which the interventions involved in the experiment are done in and upon the actual materials under investigation – physical materials in a chemistry or physics experiment, maybe other kinds of material (psychological or social) in experiments in other fields. Somewhere "between" the two are experiments using models, simulations etc. Understanding this picture of the relationship between paradigm thought experiments, experiments done upon and in the material of the target phenomena and the intermediate cases makes it easier to come to grips with the differences between and the similarities between thought experiments and other experiments in this n-dimensional space.

It is my suggestion that the distance of some experiment on one or more of the dimensions in the n-dimensional space from the location in the n-dimensional space of experiments on the actual, target materials, that is, the amount and the kind of idealization involved, can be referred to as the degree of virtual-ness of that experiment – this degree will, if my suggestion is right, be identifiable on n-parameters in the n-dimensional space.[4] At this stage of my inquiries it is not completely clear how many

[4] A word on idealizations: Idealizations have traditionally come in two forms. Aristotelian idealization involves the stripping away of features perceived as irrelevant to the inquiry at hand; this type of idealization is also sometimes called abstraction. Galilean idealizations involve deliberate distortions

dimensions we would need for the task in hand – my diagram below has three dimensions, but the diagram is just meant to be illustrative of the idea: for this reason I state my hypothesis above as a hypothesis about an n-dimensional space. Virtual experiments include models, simulations and thought experiments, and experiments which use theory-laden apparatus. Of the virtual experiments, thought experiments are, as I might put it somewhat frivolously, the most virtual.

The Spectrum of Experiments




[A three dimensional diagram the n-dimension one, the right side being experiments in which 'actual' real-world/physical materials are the media of the experiment, the left side being purely 'thought' media and in between a variety of experiments utilizing models, simulations, heavily theoretically laden apparatuses (such as electron microscopes, etc.)]

---

(often the removal of a particular relevant feature to see how and if other features relate to it). Both kinds of idealization are present in experiments across the spectrum. For a useful discussion of idealization, see (Frigg 2006). For a useful refutation of skepticism regarding the truth-value of idealizations, see Sorenson's "Veridical Idealization" (forthcoming) and (Odenbaugh 2005).

The claim that there is continuum of virtual-ness in experiments is at least partially justified by the observation that even thought experiments retain many of the key features of experiments. Indeed, Roy Sorenson argues convincingly that thought experiments retain *all* of the important features we usually associate with non-virtual experiments, as we might call them – I will henceforth for brevity call them 'real-world' experiments: real-world experiments are experiments as more traditionally thought of as "lab" and "field" experiments. Lab experiments come in two varieties: ones in which we experiment on the target phenomena under controlled conditions and ones in which we experiment on surrogate materials under controlled conditions. The former differ from "field" experiments only in that in the field we have less controls. Sorenson argues that thought experiments have these key features in common with 'real-world' experiments: both utilize "ongoing tinkering"; they share standard formats, and both use all five of Mill's Methods.[5] [6]

*The Success Conditions of Virtual Experiments*

**1. The Quality of Idealization**

Success conditions for an experiment secure the reliability of that experiment's results. By the reliability of an experiment, I mean the degree to which the experiment manages to answer correctly the question the experiment was developed in order to answer, or the

[5] See (Sorenson 1999) page 257.

[6] Here I follow Ian Hacking and J.E. Tiles in asserting that the hallmark of experimentation is intervention. Intervention can include the observation of naturally occurring interventions (as is often the case in population genetics). Such cases will still be idealizations in that they treat particular observed interventions (or sets of observations) as analogous to generalized phenomena.

accuracy of the fit of the conclusion of the experiment with the target data/phenomena. It seems plausible to suggest that the reliability will depend upon the quality of the idealization involved in the structuring of the experiment.

A word on relevancy: The point here is that any idealization is only as good as its ability to recognize and capture the features of the target phenomena which are *relevant* in the sense that they are (or depict) causal mechanisms, necessary conditions, etc. within experimental design. One might object that this smacks of circularity, since it gives no account of how relevant features are to be recognized as such. Such an account is beyond the scope of this paper, because the recognition of relevant features will depend upon knowledge (both intuitive and experiential) of the target phenomena. My goal is not to give a full account of how one recognizes relevant features, but rather to show that while they do correlate with the degree of idealization, they are not equivalent to the degree of idealization. Field experiments and thought experiments alike can manage to include or omit them.[7]

If an idealization does not include relevant features and does not make appropriate allowance for the fact that some relevant features are missing, then it will fail to give reliable results (indeed, every controlled experiment will involve some abstraction/distortion of the realities of what happens in the target situation).[8] Hence,

[7] Discovering relevant features in the field might be a process of getting one's hands dirty and repeated experimental tweaking. Discovering relevant features in thought experiments depends upon conceivability to possibility inferences, something I discuss at length elsewhere (Brumble, *Forthcoming*).

[8] Here I must say a bit about the notoriously vague term, "relevant" features.

every experiment whether virtual (i.e., using surrogate materials) or not involves idealization and thus every experiment involves reasoning by analogy. Many times, even direct experimentation on the actual phenomena can involve bad idealization if its design does not include relevant and significant features. Examples of what Ian Hacking refers to as doing the right experiment with the wrong theory are in part instances of problematic idealizations, since the theory involved in an experiment is in part, and in the least robust instances of theory, the structure of the idealization (Hacking 1983). For instance, asking how a ball will behave on a frictionless plane involves idealizing away friction, or, theorizing that friction is not a part of the relevant features.

## 2. The Choice of Surrogate Materials and Analogy: Models, Simulations, Thought Experiments, and Theory-Laden Apparatuses

If we accept Hacking's claim that experiments stand in analogous relationship to their "wild," target phenomena, then the use of surrogate materials will require more definition in that analogy inference. A virtual experiment depends upon an inference by analogy not only because it involves idealization, but also because it utilizes materials, processes or states of affairs which are inferred to be analogous to the target ones.

First, a word on analogical thinking: I follow Mary Hesse in the claim that analogical models are created by identifying positive (shared) features between model and target phenomena The experimenter does what s/he can to ensure that the model is designed so that it does share as many features as possible with the target phenomenon – of course, since the decision has been made for some reason to use a model, it cannot share all features, by virtue of being an experiment using a model. Analogical models also involve negative features (i.e., features not shared by the target phenomenon), which are identified and idealized out of the model. Finally, analogical models involve neutral features (unknowns) which may or may not be shared by the model and target phenomena. It may not be apparent at first, but neutral features, not negative ones, are the defeaters of reliability in an analogous model. This is because negative features can be accounted for, but neutral ones are unruly; greater experimental testing, or tinkering, is needed to get these features under control. Computer simulations in particular provide excellent examples of how virtual experiments undergo what Sorenson (in reference to scientific thought experiments) refers to as on-going tinkering, a process which is typical of virtual experiments.

Computer simulations may begin as formalizations of theories, but, to use Nancy Cartwright's metaphor[9], theories aren't vending machines which automatically dispense models; models which have any predictive power or results need to be tinkered with a bit, and fitted to phenomena. Sometimes called autonomous models and sometimes semi-autonomous models because their creation involves theory-independent or semi-independent tinkering, simulations are examples of models which are created from analogies but must be "fitted" to their target phenomena through small changes (Morgan and Morrison 1999) (Winsberg 2003). The difficulty of getting a complex computer simulation up and running can occupy its creators for many years before it is suitably fitted and ready to run experiments. This tinkering is the representational fitting of an analogous set of formal relations onto a system, and it is precisely the activity in which neutral features of analogy are weeded out or assumed causally related to positive and

[9] For a full discussion of the flawed "vending machine" theory/model relationship assumption, see (Cartwright 1983).

negative features, in ways which I will discuss later.

In computer simulations which depict systems of vast complexity, like weather systems, granting epistemic credibility to results realized using reasoning which uses idealization will be critical because the simulation utilizes abstract mathematical structures and entities as surrogates for physical phenomena. This will require that the construction of the simulation take into account many features of the "wild" phenomena and make assumptions as to which features must be preserved. Idealizations in computer simulations exhibit on-going tinkering to this effect; aspects of the formal systems involved are tweaked to closer represent aspects of the phenomena and produce results which more closely match those observed in "the wild."

One further observation from the literature on computer simulations is needed for another requirement, the requirement that, with a virtual experiment, the choice and construction of a virtual experimental procedure will require substantial knowledge of the surrogate medium and in fact may very well require *expert* knowledge of the surrogate medium. In the case of computer simulations, often a great deal of energy must go into their design, debugging, and tinkering to get representative results. But this is also true of all sorts of experiments which involve other kinds of modeling, including, I would argue, thought experiments. All virtual experiments will require specific knowledge of not only the target phenomena, but also the medium in which the experiment is conducted, if the needed analogical reasoning is to be robust.[10]

Some initially neutral features may be re-categorized as positive or negative because

[10] For a parallel argument regarding expertise in computer simulation, see (Winsberg 2007) and (Winsberg 2009).

they are causally linked to features we know are present (or absent). Experiments using animals to examine human phenomena typify this. I may choose to do research upon obesity in mice because I notice the presence of a particular gene in a mouse population which, in humans, has been linked to obesity. I assume as part of my experimental design the theoretical commitment that this gene is causally linked to the phenomena in which I mean to make an intervention, because I have good reason to suspect that causal connection, given the relevant ways humans and mice are genetically and physiologically similar. Thus, inference by analogy is an inductively valid form of inference just in the case that it is likely that there is a causal connection between the features depicted in my analogy, on the one hand, and the hopefully parallel causal connections in the target situation.[11]

Idealization may be present in experiments across the spectrum, but, assuming one can intelligibly speak of amounts of idealization, idealization becomes progressively greater and the importance of good idealization becomes progressively more important the further one moves away from target phenomena in the direction of using progressively more 'alien' surrogate materials. For instance, idealizations in "wet" models (models in biology and organic chemistry constructed at least partly with some of the organic, target materials) may implicitly "anticipate" (as I might put it) that the use of target materials ensures that some of the relevant features are present, without fully identifying them. The recent development at Harvard of the lung-on-a-chip is a fine example of a "wet" semi-virtual idealization. In the words of the Harvard Gazette:

[11] For a useful discussion of inductive reasoning in analogical arguments, see (Wylie 1988).

> The lung-on-a-chip microdevice takes a new approach to tissue engineering by placing two layers of living tissues—the lining of the lung's air sacs and the blood vessels that surround them—across a porous, flexible boundary. Air is delivered to the lung lining cells, a rich culture medium flows in the capillary channel to mimic blood and cyclic mechanical stretching mimics breathing (Harvard Gazette 2010).

The experiments utilizing lung-chip will "anticipate" that, as an idealization in wet material, it preserves the relevant, positive features of the target system (the human lung). Some neutral features may be accounted for in the chip-lung by bridge laws (perhaps the lung uses animal tissue which has been determined to be nearly identical to human tissue, etc.).

Further still along the axis of idealization are thought experiments which require the most exacting care in choice of and control of idealizations if they are to produce reliable results. Since a thought experiment will needs be much rougher and less exact than, say, an experiment done with a computer simulation, the experimental design will be highly idealized. Again, consider Galileo's Falling Bodies thought experiment. Vast arrays of features of the physical world are assumed as base-knowledge for the experiment, and many of them are idealized away (wind speed, friction in air etc.).

Thought experiments depend upon inferences by analogy between the imagined intervention in a mental representation (built from background knowledge and theoretical commitments) and the target phenomena. The inference done when doing a thought experiment moves from mental representation, (imagined or conceived) to target phenomena. The construction of a thought experiment will need to recognize ways in which the imagined intervention upon phenomena is analogous to the target phenomena.[12] Thought experimentation is just as (if not more) dependent upon inferences by analogy as other forms of modeling by virtue of the high level of idealization involved. If idealization, as I have claimed, is a function of the closeness of some assumed analogy between the experimental situation and the target situation, then thought experiments must have the most precisely defined idealizations in order to permit reliable inferences about the physical world.

Until now I have said very little to justify my inclusion of experiments with heavily theory-laden apparatus amongst virtual experiments. Let me now first make that claim explicit and then justify my claim. Experiments which use apparatus such as telescopes, microscopes and fMRI scanners come with embedded theories that underpin their use. As some philosophers have shown, the uses of these apparatuses are confirmed externally to the experiment in which the apparatus is used through other methods and observations.[13] However, in the absence of outside confirmation (such as alternate and

[12] The success conditions of well-formed idealizations in thought experiments in particular are unfortunately beyond the scope of this paper. I follow (Yablo 1996) and (Hawthorne and Gendler 2005) in distinguishing imagination from the more rigorous activity of "conceiving" and I also follow them in making the claim that conceivability to possibility inferences underlie successful idealizations in thought experiments, and that a detailed account of the success conditions of said inferences in modal epistemology is the key to understanding the relative probative force of different thought experiments. For further reading on the distinction between imagination and conceivability, and the success conditions of conceivability to possibility inferences, see (Yablo 1996). For the direct application of these inferences to thought experiments, see (Hawthorne and Gendler, 2005). For useful distinctions between inferences to different kinds of possibility, see (Vaiyda 2007).

