2021

Modal Understanding of Robustness Analysis

Grayson O'Reilly
University of Montana, Missoula

MODAL UNDERSTANDING OF ROBUSTNESS ANALYSIS

By

GRAYSON RIES O'REILLY

Bachelor of Arts, Christopher Newport University, Newport News, VA, 2017

Thesis

presented in partial fulfillment of the requirements
for the degree of

Master of Arts
in Environmental Philosophy

The University of Montana
Missoula, MT

May 2021

Approved by:

Scott Whittenburg, Dean of The Graduate School
Graduate School

Dr. Soazig Le Bihan, Chair
Department of Philosophy, University of Montana

Dr. Armond Duwell
Department of Philosophy, University of Montana

Dr. Sara Rinfret
Department of PA & Policy, University of Montana

Dr. Wendy Parker
Department of Philosophy, Virginia Tech

O'Reilly, Grayson, M.A., Spring 2021 Environmental Philosophy

Modal Understanding of Robustness Analysis

Chairperson: Dr. Soazig Le Bihan

In this thesis I offer an alternative framework through which we may account for the value of robustness analysis, especially in the context of climate change. In Section 1 I argue that while philosophers of science have reached the relative consensus that RA does not have confirmatory power, it is still used in practice to confirm predictions. Some philosophers push back, stating that RA may, in fact, be of confirmatory virtue or of other epistemic/cognitive value. I argued that ultimately RA fails to provide increased confirmation. In the specific context of climate change and climate modeling, RA is often used to lend increased confirmation to predictions about climate behavior. For example, the AR 5 Synthesis Report found that climate models largely agree that increased GHG emissions lead to increased global surface temperatures. The problem, then, is that "such agreed-on or robust findings are sometimes highlighted in articles and reports on climate change" providing a level of certainty to the claim (Parker 2011, 580). So if not used to confirm, then what is RA useful for? Instead of appealing to RA's confirmatory power, I offer an alternative framework to account for the usefulness and value of RA: Le Bihan (2017) and Duwell's (2018) modal understanding framework. This alternative framework does not answer the hard problem set forth by Orzack and Sober (1993): it remains unclear how we bridge the gap between models and the real world. But through the modal understanding framework, RA gives us a way to learn about the relationships between models and use their predictions in a better way, especially in the context of climate change.

Modal Understanding of Robustness Analysis

While models are necessary to represent complex systems, they incontrovertibly misrepresent the world.[1] Because of the "falsities" models contain about their target phenomena, or that which they are modeling, it is unclear how to assess how true or confirmed their predictions are.[2] Robustness analysis (RA) is often offered as a remedy to this issue: when individual, idealized models agree with one another on a given prediction, they are thought to reveal something true about the world. However, this claim remains controversial. Over the past fifty years, philosophers of science have debated the value of RA and largely undermined the notion that RA has confirmatory power for robust predictions. And yet, it is used to do so in practice.

For example, in the context of climate change, RA is routinely used to bolster predictive claims agreed upon by models. Given that "there is now a broad scientific consensus—underwritten by a substantial and growing body of evidence—that the earth's climate warmed significantly over the last century," the rate at which this change will occur remains unclear (Parker 2011, 579). Policymakers rely on climate models and RA to answer that question, ignoring the many philosophical analyses that undermine the confirmatory power of RA for robust predictions. The challenge, as Parker (2011) explains, is that notwithstanding the problems of RA, "[o]f course, it does not follow that climate policy decisions should be put on hold" but "rather to make sensible decisions despite remaining uncertainties about the details of future climate change" (598).

---

[1] For the purposes of this paper, I use Jay Odenbaugh's (2019) definition of models as "representations that abstract and idealize" (15).
[2] While "accuracy" may be more appropriate, I use the vocabulary of "truth" and "falsehood" throughout this paper to align myself more closely with the language of RA literature.