[13] See (Chalmers 2003) and (Hacking 1983).

supplemental observations), descriptions of experiments using these apparatuses should be viewed as including reference to those theories, importing all of the theories' content (that is to say, the theory of the device affects what we can claim to observe with them).[14] Additionally, these devices use the theory-dependant observation of only a small part of a phenomenon or a mechanically interpreted output (patterns generated by an MRI scan are not the direct observation of brain activity itself) as a stand-in for the phenomenon itself. Thus, we can only claim to be experimenting upon the target phenomenon in the presence of robust bridge laws, which will vary upon the application of the apparatus.[15]

***The Relative Strengths of Inferences: Heuristic and Probative Value***

Some virtual experiments, because of the extent of the idealization they involve, will not be able to draw robustly probative inferences about the world, yet can be viewed as serving to clarify and structure concepts, revise and suggest hypotheses, and suggest future directions for experiments. In fact, a great number of experiments count as virtual experiments because they are done upon materials which the experimenters hope are the actual materials of the target phenomena, but which are in fact only *known* to be analogous to the target materials. Experiments done upon the physiology of mice (with

[14] Hacking's discussion of the epistemic problems of theory-laden apparatus like microscopes is useful in understanding why they ought to be included. Within my argument, Hacking's question "Do we See through a Microscope" is the question of whether a particular theory embedded in the apparatus presupposes some relevant features, and indeed whether it presupposes the use of intervention upon actual materials when what is being viewed/experimented upon may only be a theory-dependant representation of actual phenomena (i.e. a model).

[15] For further discussion of the success conditions of experiments with theory-laden apparatus, see (Hacking 1983) and (Chalmers 2003).

the hope that they generalize to humans) are often of this type. Additionally, some experiments attempt to identify the target phenomenon's materials. For instance, currently, there are hundreds of candidate molecule markers for neuronal "tagging" (the probable mechanism by which neurons establish short-term increased efficacy). Experiments upon the phenomenon of short-term plasticity can at best defeat hypotheses and suggest correlations between candidates; that is to say, they currently weed through a set of possible configurations of markers and generate hypotheses for future experiments. And, of course, probably the greater number of scientific thought experiments do not give robustly probative results about the world, but rather serve to ground claims of a logical or conceptual kind: showing the internal inconsistencies of theories by revealing the logical and conceptual impossibilities within them. Thinking about the spectrum of virtual experiments reveals that when experiments involve comparatively loosely defined idealizations, and also when the target phenomenon itself is unclear, experiments can still have very useful heuristic functions, and that rather than being simply inferior to robustly probative experiments, heuristic experiments pay a vital role in the scientific process.

***The Project Ahead***

I hope I have made a persuasive case for the n-dimensional spectrum of experiments as a useful framework for thinking about the success conditions of experiments. In this section, I will at best adumbrate the competing views and possible objections: I will examine at least some the points in this section in some detail in subsequent chapters.

The most obvious objection to my account of the success conditions of experiments arises from an intuition about the superior reliability of experiments that work on the target materials; some philosophers suspect that experiments in target materials have a higher degree of reliability (in the sense of this term that I commented on earlier) than experiments which work on surrogates. The intuition goes something like this: experiments can, as Hacking says, "have a life of their own" independent of theory formulation (confirming and disconfirming hypotheses, suggesting further experimental design, tinkering, etc.), in part because they constitute direct interventions in the target material. Those who find it persuasive that intervention in the target material will guarantee robust results anticipate that, somehow, the closer an experiment comes to recreating the "wild" phenomena, the more likely it is that results will be reliable. Cashing this intuition out in terms of relevant features makes its mistake clear: one assumes the presence of the target phenomena/material is paramount to the presence of the relevant features,[16] but it is not. Since all experiments involve some degree of abstraction and idealization (because there must be control in experimental design and choices made about the kinds of interventions enacted, etc.) all experiments will stand in a analogous relationship to the target phenomena; every experiment will be separated (as a necessity of being an intervention) from untouched "wild phenomena," even if said experiment is done in the "wild." This means that an experiment may happen in the target materials with an experimental design which guarantees, skews, obscures, or falsifies its conclusions. Target materials may be more likely to preserve relevant features, but they are no guarantee of said preservation against the abstraction inherent in experimental design.

[16] As evidence of this ongoing theme, Allen Franklin suggests, following Ian Hacking, that interventions in target materials constitutes a condition for robust observations (Franklin 2005). Deborah Mayo suggests a similar condition (Mayo 1996).

Now for a few literature-specific worries. There has been a surge of literature recently defending the idea that there can be theory-laden apparatuses that are nevertheless robustly reliable, independent of the theory being tested in the use of the apparatus. Ian Hacking's argues that some theory-laden apparatuses are resistant to worries about theoretical contagion because they work well with almost any theory of the apparatus, and many have followed in that vein.[17] They argue further that some experimental procedures can be classed as "robustly grounded" regardless of the theory currently employed to justify the observations. While it might seem at first glance that the existence of robustly grounded procedures with apparata conflicts with my assertion that such experiments belong to the class of virtual experiments, this view is actually compatible with mine.

First, let me make the distinction between theory-independence and multi-theory-compatibility. The theory of how some apparatuses work may be compatible with a variety of theories about the phenomena under investigation (e.g., the theory of how fMRIs work may well be compatible with various views about which parts of the brain are involved in remembering , but it is doubtful that they are compatible with (in the strong sense that they entail) *any* theory. Electron microscopes should not be compatible with (entailing) theories of conscious quarks. fMRI scanners should not be compatible with theories to the effect that they are 'perceiving' the neuronal activity of anything like a *soul*. Indeed, were the theory to say this, , the theory would smack of pseudoscience. To be a robustly grounded device is to be a device compatible with any *likely candidate*

[17] For full accounts of this view, see (Hacking 1983), (Mayo 1996), (Franklin 2005) and (Baker, *Forthcoming*)

*theory*, and such theories will most likely posit many of the same relevant features: for instance, a telescope may perceive light waves moving through ether as well as the highly analogous movement of light waves posited in post-ether theories of light.

There is good reason to say that telescopes do not interact with stars, but with bits of light from stars. There is also good reason to say that fMRI scanners do not yield pictures of brain states, but rather very specific electric activity in the brain. One can hardly call interventions in the ephemeral side effects of target phenomena direct interventions (without committing category errors). Even when such apparatuses interact with only a small aspect of the target phenomena, it is possible for robustly grounded procedures to produce data which are compatible with and certainly are informed by a family of theories, and which produce observations which thus preserve a proper analogical relationship to the set of relevant features of the target phenomena. For instance, maybe there is reason not to say that an MRI scanner does not observe a particular cognitive state, but it may reliably observe correlated effects in terms of neuronal activity of that cognitive state. Likewise, uses of Sydney Shoemaker's c-fiberscope may assume that c-fibers are an important part of the physical base of pain and using the c-fiberscope we may say that we observe that the behavior of the c-fibers is correlated with pain (thus, maybe, allowing certain inferences about the causes of pain) even if it turns out later that the c-fiber firing is not all there is to the physical base for pain.[18] Apparatuses do not need to link to a particular theory in order to reliably predict or correlate a complete account of a phenomena. Nor do they need to interact directly with target phenomena; they just need to indicate and anticipate the presence of particular

[18] See (Shoemaker 1994)

relevant features and use a theory of observation which does not obscure those features.

An additional and useful worry about virtual experiments comes from the literature on computer simulations. Some have doubted the probative power of experiments using computer simulations, arguing that they involve so much tweaking in their design (in order to match existing data), that they cheat by guaranteeing their conclusions and making it impossible to refute them. Users of computer simulations often spend a good deal of time trying to get their simulations to produce results which match recorded "wild" phenomena. The fear is that these tweaks done upon the simulations' formal modeling may create a system which confirms controversial mechanisms and other aspects of the phenomena by importing them into the experimental design.[19] Thus, once a simulation begins to give "accurate" results, it merely reiterates the theory it means to test. This worry, I think, is generalizable to all virtual experimental design. Indeed, I think it is generalizable to all experiments, since tweaking experimental design is a part of the ongoing development of experiments. While confirming the consequent ought to be a genuine concern in the design of experiments, it is not necessarily solely a concern (or more of a concern) for virtual experiments. It is something all experimenters must guard against.

### *The Hunting and Cataloguing Virtual Experiments*

The aim of this thesis is to legitimize virtual experiments through identifying and classifying them according to their success conditions, and to demonstrate that they constitute a vital part of modern science. Each chapter will build on the recognition of

[19] For a useful discussion and summation of this debate in the literature on computer simulations, see (Winsberg 2001).

the Spectrum of Virtual Experiments to refine our ability to recognize the roles virtual experiments can play in modern sciences according to the inferences they warrant.

To this aim I will, in chapter 2, consider how experimental design in virtual experiments depends upon "conceivability to possibility inferences." In this taxonomy, "conceivability" refers to a specific rigorous state of affairs and where "possibility" can refer to a variety of types of possibilities, each with different inferential power. Recognizing the differing inferential power of these inferences sets us up to further refine our taxonomy of virtual experiments, and thus recognize their refined roles in the sciences.

Chapter 3, I will distinguish between two kinds of inferential powers, "Probative" and "Heuristic," which by virtue of having different inferential powers and aims thus have different success conditions. Recognizing that there exists a spectrum of virtual experiments allows for the recognition that not all experiments aim for or achieve robustly probative inferences regarding the target phenomena. Yet, many such experiments are still integral parts of scientific methodology in general and experimental methodology in particular. I will delve into these distinctions in greater detail in chapters 2 and 3, but for the moment I want to suggest here some preliminary terminology for differing inferential force, along with definitions.

**Experiments with Probative Inferences**: Experiments which result in robust inferences directly about the target phenomena I will call *probative*. This set of experiments most closely matches our intuitive sense of what experiments out to do. They query facts about the target phenomena and give *highly probable* (in that they far out match their nearest competitors) explanations (though they do not confirm or "prove conclusions"). As I will argue further in Chapter 3, some probative experiments are second order probative, in that they give probative inferences about experimental design as their target phenomena, in addition to other heuristic inferences.

**Experiments with Heuristic Inferences:** Experiments which generate hypotheses, regulate data, lend conceptual aid, test and improve experimental designs, suggest further experiments, etc. fit into the category of *heuristic* experiments. Let it be noted that while probative experiments can also have heuristic value, let heuristic experiments refer to the class of experiments which solely contain heuristic value.

Finally, in Chapter 4 I will use climatology and its dependence upon computer simulation climate modeling as example of a modern science which depends upon virtual experimentation (and refined distinctions about the inferential power various virtual experimental methodologies warrant). Along these lines, I will argue that legitimizing climatology for policy-making will depend on a revision of more traditional views about experimental design and inference in science in general; revisions already accepted widely in the scientific community.

Now that we have recognized the existence of a spectrum of virtual experiments, we ought to want to know what makes experiments in general, and virtual experiments in particular, likely to succeed in yielding salient inferences. The answer so far given, that their design depends upon the recognition of "relevant features," is frustratingly vague. What, in general, is a relevant feature and how do we know when we have one in sight? One strategy for identifying them might come from the discussion in modal epistemology of "conceivability to possibility inferences". This is because experimental design, by virtue of experimental control, idealization and abstraction, requires a conception of *how things might be* or in stronger cases *how things probably are.*

There exists some very useful literature on the role "conceivability to possibility inferences" play in the experimental design of thought experiments. Since thought experiments constitute one end of the spectrum of experiments, it would probably be productive to understand how modal epistemology in general and conceivability inferences in particular operate within experimental design for all experiments, and how the particular force and specific inferences (inferences to various kinds of "possibility" and to actualities) may change as we move along the axes of the spectrum.

## Chapter 2: "Imaginability", "conceivability" and inferences to "possibility":

### *Introductory Taxonomy*

All experiments explore possibilities. All experiments also involve a degree of idealization (even if it only consists in minor choices made in designing the experiment to secure experimental control of variable) and thus involve *conceiving* of a particular state of affairs. This conceiving involves picking out a set of features which might be relevant (and a set of features which the experimenter suspects, though does not know, might be relevant). Thus, all experiments utilize "conceivability to possibility inferences" of some kind in the designing of experiments. One trick to discovering the success conditions of these inferences, of course, is first determining what *kind* of "possibility" will be yielded by the inference. When exploring possibilities, thought experiments (and to a degree also other virtual experiments) often rely upon inferences from "imaginability" to "possibility", and, in ways to be discussed, more narrowly from "conceivability" to "possibility". Because these inferences are used widely throughout the sciences and philosophy, and are present in some of their most important arguments, it seems that these inferences are at least generally taken to be successful. That is to say, it is at least generally assumed that there must be something about them or some way of using them, which leads to successful inferences. Yet there is wide criticism of these inferences because there is no satisfactory account of how and why they work. Furthermore, just as inferences from "imaginability" to "possibility" and "conceivability" to "possibility" were used in some of philosophy's most enduring theses, they are also present in some of its failures. If we want to go on using them, and to understand how, when and why these analytic methods work, a set of success conditions for these inferences must be found, and we must come to understand why those success conditions are what they are.

Further muddying the waters, "imaginability", "conceivability", and "possibility" are used in a variety of ways throughout philosophical literature. Thus, the first task at hand is to see how these terms are being used and come up with adequate descriptions of how they ought to be used. Then, to decide whether such inferences are reliable, we will need to spell out what, in a successful inference, might be meant by "conceivability", what by "imaginability", and what type or types of "possibility" they might be useful and reliable guides to.

### *The Conflation of Imagining and Conceiving*

> Alice laughed. `There's no use trying,' she said `one can't believe impossible things.'
>
> `I daresay you haven't had much practice,' said the Queen. `When I was your age, I always did it for half-an-hour a day. Why, sometimes I've believed as many as six impossible things before breakfast.

In the following, I will argue that all virtual experiments attempt to test for the existence and nature of types of possibilities. Some even investigate necessities by way of discovering things which are not possibilities. Thus, a discussion of modal epistemology is necessary in order to discover the success conditions of various types of virtual experiments. Since thought experiments in particular among virtual experiments ask us to do a lot of imagining, it is particularly important to investigate the success conditions of "imaginability/conceivability to possibility inferences" with particular reference to the context of thought experimentation. However, because all virtual experiments will involve one type of another of these inferences, though with varying success, what can be said of the role of modal epistemology in thought experimentation holds true, to a degree, for all virtual experiments. Because of the vast range and diversity of the philosophic literature on modal inferences, it is best to begin this project with some stipulations about terminology.

When discussing how we arrive at reliable modal inferences, we might be tempted to use the words "conceivability" and "imaginability" interchangeably. However I think that the humor in the above quote from the Logician Charles Ludwig Dodgson's (Lewis Carroll) *Alice Through the Looking Glass* arises from the two characters talking past each other by referring to these two distinct activities. An Abbott and Costello-esque pair, the logical and pragmatic straightman, Alice, intends a more rigorous and methodical type of thinking, while that flighty denizen of Wonderland, the White Queen, reveals the danger of belief that comes from hasty acts of imagination. The exchange is humorous, because we intuitively know that imagination and the kind of rigorous contemplation needed before we would call an act of imagining a case of conceiving are often mistaken for one another, and not just by the denizens of Wonderland. We philosophers in particular shouldn't make the conflation, because the two terms distinguish two important mental activities, which are both part of making modal inferences, but which are different in kind and result in inferences of different probative force.

So conceiving is more rigorous than imagining, but what does it involve? Stephen Yablo (1993) constructs a careful reportive analysis of the ways in which philosophers have used the verb "to conceive", and comes to the conclusion that generally when we talk about conceiving of something, we think both that we have found it believable, and also not unbelievable. Yablo argues that to find something conceivable is to find it "believable for all we know". Conceiving is justified belief, according to Yablo, in a very strong sense; namely, in the sense that we only find anything we have good evidence for (the reliability of our perceptions, for example) believable, i.e., in virtue of our best available knowledge. The sources of the best available knowledge may include experience of the world, theorizing about the world, evaluation of a relevant group of theories and may be the product of thought experiments. It will often consist of knowledge of empirical facts, and it may be the product of sustained thought about simulations.

In at least some cases, "imagination" is a good term for the kind of preliminary, generative phase in which a thinker tries literally to imagine a particular state of affairs. A wide variety of activities ought to be counted as imagining, including imaging, drawing pictures, describing using metaphors, similes and analogies, as well as using role-playing (private mental rehearsals and role-playing with other people – (such as emergency-response training, etc.)). Thus it is crucial to recognize that imagination includes a wider set of activities than merely a set of imaging activities; one can imagine and be congenitally blind. When one begins to imagine something, one begins with what is at first relatively undefined. One imagines what a state of affairs might be like. This phase is creative, but less well defined and determined than the thinker may be capable of. Heuristics like metaphors may play a role in this phase, assisting the thinker in exploring various possible configurations of features and states of affairs. An example here is helpful. I have it on the authority of one of his students that Gilbert Ryle used the following activity to illustrate the diffuse and indeterminate nature of imagining to his students. Ryle's activity goes something like this:

> Close your eyes and imagine a dog. Now tell me, can you see all four legs? Does the dog have four legs? Is the dog black, brown, white? Is it wearing a collar, etc.

One might object to this characterization of imagination along Wittgensteinian lines, that one can be sure one dreamt of King's College (Wittgenstein 1958) (and not some other building) and that thus, imagination can grasp identity without determinateness. To this I reply that one can be very sure that one is in fact imagining a dog (and not a cat, especially if one has seen a dog), but one does not imagine an exhaustive list of features for the dog, let alone even all the features that we might take it a dog has got to have (a particular number of legs, a color, the presence or absence of a collar, etc). Whether you think essential features, necessary features, a good definition or particular relevant sortals of features make something what it is, those features will need to be present to do more than imagine in a weak sense some particular thing or property. This feature of imagination makes some dreams very strange. Who hasn't had a dream in which their friend/mother/ brother was present, but was also somehow a monster/Albert Einstein dog? We can know what we think we are imagining in the sense that we can call it by a particular proper name, but it does not follow that we know it's nature in any robust sense (on any theory of identity, my mother is not a dog, yet I am capable of imagining that specifically she is, somehow).

We can see that what one imagines does not even have to obey the laws of logic (in the sense that an imagined dog can indeterminately neither have a collar nor not have a collar). And as I will now suggest one can count something imaginable even if it involves what we would take to be a physical impossibility.

We readily imagine physical impossibilities. As a cursory glance at the artwork of M. C. Escher can show us, vividly, imagination is capable of broader and wilder minglings than more ordered and determined conceiving. We can, like Carroll's White Queen, freely imagine impossibilities, both known and unbeknownst to us. Thus, if "imaginability" is not a reliable guide to "possibility", we need a term for the mental activity that it might be.[1]

"Conceivability" is a good term for describing what might be thought of as a later, more refined phase, in which the thinker considers whether he or she can generate a robust theory for how such a state of affairs might be. This phase involves a more concrete and intentional synthesis of knowledge about the world and its application to the case in point.[2] At this point, we might suggest that "conceivability" might be viewed as imagination in conjunction with a good working theory about how what one is imagining could be possible. My reasons for taking this particular definition seriously will become clear later, when I spell out the workings of "conceivability" to "possibility" inferences.

---

1 One can also conceive of impossibilities while knowing them to be impossibilities, provided one adds a premise stipulating the absence of a particular part of the system one is conceiving in. For instance, one can conceive of a scenario involving an inclined plane without friction, because one clearly stipulates that it is a frictionless world.

2 (Chalmers 2006) makes a similar distinction, naming his two processes primary and secondary conceiving, where secondary conceiving is "ideal conceiving." I avoid Chalmers' distinction for two reasons. One, because imagination is already a perfectly good and distinct term for *prima facie* conception, and secondarily, I wish to avoid all talk of ideal conception and ideal conceivers, which will become clear later when I discuss platonic rationalism in the work of Brown on thought experiments.

Of course, what you will need to know in order to conceive of something, that is to say in order to imagine it and have a good working theory about how it might be possible, will be determined by what type of "possibility" you have in mind. Let us consider the many kinds of "possibility" philosophers make use of.

So "conceiving" means forming a rigorous theory based upon analyzed evidence, both conceptual and empirical. That is to say, conceiving of something means rigorously consulting all the information one can have at hand to form a judgment about how that thing most likely is, in light of our best relevant knowledge. The bar for conceiving of something being possible is thus a rigorously high one, and will depend upon well-formed and informed thought. Immediately we see that acts of conceiving exist on a spectrum of are better informed and worse informed instances, moving away from imagining and towards more and more rigorous thought experimentation.

A critical aspect of this definition of "conceiving" is that it keeps "conceivability to possibility inferences" from collapsing into circularity, because more is involved in conceiving than just assuming premises needed for a conclusion. Some sort of theory is created, and tested before the inference moves forward. This testing, I will argue, is a fundamental part of virtual experimentation. To understand better the features of conceiving, let's consider some standard philosophical criticisms of "conceivability to possibility inferences" and the strategies for refuting them.

***The Objections***

Considering the naive collapsing of "conceivability" into "imaginability" brings us to our first group of objections to the value of inferences to possibility from conceivability:

these are objections which claim that such inferences involve various sorts of circularity.[3] Since the constraints on what we can imagine *prima facie* look less rigorous than those on what we can conceive, to conflate them is to bring onto "conceivability" to "possibility" inferences the familiar objection, classically brought against "imaginability" to "possibility" inferences that they suffer from circularity. The standard framing of the objection, according to Yablo[4], runs something like this:

> If what "conceivability" means is imaging that something is possible, then in order to draw the inference that something is possible, we must first imagine that something is possible. This means that finding something possible is predicated upon assuming its "possibility" or finding some premises which make it possible (as part of imagining it). This is surely circular reasoning.[5]

If to conceive of something is just to imagine that it is possible, than anything that we can imagine could be grounds for finding something possible. But of course, like the White Queen, we can imagine all sorts of clear impossibilities. That is part of the fun of

[3] For a detailed discussion of the circularity objections, see Yablo's extensive discussion of the circle in section V and VI, pp. 49 (Yablo 1993).

[4] See page 13 of (Yablo 1993) for a detailed discussion.

[5] This objection to imaginability to "possibility" inferences is often first credited to Arnold's objection to Descartes, "how does it follow from the fact that he is aware of nothing else belonging to his essence, that nothing else does in fact belong to it?" Here Arnold points to the fallibility of pure imagination as a guide to metaphysical "possibility" (CSM II, p. 140).

imagination, I suppose, and part of the pleasure that comes from looking at M. C. Escher's drawings!

Conceiving then must be viewed as being more rigorous and robust of a guide to possibility than imagining. One might object here that "conceivability" to "possibility" inferences are just as circular as "imaginability" to "possibility" inferences, since presumably they can be formulated thus:

> If to conceive of something's being possible is a guide to what is possible, then surely we run afoul of the same sort of circularity wherein in the act of conceiving of something's "possibility", we assume the very premises needed to make something possible.

But let's consider some of the strongest objections against "conceivability" to "possibility" inferences; doing so should help further convince the reader both to view "conceivability" as being not just "imaginability," and that my formulation of what is involved in conceiving best answers these objections.

Stephan Yablo identifies the following classic objections to conceivability to possibility inferences (Yablo 1993). The first objection, which is an objection closely related to the circularity objection I have just commented on, is that finding something conceivable is just finding it believable-possible: that is to say, possible to be believed. This consists of the conflation of the everyday sense of "conceivable" ("believable") and the philosophic sense of the word (which requires the appearance of a "possibility"). We know that we can believe all sorts of impossibilities. Believing impossibilities is not the same thing as imagining impossibilities; after all, one can believe something P because one has a good working theory of how something could be, but it often later turns out that

P is impossible because of a particular feature of the target phenomenon instantiating P about which you were ignorant.[6] Imagining something is possible is prey to the same problems that believing something is possible is prey to, for seemingly the same reason (i.e. a possible lack of relevant information AND a possible lack of a strong theory about how it could be the case). Refusing to conflate "conceivability" with "imaginability" does not protect our inferences from missing defeating information, but it does make them less prey to a lack of a good theory, and it does make them of a different kind than believable-possible to possible inferences.

But Yablo identifies more serious objections that take the form of circularity objections. His second and third objections are closely related. The first insists that if you don't know all the relevant information for sure, you should hold off drawing the inference until you have all the relevant information. In this case we infer that if something is conceivable (we have all the success conditions for "possibility" identified) then it is possible. But since on this account we can infer that something is possible from the fact that it is conceivable if and only if all the success conditions are met, this is circular. From this objection we may also conclude that we ought not conclude with too much confidence that our theory is correct until we have ruled out all competing theories, which could also account for the target phenomena. This second objection makes the same move with regard to competing theories, which the first objection did with regard to information. Yablo notes that both these objections fail to convince us that we ought to abandon "conceivability" to "possibility" inferences because they assume that we can

[6] Yablo's example of this is that of Aristotle thinking matter is infinitely divisible, because he is missing some particular knowledge about matter

only be justified in finding something conceivable if we can reasonably rule out the "possibility" of unknown defeating conditions or remote rival theories.

I follow Yablo in holding that these objections to "conceivability" to "possibility" inferences are instances of an unacceptable form of skepticism, which if accepted brings in question much of our ordinary empirical knowledge and our most rigorously tested scientific knowledge. There is no reason to assume we are unjustified in using a theory because all of its poorly formed and improbable rivals have not been defeated. Likewise, if we agree that one ought to only accept a "conceivability" to "possibility" inference if all the things that could even remotely possibly be success conditions are accounted for, we must commit ourselves to one of the most virulent forms of skepticism. Following Yablo, I conclude that these objections fare no better against "conceivability" to "possibility" inferences than they do against many other forms of inference. They do not distinguish between well-formed theories and poorly formed ones, and between clearly relevant information and "long-shot" success conditions for possibilities. As I will show, inferences that adhere to these strict rules will yield only logical possibilities, conceptual possibilities and the most trivial forms of metaphysical "possibility" (if any).

Yablo deals with a further objection (which he attributes to Arnold), that one must know all the essential properties of the state of affairs in order to make well-grounded modal claims. Like the objections before this one, insistence upon the need to identify all the essential features (and establish conclusively which features are accidental) commits us to a very limited set of inferences with very limited results, and also fails to distinguish between features which are very likely to be essential (say, that I am human) and features which are unlikely to be essential (that was born on March 31~~f~~st). I think this becomes apparent further in this chapter, when I discuss metaphysical "possibility".

"Conceivability" is achieved only if one has the best available knowledge. Establishing that some claim is conceivable requires that the conclusion that it is conceivable rests on a very well-formed theory. It should be clear now that even the strongest objection to "conceivability to possibility inferences", namely that they treat as a sufficient condition for finding something S conceivable that one does not realize that S is impossible, is defeasible. Much, much more is involved in concluding something is conceivable in that it is a more rigorous form of verification than finding something imaginable. Whether one agrees that this is how "conceivability" has been used in the past, it is clear now that that is the kind of "conceivability" inference we ought to be interested in using to infer robust possibilities[7]. It should also be clear that the amount of knowledge we have about a particular state of affairs upon which we may develop the theory involved in conceiving, will be directly proportional to the probability that the inference is successful. I will discuss these knowledge bases, and the "possibility" of achieving them, in a later chapter on the domains of virtual experiments.

### *Parsing "possibility"*

The different aims of heuristic and probative virtual experiments rely upon inferences to different types of "possibility". Thus, it is critical that one distinguishes between the

[7] One might argue that imaginability to possibility inferences are good for querying conceptual possibilities. Perhaps some defenders of empirical investigations into intuition would take this stance. However, concepts often need robust knowledge and theory to explore, so I am sticking with conceiving as the proper inferential mode even for conceptual possibilities.

types of "possibility" that can frequently be found at work here. Consider Anand Vaidya's clear definitions of the types of "possibility" often referred to by philosophers:

> **Logical "possibility"**: Scenario S is *logically possible* in system *L* if and only if a description of S is consistent with the axioms of system *L*. *Example*: Given propositional logic, it is logically possible that 'Italy will win the 2008 world cup' (i.e., it is logically possible that the proposition is true and logically possible that it is false.)[8]
>
> **Conceptual "possibility":** Scenario S is *conceptually possible* if and only if S is not ruled out by the set of all conceptual truths. *Example*: It is conceptually possible that the earth is flat.
>
> **Metaphysical "possibility"**: Scenario S is *metaphysically possible* if and only if S obtains (or exists) in some metaphysically possible world. *Example*: It is metaphysically possible that some physical particle moves faster than the speed of light.
>
> **Physical "possibility"**: Scenario S is *physically possible* with respect to physical laws *L* if and only if a description of S is logically consistent with *L*. *Example*: Given the actual laws of physics, it is physically possible for a train to travel at 150 mph.

Of course, these terms are not always defined in these ways and of course a lot of work has been done on trying to refine the definitions — see for example the work of Lewis on

---

[8] This definition is a slightly revised version of Vaidya's which should appear in the Stanford Encyclopedia entry shortly.

physical "possibility". Moreover, there is considerable debate among philosophers as to whether metaphysical "possibility" and logical "possibility" are coextensive: are they extensionally equivalent (i.e. do they have same members or more strongly), intensionally equivalent (i.e., do they have the same members because the test for membership is the same in both cases).[9] One may take logical and metaphysical possibilities to be extensionally equivalent (equivalent maybe merely by chance, but not by character) or intensionally equivalent (equivalent because they have the same criteria and the same test for membership.) Since there are certain things, which are logically possible, but only appear metaphysically possible when imagination is used as a guide to "possibility", such as that water is not $H_2O$, it seems unlikely that logical and metaphysical possibilities are ever entirely coextensive. In fact, all metaphysical possibilities must be logical possibilities, because nothing can be possible in any world if it is not logically possible according to some logical system, assuring internal cohesion of a particular world.

To elaborate, if something is a logical "possibility", then it is not logically necessarily false. It is not the case that in all worlds it is not true. At least some propositions are metaphysically possible and that they are possibly true depends on *a posteriori* facts. However, some statements of metaphysical necessities are *a posteriori* true, so presumably some statements of metaphysical possibilities are *a posteriori* true – meaning some are world-relative, and thus relative to a particular logical system (and a particular type of logical "possibility"). Since probative virtual experiments query specific facts about this world or very similar worlds (and thus query physical and

[9] I borrow these concepts of equivalence from David Lyons' "Forms and Limits of Utilitarianism."

metaphysical possibilities) a probative virtual experiment may at once query a metaphysical "possibility" *and* a logical "possibility", but may not only query a logical "possibility" (such a virtual experiment would be heuristic).

So following Kripke, there are a great number of statements of metaphysical possibilities which are *a posteriori* true, and which must be specified world by world. Some statements of metaphysical possibilities are undoubtedly statements of possibilities about how things could have turned out in a particular world, given its particular set up and constraints. For instance, in this world it is not possible that I could have been born a year before I was, or that Richard Nixon was not once the President of the United States.[10]

Physical "possibility" we may understand as a subset of metaphysical possibilities, which track particular laws and systems of laws, albeit in different worlds which adhere to those laws. Thus, there is no general physical "possibility", but rather world-specific physical "possibility" systems. Generally though, when we refer to a physical "possibility", we mean implicitly physically possible *for this world.*

The set of metaphysical possibilities is wider than the set of physical possibilities, because the physical laws of a particular world could be such that a metaphysical "possibility" could not realized in that world, and also that a metaphysical "possibility" in that world could re realized but not realized in our world. What we want here is to be

[10] Although I may be able to imagine being born in 1981, I know that it is necessary that (not possible that I was not) I was born in 1982 (given certain features of the timing and distribution of genetics, etc.) Likewise, I may be able to imagine that Nixon was never president, but when I examine the relevant sortals of features I ascribe to Nixon's identity, he is hardly the same person if not a president.

able to think about a world in which, say, it is possible to travel at the speed of light, because thinking about such scenarios helps us to get straight particulars of what changing a Natural Law might entail. To consider why this is the case, let's consider an objection to metaphysical possibilities which outrun nomological ones, posed by Sydney Shoemaker.

Shoemaker (1998) objects to the suggestion that metaphysical "possibility" can outrun nomological "possibility" on the grounds that those "metaphysical possibilities" which are not nomological possibilities are actually merely conceptual possibilities masquerading as more robust inferences:

> I say about [non-nomological metaphysical] "possibility" what I said above about the "possibility" of scientific hypotheses that are not nomologically possible –there is no reason to think that this is anything other than epistemic "possibility", relative to some envisaged body of background knowledge. So regarded, it gives no support to the idea that "imaginability" gives us access to metaphysical "possibility" that outruns nomological "possibility". It may further be observed that were this epistemic "possibility" to be realized then the "imaginability" of phenomenal states of affairs would establish nothing more than their epistemic "possibility". There is another sort of "possibility" that "imaginability" might be said to give us access to, namely conceptual "possibility".

It is unclear what could be meant by "metaphysical "possibility" beyond the adherence to physical laws combined with particular states of affairs. In order to tease the two apart in a particular thought experiment (say, about whether something metaphysically impossible is actually physically possible, like my birth in 1980 instead of

1982) one will need to be very careful in specifying just how similar and dissimilar a particular world is, when guided by the same physical laws.[11] Thus, it seems as if metaphysical possibilities must all be nomological possibilities. Shoemaker is certainly right that such inferences must be made with reference to some envisioned background knowledge, but whether "envisioned" comes to mean imagined or conceived will be the rub. Undoubtedly, to conceive of a world in which one or more Natural Laws are different than they are would be a difficult task, if one wants to do it in anything beyond a schematic way, without details or clear relevance in our world. One would have to parse which Laws are dependant and how upon which, and what might reasonably stay the same in such a world, but unless we want to give up upon conjecturing about the nature of Natural Laws not yet discovered or confirmed, or what might change if said laws are different or change themselves, we had better not give up entirely upon these problematic and less reliable inferences. In a sense, Shoemaker is right to call these inferences inferences to conceptual possibilities, since they do stand in relation to a body of assumed background knowledge about how features of a metaphysically (but not nomologically) possible world might be. We don't *really* know what a world without friction would look like in its entirety, and so thus inferences about such scenarios will be a bit more haphazard than their nomologically possible cousins. But this is no reason to give up entirely on using them probatively, since sometimes, as in the case of the frictionless

[11] This view of the co-extension of metaphysical and nomological/physical "possibility" I believe is shared by (Shoemaker 1998), in which he argues that the essence of a particular property is constituted by it's causal features, both forward-looking and backward-looking. Shoemaker argues that the possibilities of different worlds be the same as the possibilities of a thing in this world, and thus that genuine possibilities are only revealed to us by what we can conceive, rather than imagine.

plane, we have a good grasp on how breaking a Natural Law effects the overall results of the experiment, while in other cases (moving at the speed of light?) perhaps the more cautious and reserved conclusions of a heuristic are warranted.

Since "conceivability", not "imaginability", must prima facie be an indicator of "possibility," and since "conceivability" must involve the conception of rigorous theory, which is presumably in conjunction with the assumption of regular, or natural laws, then what would it look like to conceive of a world, which is not a nomological "possibility"? We would have to specify very carefully what the laws of the world were, and how they could in fact be different, while presumably some of the world's facts were the same. Such a conceiving would be very, very complicated and of very limited use, I think.

***"Conceivability" as a Guide to the Different Types of "possibility"***

How effective a "conceivability-to-possibility" inference within a virtual experiment will be will depend upon what reliability constraints are placed upon the inference, both by the subject matter which it is about and by the type of "possibility" to which it is meant to infer. Different types of "possibility" have different sets of laws and truths to which they must adhere. First, let's consider logical "possibility":

> **Logical "possibility"**: Scenario S is *logically possible* in system *L* if and only if the description of S is consistent with the axioms of system *L*.

Conceiving as guide to logical "possibility" will require that the theory formed be consistent with the axioms of the system. Thus, if I want to investigate a "possibility" within a particular system *L* I had better be sure that the theory I am conceiving as a guide to "possibility" does not violate the laws of system *L*. Presumably, knowing that my theory does or does not violate the laws of system *L* within a definite system ought to be fairly straightforward, because of the explicit nature of the logical system and the knowledge that it pertains to everything possible in this world (and any possible world). Thus, it seems that "conceivability", as a guide to logical "possibility" will be a very reliable, very verifiable and relatively uninteresting sort of inference. Since such a vast array of situations are logically possible in system *L*, these inferences tell us very trivial things about what specifically may be possible in this world or worlds closely related to it. Consider Galileo's Falling Bodies thought experiment. Remember that this thought experiment shows a physical impossibility (and thus a physical necessity) by way of a logical impossibility: a contradiction. The tied together balls cannot fall both faster and slower because physical laws abide by logical ones (there are special problems here about my account when in discussions of scenarios conceived in quantum mechanics setting – there are special problems there discussion of which would be beyond the scope of this investigation). But of course, the logical inference is not the interesting one; it is the base of the probative physical inference, which shows that, by containing a logical contradiction, Aristotelian physics contains a physical contradiction.

Embattled as a description or definition of conceptual "possibility" may be, there are some tentative things to be said about it, if Vaidya's definition is to be accepted.

"Conceivability" as a guide to conceptual "possibility" inferences will hinge upon the full understanding and spelling out of the particular concepts involved. Thus, if I want to explore the "possibility" that the mind is a separate thing from the brain, and not ~~ir~~reducible thereto, I need to have a fully spelled out concept of the mind (all the necessary and sufficient conditions for being a mind- that is, a full and complete definition, if you like) and, likewise, for the brain. These definitions will most likely be at least partially incomplete (and possibly contradictory between accounts of the concepts) so the reliability of my inferences to conceptual "possibility" will be limited by the preciseness of the definitions of the concepts themselves.

"Conceivability" as a guide to conceptual "possibility" is the sort of inference one finds in a thought experiment meant to pump intuitions, i.e., in heuristic roles. Presumably, conceptual "possibility" tracks the set of things, which we may conceptualize with relative clarity (perhaps have a reasonable theory about). Thus, a "conceivability" to conceptual "possibility" inference may guide us to see whether a situation is something which we can actually conceive of with reasonable definitions of the concepts involved. These inferences will be largely heuristic, serving to iron out internal features of "conceivability", but not moving beyond it.

While physical "possibility" may be constrained by logical "possibility", it is not constrained by conceptual "possibility", despite the fact that both inferences to logical and conceptual "possibility" are heuristic. Thought experiments like Kripke's C-Fiber Firing show us that examining what our conceptual systems make possible may lead us to accept physical impossibilities, and that some physical necessities are conceptually impossible. As I will discuss in more detail in the next chapter, conceptual "possibility" inferences can be extremely helpful in generating and modifying hypotheses, when properly interpreted. They are excellent epistemological tools and poor ontological ones, since the set of metaphysical possibilities surely outruns conceptual ones, and visa versa.

Inferences to metaphysical and physical possibilities are, by contrast, of great ontological interest, and it would be a great boon to probative thought experiments to spell out their success conditions. However, distinguishing the set of metaphysical necessities from physical necessities, nomological necessities and logical necessities has been a subject of great debate in the literature. Further frustrating, our tentative definition verges upon circularity:

**Metaphysical "possibility":** Scenario S is *metaphysically possible* if and only if a description of S is true in some metaphysically possible world.

This definition serves to reiterate the basic assumed modal structure of "possibility". Metaphysically possible worlds are used to examine what is metaphysically possible in those worlds. Unless one believes that the non-actual metaphysically possible worlds are somehow real metaphysical entities (and not mere epistemic tools for thinking about metaphysical possibilities), one cannot define metaphysical "possibility" in terms of metaphysically possible worlds alone; this is circular. We will need an idea of what we mean to count as metaphysical "possibility" before it makes any sense to talk about metaphysically possible worlds.

So what counts as metaphysical "possibility"? Some have thought that metaphysical "possibility" is co-extensive with logical "possibility", but since there are a vast number of logical possibilities that are certainly not metaphysical possibilities in any more robust sense, metaphysically possible worlds must be taken to be a subset of logically possible ones. They are a proper subset, and we need a marker that puts some of the worlds in a subset, and one marker, following Saul Kripke, might be essences. Perhaps there are words which designate rigidly, that is, which pick out the same objects, processes, or states of affairs in every possible world, and imply that those objects, processes, or states of affairs are identifiable because they adhere to a set of essential features. Maybe then, inferences to metaphysical possibilities are inferences which query the set of worlds in which essential features are preserved where rigid designatators pick out objects, processes, or states of affairs.

But there is a serious circularity objection to this program of defining metaphysical possibility: on the rigid designator program essences are queried via modality, making it circular to then query modality through an appeal to essences without further explication. Kit Fine outlines some serious problems with the definition of essence in terms of modality (Fine 1994). Fine argues that necessity is a special kind of essential feature. Furthermore, Fine demonstrates that essence is not a reflexive feature, and that reflexivity would be needed to avoid circularity in querying metaphysical necessity and essences through mutual association. Consider Fines example: it is necessary that Socrates, if he exists, belong to the singleton Socrates, but there is nothing about Socrates' identity that makes the singleton Socrates essential *to him*. Likewise, it is necessary that Socrates and the Eiffel Tower be different things, but there is nothing essential *to* Socrates *about* the Eiffel Tower, etc. Let's consider another example. Surely, a particular token genetic makeup is an essential feature of being a dog. Thus, every instance of a dog will contain that genetic makeup. Additionally, the genetic makeup of a dog will never enable speech. That is, it is essential that dog genes do not afford speech. However, there is nothing essential about speech which makes it not appear with dog genes. While ¬speech is essential to the dog genes, ¬dog genes is not essential to speech. Thus, not all essences are reflexive, but surely all necessities are.

According to Fine necessity is "a highly refined version of [essence]; it is like a sieve which performs a similar function but with a much finer mesh" modality captures those essential relations which are reflexive in that what it necessary to A about B is also necessary of B (in relation to A). Rather than think of essence in modal terms, Fine, "argue[s] that the traditional assimilation of essence to definition is better suited to the task of explaining what essence is." I agree with Fine that essence had better be defined rather than explored with modality, because definition allows us to make two distinctions. First, it allows us to distinguish between instances of a particular thing or property that are token, rather than type identical. It does so because a definition can make reference to a class of things being present in different instantiations, rather than a specific instantiation (which will have many accidental features we don't want to treat as essential ones). Thus, it is essential (perhaps) that a human brain have a neocortex, but it is not essential that it have *my* neocortex. Similarly, definitions allow us to distinguish between identity and sameness (with respect to particular sortals of features). Let it be enough to assert that there are some essences which are not the reflexive kind we can identify with necessity. What really matters for discovering essences, Fine argues, is the origin of the property or thing in question. By this, Fine means the subject matter in which the thing to be defined appears. For instance, a physical entity had better be defined in physical terms, a mental property in mental terms, etc. Of course, one result of adopting this view is to realize how metaphysical possibilities (with regard to essence) might be limited by conceptual possibilities, epistemically, but not ontologically.

It should be clear now that metaphysical "possibility" is by far the most embattled kind of "possibility" about which philosophers speculate. We have seen that "metaphysical possibility" may mean something in the context of talk about essences (in very particular Finean ways), but that essences had better not depend upon modal inferences. Likewise, we have seen that metaphysical possibilities may be defined partially by necessities about the past as well as a set of possibilities about the present and future restricted by either the laws of this world (nomological "possibility") or the very specific and restricted laws of another world, with very limited inferential power. Rather than insist upon a particular definition, then, I think this discussion has been productive in outlining what the consequences of using different definitions of metaphysical "possibility" for "conceivability" to "possibility" inferences.

Perhaps we can assert this:

> A scenario is metaphysically possible iff the entities mentioned in a description of the scenario could be configured the way the scenario configures them without it being the case that we have to assume that there is change in the essence of those entities. For the moment let the issue of metaphysical "possibility" and essence lie there. At the very least we should see the consequences of adopting particular views of metaphysical "possibility" for our inferences.

Some of the candidate markers of metaphysical possibilities are constraints that come from facts about Natural Laws (Swoyer 1982), constraints that come from the fact that some facts are in the past (Kneale 1938) and of course constraints that come from the fact that some properties are essential properties (Kripke 1980). Swoyer and Kneale's contributions I mention in order to sort out a working conception of what metaphysical "possibility" might be accessed through "conceivability".

Swoyer argues convincingly that Natural Laws define what is metaphysically necessary or possible. If this is the case, and if we are committed physicalists, than physical possibilities will be extensionally equivalent with metaphysical "possibility", unless of course we can identify some unproblematic examples of metaphysical possibilities which are not physical possibilities. Kneale argues that past events are necessary in a way which is not logically necessary (it is not logically necessary that I was born on March 31rst, 1982), but we are tempted to say it is a metaphysically necessary. It is also not a physical necessity that I was born when I was, but it does seem to be necessary that I was born within a few months of when I was. So it would seem that if there were metaphysical necessities, which are neither logical nor physical possibilities, that there should likewise be metaphysical possibilities, which are neither logical nor physical possibilities: namely, possibilities about how the past might have turned out. Once again, if we are to specify a metaphysical "possibility" which is neither harmonious with physical nor logical "possibility" (of a certain logical system) than we had better be very, very sure that we know how we can track intuitions into such a strange and unfamiliar possible world. Likewise, if we want nomological possibilities which exceed the metaphysical possibilities of this world (or visa versa) we had better specify them as we specify logical possibilities, with reference to a very definite and limited system. Alas, I fear that inferences within these worlds will be of limited value beyond sorting out inconsistencies within those definite, artificial (or at least highly hypothetical) and strange worlds, and possibly mere conceptual possibilities.

***"Conceivability to Possibility Inferences" and Virtual Experiments:***

Let's consider where we've been so far:

Since all virtual experiments query possibilities of one sort or another, it will be imperative to discovering the success conditions of virtual experiments to identify which type of "possibility" a particular virtual experiment means to query. Since the scope of this work is mostly limited to virtual experiments in the sciences, metaphysical "possibility" and physical "possibility" will be my primary targets. Of course, logical "possibility" and conceptual "possibility" also play key roles in heuristic thought experiments, but since their definitions are less embattled, determining whether a virtual experiment successfully tracks a logical "possibility" or a conceptual "possibility" is a simpler task. Let's turn our attention, for the moment, fully to metaphysical/physical possibilities as they are investigated through thought experiments.

Since many virtual experiments take the form of cut-and-paste investigations (i.e. they aim to investigate whether one thing can or cannot exist without some other state-of-affairs), metaphysical "possibility" inferences in scientific virtual experiments often aim to predict whether a particular phenomenon can occur with or without a particular condition. This is certainly the aim of Kripke's C Fibers Firing thought experiment: can we imagine pain existing without a particular physical base, and is this grounds for assuming that pain can in fact exist without a particular physical base. Since virtual experiments which investigate physical possibilities about this exact physical world in which we live investigate a very narrow subset of possibilities, and do so under very restricted conditions, it is imperative that one know the conditions to which one must restrict one's experiment, just as one investigating a logical "possibility" needs to know all the restrictions of a particular system within which one's inference runs. The problem with this is that our knowledge of this physical world is anything but complete; knowing all the properties of our physical world (i.e. a complete physics) is not within our grasp. Therefore, when one conducts an experiment about this physical world, one tries to severely limit the variables which may have significant effects upon the outcome. For instance, if I want to investigate the breathing patterns of dolphins, I had best do so in a terrestrial body of water, since an imagined Martian body of water may contain unknown differences which could skew my results. Thus, the more of the actual materials/ conditions/properties of the scenario I want to investigate are present, the more likely it is that I have not included or excluded a relevant feature. We set up our experiments to be as similar to the conditions we wish to query as possible in order to replicate the conditions we want to produce certain results. However, that is not to say that the more of the actual conditions of a scenario I include, the more will de facto be relevant. Perhaps I notice that a group of non-verbal autistic children grow agitated in a particular classroom. I want to discover why this is the case. I suspect that it is because of the fluorescent lights in the room. I only need to produce the requisite conditions (flickering lights) to investigate whether the students have increased sensitivity to them. I may do this either by removing the lights from the first room (cut) or placing the students in a second fluorescently lit room (paste). I do not need replicate the room in it's entirety to discover if this is the case, I merely need to see if the students grow agitated in other similarly lit rooms. There is no need to hang the same posters, or produce the same carpet in the new room, because I am pretty sure these are accidental conditions. While there remains a faint "possibility" that the students were in fact upset by motivational kitten posters, I know that these are unlikely to have been the cause of their distress. Thus, the predictive value of a particular experiment does not hinge necessarily upon the proximity of recreation; rather it depends upon the presence of, and whittling down to, merely the candidate success conditions' presence or absence.

The more I know about the conditions under which are particular phenomenon appears, the more likely it is that can produce it without recreating its original conditions. This, I think, accounts for R. A. Snooks' observation that there are a vast number of thought experiments in physics, but relatively few in chemistry; chemical reactions happen under very specific conditions, while physics investigates properties which hold constant throughout vast areas of the universe. You need to know some very specific information about a very, very specific physical world in order to run a thought experiment in chemistry and have it be genuinely predictive. However, as I have hinted already, if a virtual experiment need not be genuinely predictive (and indeed not all experiments themselves are), than there are indeed heuristic thought experiments in chemistry. In the following chapter, I discuss the distinction between heuristic and predictive thought experiments, with relation to the bodies of knowledge required to run them (their domains) and the types of "possibility" they query.

## Chapter Three: Inferential Force in Virtual Experiments: Heuristic and Probative Value

In chapter 1, I argued that one consequence of recognizing the spectrum of virtual experiments is a necessary revision of naïve assumptions about the probative power of experiments. Additionally, I hinted that with regard to success conditions of experiments, there is a useful distinction to be made between probative and heuristic experiments, wherein at a first pass, probative experiments make robust inferences about states of affairs directly, while heuristic experiments serve as a tool for the generation of hypotheses which may or may be later tested in experiments or otherwise. Subsequently in chapter 2, I examined how virtual experiments depend upon conceivability to possibility inferences, and how those inferences to different kind of possibility (physical, and metaphysical versus logical and conceptual) warrant probative and heuristic inferences, respectively. In this chapter I will spend some time unpacking the distinction between probative and heuristic experiments generally, while acknowledging that this distinction is just the first in a taxonomy of the inferential powers of experiments (and their related roles in successful sciences).

A very serious objection that has been leveled against the relative success of thought experiments is that there are no clear success conditions for thought experimentation. If we accept that, as I have argued in previous chapters, thought experiments are the limiting cases of the spectrum of virtual experiments, then it becomes recognizable that all experiments involving at least some degree of abstraction and/or idealization (or even just experimental control) will should be be prey to skepticism of the same kind or other. One plan of attacking the problem has been to investigate what particular virtual experiments are meant to achieve, and, of course, what they do manage to accomplish. Along these lines, I have discussed how virtual experiments rely upon a variety of underlying conceivability to possibility inferences, which in turn partially determine the inferential probative force of an experiment.[1] As previously discussed, there are many strategies for investigating what reliable success conditions for conceivability to possibility inferences might be identified, but here I wish to focus on one particularly revealing idea for how think this problem through for scientific virtual experiments: the distinction between inferences within relatively well-defined and less well-defined or, perhaps, less "refined" domains. I elaborate on a distinction originally made by R. J. Snooks in reference to the difference between thought experiments in Chemistry and in Physics: the degree of definition of a subject domain for a thought experiment will determine the success of said experiment (depending also on the goals of the experiment) (Snooks 2006). Using the concept of a relatively well-defined domain, I distinguish between two categories of experiments with different investigative aims – heuristic and probative experiments: these as individuals and as classes have differing probative force and thus different success conditions.

The success conditions of experiments will depend of course in part on what one expects success to entail. The paradigmatic experiment (in many people's minds) gives us reliable knowledge directly about the world. Erroneously, we sometimes want experiments to *prove* something to be true or at least illuminate something – in our more

[1] There is great debate about what exactly even happens in conceivability to possibility thought experiment, and the meanings of conceiving and possibility. I set this debate aside here because I think even a naïve understanding of these terms will suffice for discussing the present case.

idealistic moments we want experiments which result in a "eureka!" We would like a clean narrative about discovery, in which clever observation and testing result in the discovery of startling and lasting truths. But this is not how science works (at least not most of the time). Vast amounts of corroborating evidence, experiments, models, theory revision etc. go into the testing and refining of a theory like Natural Selection, Neuronal Plasticity or even Newtonian physics.[2]

A great deal of scientific effort goes into experimental design and regulation. Experiments with experiments, as I might put it, are among a set of tools available for

[2] There may in fact be a separate class of processes by which one ends up in a eureka or 'gotcha' realization and with a hypothesis to be tested. Call these set (a). If this is the case, then set A must be distinguished from the set of processes which involve and maybe are designed to check some hypothesis already formulated but which do not have the eureka-effect. Call these set (b). Furthermore, there may be processes which involve and are designed to refine hypotheses. Call these set (c). Experiments can be involved in (b) and (c). Experiments *may* but may not be involved in (a). Gotcha can occur in (b), (c) and (a), but some "gotchas" which feel paradigmatic may be (a) rather than either (b) and (c). (a), (b) and (c) may be useful in other contexts of distinguishing discussions of (i) methods of discovery and (ii) methods of validation/proof/checking/corroboration, etc., and their negations. Methods of discovery *may* be involved in the context of hypothesis discovery – they *may* work also with hypothesis checking. An interesting discussion and worth further thought, but for now I set it aside perhaps to be later examined in particular cases

designing, inventing, and refining experiments. Since there is no method of discovery for when inventing an experiment, scientists often must use guesses and guidelines that might help them to think of experimental designs.

So experiments often beget experiments. When Ian Hacking famously noted that experiments can "have a life of their own", he illuminated the way in which experimentation involves a lot of testing and regulating of experimental methodology and apparatuses. You can do some experiments in order to prove that some experiments have probative power or to prove they need refining. Thus, within the class of probative experiments, the probative power of some of these probative experiments is second order in that they test/probe experimental design.

The vast array of needed corroborative steps discussed above tends to spread the eureka-effect rather thin, making it so that few experiments can (or should) aim to prove beyond further confirmation their hypotheses. Additionally, we know that, in addition to "harder" data-driven methods, there are seemingly irreducible models, metaphors and methods. This mess is often relegated to a holding pen with a vague and sometimes dirty word: "heuristics". In this chapter I want to pluck a subset of virtual experiments out of the slighted heuristic holding pen and save for them the title of "experiment" against the charge that because they can be said to hold heuristic value, they are "mere heuristics." But I'm getting ahead of myself: let's begin my considering when, how, and where virtual experiments with less-than-eureka-effect discoveries occur. I turn my discussion to the tendencies of the most beleaguered of virtual experiments: thought experiments.

In an 2006 essay, R. J. Snooks tracks and describes a trend in scientific thought experiments: there seem to be a plethora of thought experiments in physics and relatively few thought experiments in chemistry. Because Snooks' considers thought experiments to aim at predicting and ultimately discovering truths about the physical world, Snooks finds that these predictive thought experiments only appear in mature sciences, like physics.

Snooks suggests that this is the result of the relatively well-defined domain of knowledge in physics -- the particular feature of the maturity of physics that is operative here is the fact that as a mature science, physics forms a well defined domain of knowledge; particular success conditions of thought experiments, like regular natural laws, as well known in physics. Furthermore, physical laws can be expected to be 'realized' in similar ways from situation to situation, making it easy to abstract them into idealized thought experiments without fear of leaving out relevant conditions. As a settled science, many areas of physics probably have fewer spaces in the theory, allowing thought experiments more settled ground to rest upon.

Chemistry, on the other hand, involves very specific processes under very specific and localized (for our universe) conditions. We just can't know all of the success conditions for a particular instantiation of a chemical reaction, because we only know about it in the highly specific yet less well defined, less refined domain of our planet's physical set-up. We can't yet predict reliably about what happens to these reactions under very different circumstances (say, in a black hole, etc).

I follow Snooks in using the useful term: "domain". The domain of a thought experiment is the knowledge base that one would expect an experimenter (and his or her audience) to bring to the consideration of it.[3] A domain encompasses the class of phenomena that a certain knowledge base concerns. A knowledge base will include both a set of well-established hypotheses and also a set of experimental procedures together with a set of views about what those experiments reveal, given the hypotheses. Additionally, this may include the regular or natural laws assumed, the specific possible effects assumed not to be present (friction in physics, etc), and a variety of other features of the target phenomena. Thought experiments are prevalent and highly productive in physics because they work within a relatively well-defined domain with relatively reliable natural laws. Thus, they are able to be highly predictive. Thought experiments in chemistry, on the other hand, are virtually non-existent, and when they do appear they serve more heuristic, rather than predictive functions. Snooks argues that this is because chemistry is a relatively less well-defined, less refined domain with more unknown variables for particular instantiations of phenomena. That is to say, chemistry takes place in highly specific (for the universe) environments in which we do not know precisely why certain conditions hold or how they hold, and the laws associated with these very specific interactions rely upon these highly specific and less well-defined, less refined conditions. This makes it very hard to run predictive thought experiments.[4]

[3] Here perhaps the notion of textuality is helpful in understanding the defuse boundaries of a domain. The domain of a thought experiment, much like the Text, extends well beyond what is presented on the page, but includes an assumed familiarity with the subject, various cultural influences, etc.

[4] It many also be the case that physics is a science in which, because of the nature of the phenomena, thought experiments are all that we *can* do by way of experiment. I

One result of fully embracing Snooks' view that thought experiments in chemistry are only of limited predictive value is that one must concede that they can only exist in subjects and disciplines with very highly refined domains. Yet, thought experiments appear throughout the sciences and are enormously helpful, without necessarily predicting results in the way that some thought experiments in physics do. Thought experiments in such less refined and highly specific domains have considerable heuristic value; they act as excellent intuition pumps and generators of hypotheses.[5] Good heuristic thought experiments also make clear that they require the assumption of tendentious success conditions and black box mechanisms without which they are not reliable intuition pumps *or* guides to possibility. Predictive thought experiments, because of their relatively well-defined domains (and regular laws) do not need to assume tendentious claims, as the success conditions are ideally laid out before hand in the agreed upon reliable domain. Controversy over interpretation of intuitions and of what (if anything) heuristic thought experiments can reveal is bound to appear, but the heuristic thought experiment makes these conditions explicit, because the domain cannot assume them.

What makes conceiving of something within a relatively well-defined domain a

will argue that this is true of some phenomena in some sciences, in particular climatology, in Chapter 4.

[5] There has been much discussion of what a heuristic, in general, might be, both in the sciences and in philosophy. Here I use the term to denote a thought experiment, which clarifies intuitions and generates hypotheses, which can then be tested or used predicatively in our world. A heuristic thought experiment points us in the direction of possibilities, without asserting them.

more straightforward guide to possibility? As explained in the previous chapter, I follow Stephan Yablo in the assertion that "conceiving" means "imagining" something, *given everything you can know about it* (Yablo 1993). If you can imagine of something clearly and within a relative settled body of science, then this can act with relative reliability as a *prima facie* guide to possibility. You can, on this model of conceiving, end up being wrong, if you don't consider or know about a relevant feature of the domain under consideration. For instance, if I think I can imagine a tiger with spherical stripes, perhaps I am missing some relevant aspect of the definition or on the current theory of stripes, which renders this an impossibility. The more defined or refined the theory covering some domain, that is the more you know about the particulars at play in the domain you are imagining within, the more likely it is that you aren't missing a relevant or necessary feature (like a reliable natural law), and thus the more likely that you are perceiving something which is a genuine possibility.

We might be tempted here to arrive at a hasty conclusion. If we accept the evidence that highly virtual experiments have probative force within highly refined and well-defined domains, than we may shrug our shoulders and say something like this:

> Probative value is real experimental value. It's only possible in mature sciences with relatively well-defined/refined domains. So much the worse for thought experiments and their near relatives, other highly virtual experiments, ought to be restricted to highly defined domains or not be used at all. So much the worse for young sciences, which must muck through data in the field, if their adherents want any semblance of respectability. Observation first, speculation next.

But this conclusion jumps the gun. In fact, it violates its own advice, drawing a conclusion about virtual experiments and the young sciences hastily before observing them in action. To put it another way, ask first: how does a domain become refined and/or well-defined?

The background knowledge which calibrates domain definition is the result of a build up of experiment upon experiment upon experiment, each building upon the reliability and testing of previous experimental regularity all of this within well developed hypotheses, hypotheses framed in terms of precisely specified laws of nature.[6] Confirmation and disconfirmation serve to make results more likely. The more scrutiny a hypothesis undergoes, especially from multiple independent methods of confirmation and especially within well substantiated hypotheses both hypotheses about the phenomena in the domain and hypotheses about why the experiments work as they do, the more likely it becomes and thus the more relatively confirmed. Disconfirmation serves to suggest changes in experimental design, in theory, and sometimes in background assumptions. What I am describing is the domain-wide tinkering in a particular science, which serves to define the domain; tinkering occurs in theories, in experimental design, in hypotheses.

So conceptual possibilities and logical possibilities, and other tinkerer's inferences serve to define the domain of a science, and they are every bit as vital to good science as robustly probative inferences. Call this second class of experiments which employ these inferences heuristic experiments. I think recognizing this category of experiments, and their vital role in the development of the domain of inquiry for a specific science goes a

[6] To simplify the process very considerably and leaving aside various complexities, including of course the reality that often in investigations there are various serious set backs.

long way towards revealing and legitimizing the role of virtual experiments performed in sciences with relatively less well-defined, less refined domains. Rather than failing to recognize the relevant features of an experiment's domain, a heuristic experiment (in the way I will use the term 'heuristic experiment') does not assume them, but investigates them. Let's consider some examples which reveal the utility of heuristic virtual experiments.

**1. Heuristic Experiments may induce/ reveal the need for major fundamental theoretical shifts.** Abstraction and idealization in virtual experiments can isolate a particular aspect of a theory, acting to place focus on a possible and maybe fundamental conceptual error in the theory. Consider, by way of example, the overriding thought experiment of Stephan Jay Gould's book, "It's a Wonderful Life." Gould suggests that, if we imagine rewinding the VHS tape of the history of life on earth and then playing that tape forward, it is very likely that minor changes in past events will reveal a vast number of contingencies in the development of life forms. The recognition of these contingencies in the fossil record (Gould asks us to consider the wildly alien and extinct creatures of the Burgess Shale fossil record, wiped out by catastrophic events) reveals a major inconsistency in the adaptationist tendency to assume that organisms are composed entirely particular traits, each with a definite function, selected for overtime. Historically contingent evolutionary paths reveal that some creatures go extinct because their populations are wiped out by freak catastrophes, rather than failure to adapt. Thus, the fact that populations extinguish in this way refutes the assumption that the traits we observe represent "the best of all possible worlds."

**2. Heuristic Experiments can be the mechanism of ongoing experimental tinkering-** Ian Hacking has noted that experiments can have a life of their own beyond theory confirmation – they can serve to improve experimental design, by, for example, calibrating and perfecting instrumentation or fixing bugs in simulations and in the use of models. It is difficult, and probably counter-productive, to draw a firm line of distinction between tinkering and "actual experimentation," because tinkering may also reveal features of the target phenomena, prompting further experimentation and providing the first bits of evidence for robustly probative inferences and theory confirmation. They can of course be probative of some experimental procedure (e.g., a technique for checking for trace elements in water without doing high tech lab word (feed it to a certain plant to see if it changes color). These constitute our second-order probative experiments; a set of hybrids which also contain both probative and heuristic value.

**3. Heuristic Experiments can serve to sketch out what will be needed to further define the domain-** Ongoing tinkering, as I suggested above, is the method by which experimental methods, theory and other important aspects of a young science are developed and confirmed, and thus, heuristic experiments which tinker serve to define the domain and determine relevant features of the target phenomena as part of that domain.

4. **A Heuristic Experiment may suggest or even model possible explanatory gaps-** some virtual experiments, such as simulations and the use of models, import metaphors, which supply possible explanatory structures, which can then lead the search for the actual mechanisms by analogy. Purportedly. Watson and Crick's physical double helix model suggested itself from the physical constraints of the nucleic acid molecule models.

The above is only a tentative shortlist of some of the valuable and sometimes invaluable uses of heuristic experiments in relatively less well-defined, less refined domains. Difficult as it is to come up with a unifying, regulative definition of experimentation, it would be foolish to think that this chapter (or even this thesis) could give an exhaustive list of the productive roles of heuristic experiments in the vast variety of young and specialized sciences. Suffice it for my purposes to demonstrate that they are vastly prevalent and hugely important in the development of a science.

Heuristic value and probative value are not mutually exclusive, or even necessarily sharply distinguishable. The aim of the distinction is rather to assert the existence of experiments which have other useful functions besides directly probing a truth about the target phenomena. After all, a logical impossibility is a physical necessity, and thus a heuristic inference may also carry probative force; a thought experiment in evolutionary biology may reveal a conceptual inconsistency in adaptationist account, which in turn reveals the impossibility of said theory mapping onto the world.

One major advantage of the recognition of heuristic experiments, especially the recognition of heuristic virtual experiments, is the explanatory power the category gives to heuristic thought experiments. As I have noted earlier, there exists considerable debate among philosophers as to what thought experiments are and whether they can be considered genuinely experimental. A subgroup of thought experiments, such as those Snooks considers in physics, can be argued to have the robust probative power we usually associate with more traditional notions of scientific experiments, but there are close relatives of these thought experiments which arise in less defined domains and which seem instrumental in the development of younger sciences. Gould's rewinding the tape of life is one such example. Recognizing heuristic value in virtual experiments allows for an account of the role of thought experimentation within specialized and young sciences.

Furthermore, the heuristic/ probative inference distinction allows that thought experiments in particular, and virtual experiments in general, have value even if their results remain unconfirmed or after they have been disconfirmed. This is particularly important in understanding thought experimentation as a vital part of the history of the philosophy of mind, where empirical experimentation has only been able to confirm or disconfirm virtual experiments very recently.

Finally, the heuristic/ probative distinction sheds valuable light upon the tinkering particular to the development and running of vast computer models, especially those which not only crunch data but also generate predictions. By viewing the on-going tinkering development of vast simulations as heuristic experimentation fading into probative experimentation, the history and development of simulations such as climate models is illuminated, and skepticism about the reliability of such experimental methods faces robust challenges.

A final literary metaphor on the utility of heuristic experiments. When a great work of literature is written, an author (or authors) generate(s) something startling and new. However, that great work more often than not arrives to us after the author collaborates at length with an editor (or editors) who amplify, crop, modify and enhance the power of the work. The relationship between author and editor is often deeply creative, sometimes generative, and often ignored by a society interested in easy-to-digest instances of "flashes individual of brilliance" and genius. Just as editor assists author to create a text, so heuristic experimentation (indeed, perhaps all irreducible heuristics) serves to suggest, edit, defeat, augment, and often pave the way for probative inferences in experiments. A defined domain, a mature science, takes a lot of editing.

But of course, virtual experiments are productive in young sciences as well. Furthermore, as I will show in Chapter 4, young sciences may employ them probatively, so long as they are utilized as part of a methodology which grants them probative power and which involves rigorous and systematic reviewing of methodology. Moving forward, let's consider how climatology, a science less than a hundred years old, is able to utilize models and simulations with probative force.

**Chapter 4: The Modal Apocalypse, Now: Climate Simulation Models and Virtual Experimentation**


1

In this final chapter I will motivate and ground my account of the spectrum of virtual experiments, and their success conditions, by giving an account of how virtual experiments appearing in a young science serve to define their domain and to meet the success conditions of probative inferences. The subject will be climate modeling in computer simulations. The aim will be to show how climate models represent accurate and robust representation and experimentation within a vast system. Along the way, we will see how looking at the way vast simulation computer models work within modern sciences demands an account of experimentation and the relationships between data, model, theory, and world, which enables genuine virtual experimentation.

[1] Image courtesy of http://www.toothpastefordinner.com/122510/laymans-terms.gif.

### *Introduction:*

If one wants to take climate modeling, and its discoveries and conclusions seriously, one must revise old accounts of the relationship between model, world, data, and theory. Not surprisingly, climate change skepticism has focused on the very features of climate models which make experiments utilizing them count as virtual experimentation: the concerns raised inevitably challenge the view that modeling can be interpreted as experimentation, and that simulation can be interpreted as producing genuine knowledge about the world. These challenges are really challenges from a traditional view of data/model/theory/world relations, a vantage point from which the innovations of climate models as virtual experiments (including the symbiotic relationship between data and theory/model, the idealization necessary to conduct a simulation of a vast system, and issues regarding the reliability of the background knowledge needed to simulate that vast system) appear to be indications of shaky science. That is to say, skeptics think there is an explanatory gap between what climate models reveal and what their users claim to infer from them. However, within the scientific community, where virtual total consensus exists as to the reliability of the findings of these virtual experiments, climate models represent a monumental shift in the way legitimate science is perceived to relate theories data, models and probative inferences. This revision in turn changes what can count as an experiment, and more specifically what we ought to take to be sufficient definition in a domain to license probative inferences.

### *Climate Models as Virtual Experiments*

While you can observe interventions on the global climate, you cannot (practically, ethically or prudently) create them. Yet, establishing first the factuality, then the extent, and features of human interventions on the global climate have been some of climatology's main goals. Climatology must use virtual experimentation to study the real effects of unintentional human intervention upon a system which we cannot prudently manipulate. . Thus, climatology is in large part a science which relies upon virtual simulation and modeling.

Climate modelers have responded to these challenges by adopting and greatly enhancing simulation modeling methods from earlier meteorology models. Climate models have become increasingly complex as climatologists recognize the many kinds of inputs which affect the global climate. Models of varying complexity have been developed to meet the needs of particular investigations. Correspondingly, there are a variety of climate simulation modeling methods, ranging from relatively simple single dimensional radiant heat transfer models (which calculate the heat transfer on the earth as if the earth were a single point) to vastly complex General Circulation Models (GCM), which can include inputs from Ocean General Circulation Models (OGCMs), Atmospheric General Circulation Models (AGCM), Ice Sheet Models (for Greenland and Antarctica ice masses) and Chemical Transfer Models (such as a Carbon Transfer Model for Carbon Emissions) to create a more complete picture of global climate. These complex models are both used to input, to groom, complete, and to produce the data of climate science (Edwards, 2010).

There is no Twin Earth on which to run experiments in climate science, but there are computer simulations which model the entire system of the earth's climate in increasing detail. These simulations, modeled to accurately depict previous climate patterns and readings, can have variables changed within them to give us a picture of what might happen if a particular prediction or input where to change.

So, in addition to modeling the past and modeling the probable future, complex climate models allow us to simulate various possible earths. One should say that climate simulation models allow us to model a variety of closely related possible worlds as the future world, because climate predictions can (arguably) never be totally pinpoint precise, given minute fluctuations in the enormous system. What climate modelers do in order to produce a prediction of the future climate is produce a small spectrum of outcomes (for instance, within a set period of time the temperature of the globe will rise between 0.5 and 1.0 C°) by running prediction models for a set of likely possible variations. As climatology matures and models has become more sophisticated, this range in predictive accuracy has shrunk, but it is unlikely it will ever become a single, confirmable number; it will certainly never be completely certain (more on this later, but of course, no science does give us complete certainty!). What climate models can make us reasonably certain of is that our climate future will fall within a shrinking range of closely related temperature scenarios. Amongst the closely related simulations of possible worlds, we infer that our future world exists as one of a set of highly similar contestants.

So to sum up, climate simulation models are the way, for the most part the only way, to ask and test questions about the global climate's past, present and future, since direct experimentation on the target phenomena is for the most part not an option. Despite the youth of the science, and the unruly nature of atmospheric data and the vastness of the system for which and within which predictions are being made, they give us robustly accurate simulations of the past and surprisingly reliable predictions of the future. We can conclude, then, that computer simulations and analysis models of climate change fall on the spectrum of virtual experiments, utilizing and crunching data taken from the world, but not initiating interventions upon the target phenomena (select factors in the global climate), and that as long as we are comfortable with a small range within which predictions, they count as robustly probative.

Still, climate skepticism has often taken the form of skepticism about virtual experimentation in general, and about whether modeling and simulation can act as substitutes for intervention upon real phenomena and complete data, in part because of older accounts of the role of experimentation within the sciences and of what can account as probative inference in science (namely high probability rather than traditional notions of confirmation). But since complete data for such a vast system will never be attainable, and since hypothetical climate disasters will, it is to be hoped, remain hypothetical, many interventions upon the phenomena need to remain credible projections only. Thus, climate science must rely in part upon experimentation in climate simulation models, and in turn must rely upon ways of legitimizing the inferences of virtual experiments. Accepting climatology and climate modeling as legitimate science and legitimate experimentation requires useful explication and evidence of some core concepts in the account of the spectrum of virtual experiments. Below, I will discuss how climate models enrich and motivate accounts of domains, inferential power and confirmation.

***Climate Models and Data***

The Regan and George W. Bush administrations couched their aggressive campaign against climate science in the terms of data wars, arguing that climate models are merely models, producing their own data and predictions independent of "real data" and evidence.[2] Overt political agendas aside, concern about the relationship between data, model and prediction ought to initially arise within the framework usually assumed by traditional conceptions of data-model relationships and how climate models actually interact with their data. Climate models present interesting complications for previous accounts of data/model/world accounts. On a simple traditional conception of the relationship, mathematical models relate variables through mathematical relationships (Frigg and Hartman, 2006). They crunch "raw" data in order to test theories. If the output of the model matches the observed phenomena, then one concludes the theory fits the world.

But climate models typify a different relationship, one of model/data symbiosis.[3] Climate models incorporate data which is already theory laden, through the theory-laden sensors which record the data, through the assembly and modeling of data in simulations (which in turn produce the data of climate analysis models) and through the prediction of missing data points by simulation models. Thus, models produce data, and data shapes models. The metaphor runs thus: the relationship is symbiotic, rather than parasitic, because computer models assist data generation by checking for outlier data points (indicative of error) thus suggesting data checks and revision. Conversely, data constrains model-building, because simulations can be checked against previous data sets, both retrodicting the past climate patterns and thus predicting current and future patterns.

This acknowledged interdependency between theory-laden data and data-laden

---

[2] For a historical account of the policy wars, see Edwards' historical overview.

[3] I borrow this useful term from (Edwards, 2010), p. 281-282.

theoretical modeling warrants the advent of "Reproductionist" accounts of experimental science. Reproductionism is the account of data/model/theory world relationships which comes closer to matching the methodology of computer simulation experimentation than previous reductionist accounts on which models were merely heuristic expressions of theory, used for crunching "pure data". "Reproductionism" is climate modeling researcher Paul N. Edwards' term, so here is an account in his own words:

> [Climate Modeling reveals]...the necessity of parameterization, much of which can be described as the integration of observationally derived approximations into "model physics"...[Parameters] are sometimes referred to as "semi-empirical," an apt description that highlights their fuzzy relationship with observational data...Though this looks very little like our idealized image of science, in which pure theory is tested with pure data, that image was arguably always a false one...Complementing [reductionism], in computational sciences such as meteorology, a new ideal has emerged- Reproductionism seeks to simulate a phenomenon, regardless of scale, using whatever combination of theory, data, and 'semi- empirical' parameters may be required. It's a "whatever works" approach- a 'Pasteur's Quadrant' method that balances practical, here-and-now needs for something that can count as data, right up alongside rigorous physics (Edwards, 280-281).

The reductionism which Edwards refers to is the form in which higher order phenomena ought to (and eventually will) be accounted for only in terms of lower order phenomena, and in which models can and are simply reduced to mathematical expression of theory, crunching "pure data." Reproductionism in contrast allows that in sciences requiring simulations of vast systems, a more ad hoc method of interdependency: symbiosis will exist between data, model and theory, rather than asymmetrical development of one from another. This symbiosis may sound like circularity, but in practice it describes the feedback loop between model, theory and data which ensures the better expression of each (and in the big picture, the definition of the domain of each). Empirical and semi-empirical checks exist upon data, models and theories.

Data model symbiosis looks like this: because vast climate models depend upon many, many data points for accuracy, and because the vast set of sensors on the planet is both financially and physically prohibitive (they cannot be placed in some areas of the open ocean, for instance), climate modeling utilizes past weather and sensory patterns to fill in and predict, given patterns around these missing data points, their values.

Such vast models interpret, interconnect and generate data sets, which are then analyzed in analysis models, making the model/data relationship especially complex.

***Some Conclusions about Climate Models, Climatology and Virtual Experimentation***

There is broad consensus in the scientific community that climatology, although a young science done with simulations and a reproductionist ontology, is a respectable, reliable science (a science with the reliability and evidentiary respectability of sciences in which experiments are possible directly on the phenomena under investigation), and that it's inferences are robustly probative. Though sceptics may cry against simulations and the need to use them to generate data, the rigorous and respectable scientific community is embracing the inferences of these models as reflective and predictive of the real earth. As some of the most robust virtual experiments, climate models have much to tell us about the present and future of virtual experimentation, and perhaps that of scientific methodology. Because climatology requires extensive simulation and modeling, it has developed a sophisticated methodology for use of and the development and refining of these virtual methods. Let's consider a few lessons from the use of virtual experiments in climatology:

***The spectrum of Probative and Heuristic Inferences, in light of reproductionism and Simulation models of Vast Systems-***

The probative inferences of climate models, in the face of data shimmer, become inferences which converge upon robust knowledge about the world. But then, scientific inferences have always existed on a scale of probability of accuracy; they can be more or less "confirmed" (found to be highly, highly probable in comparison to alternative explanations).

Climate models depend upon a two-step modeling process, wherein one simulation generates and checks the data set which is run in a second analysis model (a model which analyzes the data generated and regulated by a simulation model). This process highlights the fact that heuristic value and probative value can exist in one and the same experiment in the reproductionist model of simulation methodology. Climate simulation models "have a life of their own," as Hacking would put it, but ultimately their data feeds into analysis models, which are in turn robustly probative.

**Domains:**

In an earlier chapter, I introduced and expanded on a term I borrow from R. A. Snooks: the domain of an experiment. The domain of an experiment consists of the knowledge base which those conducting an experiment have at their disposal for creating the experimental design. Well-tested methodology, previous experiments, confirmed and unconfirmed theories, models, access to phenomena, and other forms of relevant background knowledge are some examples. Recalling chapter three, the domain will determine the robustness of the inference the experiment may draw (but not its total scientific or epistemic value). While experiments (virtual or not) may not always be able to produce robustly probative inferences reliably in newer and specialized sciences, they experimentation is key to the tinkering which produces a defined domain.

Examination of climate modeling modes corroborates the view that experiments define domains. Specifically, the observation of data/model symbiosis and the "semi-empirical" nature of reproductionist experimentation in climate modeling demonstrate how ad hoc yet sophisticated and technically informed feedback loops serve to define the domain of climate simulation model formation and implementation in experiments. Climate simulation models were developed in part out of older methods and models of recording, analyzing and predicting meteorological events and patterns. As meteorological models became more sophisticated and expansively able to account for larger portions of the globe and its climate patterns, researchers began to think about how one might in fact model planet-wide weather and heating and cooling patterns. Thus meteorology taught climatologists to think of the world in terms of four dimensions (three spatial plus time) and in terms of multiple interconnected levels of air patterns and systems.

**Data Shimmer**

Climatology may be a relatively young science, but it has developed and implemented an intricate reproductionist methodology which ensures that it produces an increasingly narrowed spectrum of possible outcomes in response to data shimmering and complexity so vast it rules out both epistemological reductionism and the absolute certainty of an exact prediction.

**The Robust Peer-reviewing within the ICCP Collective Conceivability to Possibility inferences in Simulation Models**

Climatology corrects for the differences in modeling systems through the analysis and evaluation of the Intergovernmental Panel on Climate Change (ICCP). The ICCP is an association of hundreds of climate scientists who work in massive collaborative teams to produce reports on the state of climate science and its collective predictions. . Each section of a report is headed by one or more individual team leaders, but written and edited by hundreds of climatologists. The end product is the most robustly peer-reviewed publication to exist in the scientific world.

In addition to producing one of the most rigorously peer-reviewed reports and syntheses of a discipline, The ICCP investigates and reports on the current methodology of climate science. More so than many other fields, climatology utilizes a homogenized peer reviewing system to analyse models and their predictions, placing them in dialogue with each other.

Put in terms of virtual experimentation, the ICCP produces a vast, collective conceivability to possibility inference about the present state of the global climate and the forecast of what is likely to happen to the climate. Hundreds of researchers collectively pool expert knowledge to predict the future of the world's climate and make suggestions for changing the projected course of disastrous climate change (and regulate and refine that prediction). In this sense, the inferences upon which the ICCP bases its predictions are massive, collective "Conceivability to Possibility" inferences, utilizing the expert knowledge, research findings and experience of hundreds of members.

Climatology gives us a picture of current and future science in which, with increasing access to complex computer simulations, virtual experimentation will become an integral part of reproductionist methodology. Legitimization of virtual experimentation as a genuine and robustly probative scientific procedure will come about as, increasingly, our sciences require previously unfeasible interventions upon vastly complex systems, both large and small, some living and some subatomic and invisible.

**Glimpses Beyond[1]:**

I hope now that I have convinced my reader that thinking about virtual experiments in light of the recognition of the Spectrum nuances and shifts how we understand methodology in general and scientific experimentation in particular. Let's consider where we have been:

Beginning in Chapter 1, we have seen how the all experiments fall somewhere along the spectrum with regards to experimentation directly upon the target phenomena. Furthermore, we have seen how the materials of experimentation make the recognition of relevant features of the target phenomena likely or unlikely to be recognized, but that the use of the target materials do not guarantee said recognition. In Chapter 2, we have examined how modal epistemology can illuminate idealization in experiments, particularly how "conceivability to possibility" inferences structure and inform idealization in experiments. Chapter 3 undertook the exploration of inferences to different kinds of possibility and of how they enable different degrees of inferential power in experiments, which in turn enables experiments to play the wide set of roles that they do. By way of example, Chapter 4 demonstrated how thinking about virtual experiments in climatology reveals big implications for future philosophy of scientific methodology. If we are to understand these advances and remain of use to the sciences, philosophers of science must accept the inferences of virtual experiments as robustly probative, while continuing to explore their success conditions.

For philosophy in general, the spectrum of virtual experiments sheds some light upon the methodology of thought experiments, and also upon their success conditions relative

---

1

to their particular aims and uses. While I see no real need to come up with a normative definition of a thought experiment, it is critical in light of the spectrum of experiments to admit that not everything to which the title thought experiment is applied is of a kind with virtual experiments. Yet, for those thought experiments we philosophers design, utilize, and discuss as virtual experiments, the spectrum of virtual experimentation gives us a taxonomy and methodology for determining whether thought experiments achieve their goals. The family relationship between various virtual experiments demonstrates the nuanced interplay between a priori and empirical investigation, and thus, I hope, suggests a strong relationship between philosophical inquiry and empirical research.

For new sciences utilizing computer simulations and other reproductionist methodology, I hope that the articulation of the spectrum of virtual experiments lends increased clarity in explaining the need for and legitimacy of virtual experimental methods. This clarity can translate into reduced opacity if those methods should, as they have for climatologists, cause scepticism among other scientists, academicians and policy makers.

To this aim, I hope also that the spectrum of virtual experiments may help educators explain the current state of experimental methodology to students in such a way that they can become scientifically literate, critical thinkers regarding the production and regulation of scientific theory and research.

## Bibliography


Baker, John A. (*Forthcoming*), "Some Thoughts on Experiments."

Brumble, K. C. (*Forthcoming*), "Conceivability to Possibility Inferences as Guides to the Relevant Features of Thought Experiments."

Cartwright, N. (1997), "The Blueprints of Laws", *Philosophy of Science*, Vol. 64, Supplement. *Proceedings of the 1996 Biennial Meetings of the Philosophy of Science Association*. Part II: *Symposia Papers* (Dec., 1997), pp. 292-303.

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

Chalmers, A. (2003), "The Theory-Dependence of the Use of Instruments in Science", *Philosophy of Science*, vol. 70 (July 2003) pp. 493–509.

Edwards, N.P. (2010), *A Vast Machine*: *Computer Models, Climate Data, and the Politics of Global Warming*. The MIT Press.

Fine, Kit (1994), "Essence and Modality: The Second Philosophical Perspectives Lecture" *Philosophical Perspectives*, Vol. 8, *Logic and Language*, pp. 1-16.

Franklin, A. (2005), *No Easy Answers*: *Science and the Pursuit of Knowledge*. Pittsburgh, PA: University of Pittsburgh Press.

Frigg, Roman (2009), "Models in Science." Stanford Online Encyclopedia of Philosophy. http://plato.stanford.edu/entries/models-science/. October 6th 2010.

Gendler, Tamar Szabo and John Hawthorne (2005), editors, *Conceivability and Possibility*. Oxford University Press.

Gould, S. J. (1989), *Wonderful Life*: *The Burgess Shale and the Nature of History*. New York: W.W. Norton.

Hacking, I. (1981), "Do We See Through a Microscope?" *Pacific Philosophical Quarterly* 63: 305-322.

—— (1983). *Representing and Intervening*. Cambridge: Cambridge University Press.

—— (1992). "The Self-Vindication of the Laboratory Sciences", in *Science as Practice and Culture*. A. Pickering (Ed.). Chicago, University of Chicago Press: pp. 29-64.

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

Kneale, M. (1937-1938), "Logical and Metaphysical Necessity", *Proceedings of the Aristotelian Society*, New Series, Vol. 38, pp. 253-268.

Kripke, S. A. (1980), *Naming and Necessity*. Cambridge, Mass: Harvard University Press.

Lyons, David (1965), *Forms and Limits of Utilitarianism*. Oxford: Clarendon Press.

Mayo, D. G. (1996), *Error and the Growth of Experimental Knowledge*. Chicago, IL: University of Chicago Press.

Morgan, Mary, and Margaret Morrison (1999), editors, *Models as Mediators*. Cambridge: Cambridge University Press.

Odenbaugh, Jay (2005), "Idealized, Inaccurate, and Successful: A Pragmatic Approach to Evaluating Models in Theoretical Ecology" *Biology and Philosophy* 2005, 20: 231-255.

Quine, W. V. (1982), *Methods of Logic*, Boston: Harvard University Press.

Shoemaker, Sydney (1994), "The First Person Perspective." *Proceedings and Addresses of the American Philosophical Association* Vol. 68, No. 2 (Nov., 1994), pp. 7-22

Snooks, R. A. (2006), "Another Scientific Practice Separating Chemistry from Physics – Thought Experiments." *Foundations of Chemistry*, Volume 8, Number 3, October, pp. 255-270.

Sorenson, R. (*Forthcoming*). "Veridical Idealizations".

— (1992), *Thought Experiments*, Oxford: Oxford University Press.

Swoyer, Chris (1982). "The nature of natural laws", *Australasian Journal of Philosophy*, 60:3, pp. 203–223.

The Harvard Gazette. (2010). *Living, Breathing Human Lung-Chip*. http://news.harvard.edu/gazette/story/2010/06/living-breathing-human-lung-on-a-chip/. Accessed: Oct. 19, 2010.

Tiles, J.E. (1993), "Experiment as Intervention", *The British Journal for the Philosophy of Science*, volume 44, no. 3, September, 1993, pp. 463-475.

Tiles, Mary (1988), "Scientific Dream Space: Symbolic Forms and Scientific Theories", *International Studies in the Philosophy of Science*, 2: 2, 189–204

Vaidya, Anand, (2007), "The Epistemology of Modality", *Stanford Online Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/modality-epistemology/. October 6, 2010.

Winsberg, Eric (2001), "Simulations, Models and Theories: Complex Physical Systems and their Representations", *Philosophy of Science* 68 (Proceedings): 442-454.

— (2007). A tale of two methods", unpublished manuscript can be found under http://www.cas.usf.edu/~ewinsb/papers.html, 2007.

Wittgenstein, Ludwig (1958), *The Blue and Brown Books*. New York: Harper Torchbooks,

Wylie, Alison (1988), "'Simple' Analogy and the Role of Relevance Assumptions: Implications of Archaeological Practice", *International Studies in the Philosophy of Science*, 2: 2, 134–150

Yablo, Stephen. (1993), "Is Conceivability a Guide to Possibility?", *Philosophy and Phenomenological Research*, Vol. 53, No. 1 (Mar., 1993), pp. 1-42.