So if not adequate to confirm predictions, especially in the context of climate change, what is RA useful for? Some have offered alternative frameworks to account for the epistemic value of RA. In this paper I offer one such framework, modal understanding, that may address the issue described above. Through the modal understanding framework, RA may be better epistemically understood and its usefulness in climate policymaking be better warranted.

In Section 1, I explain the relative consensus that RA does not have confirmatory power, but is still used in practice to confirm predictions. I investigate arguments that push back against this consensus and attempt to recover confirmatory power and other epistemic or cognitive value. I demonstrate the ways in which RA fails to provide increased confirmation and explain why recent analyses of RA are unsatisfactory. Finally, I look at RA in the specific context of climate science as RA is often used in climate modeling to lend increased confirmation to predictions.

In Section 2, I give an account of Le Bihan and Duwell's conception of modal understanding and argue that under the modal understanding framework we are able to account for the epistemic value of RA, making sense of the usefulness of RA without appealing to its confirmatory power.

In conclusion, I hope to convince the reader that the kind of understanding of phenomena that RA provides allows us to engage in some forms of reasoning that may be useful for policy decision making.

1. Robustness and Confirmation

Conversation regarding RA in philosophical literature is often traced back to biologist Richard Levins' influential 1966 article, "The Strategy of Model Building in Population Biology." Levins' main concern was how multiple models relate to a hypothesis. He claimed that we can access truths about nature when we identify what he calls "robust theorems."[3] He argues that when a set of models points to a common prediction, then the prediction is robust and likely says something true about the world. This is accomplished by looking closely at models' results and "if these models, despite their different assumptions, lead to similar results, we have what we can call a robust theorem that is relatively free of the details of the model. Hence, our truth is the intersection of independent lies" (1966, 423). So model agreement is supposed to be indicative of truth about the world (Levins 1993). Robustness, in this way, is taken to have some level of confirmatory power for the robust prediction.

Biologist Steven Orzack and philosopher Elliot Sober (1993) famously published a direct response to Levins wherein they argue that whether or not a prediction is "robust" in Levins' sense does not have any clear bearing on the question of whether that prediction is true. Eric Winsberg (2018) summarizes Orzack and Sober's response in clear terms:

> "Orzack and Sober argued that anyone who wants to employ [Levins'] principle faces the following dilemma: unless the employer of RA knows that at least one member of [the model set] is the "true model", then the fact that all of the members of [the model set] detect [the robust theorem] is no reason at all to believe [the robust theorem]. But if she already knows this, then it is unclear what role the rest of the set is playing." (4)

In other words, when models produce similar results they are not pointing at something true in the world unless we already know that one of the models is an accurate representation of the world. Instead, agreement among models demonstrates that these models share some mathematical properties (Orzack and Sober 1993).

[3] I take a robust theorem to be the conditional proposition in which the robust prediction is the consequent. As Weisberg (2006) describes it, a robust theorem is a hypothesis that "simply describes the regular connection between the effect of a causal structure on a highly general class of systems" (13). For example: if structure S is in place, then prediction P. RA literature also references robust outcomes, results, and properties, all of which I refer to as robust predictions.

Orzack and Sober asked that Levins write a response in the same issue of *The Quarterly Review of Biology*. In his response, Levins (1993) disputed all claims against his original work stating that the "formal/analytic framework" used by Orzack and Sober "is not the appropriate domain for evaluating a model of research strategy" (547). Despite this, Orzack and Sober's criticism has remained unresolved since. As we shall see in this section, the various attempts at articulating the epistemic value of RA in recent literature do no not fully address the fundamental challenge that Orazack and Sober posed, i.e. that RA reveals something about the mathematical properties of models and theories rather than something about the world.

1.1 Consensus about the Confirmatory Power of RA

The Orzack and Sober challenge is that RA lacks confirmatory power. Yet it is used as lending confirmation to robust predictions in scientific practice (this is especially the case in climate science). Confirmation occurs when evidence supports a hypothesis. It is common scientific practice to take the fact that some evidence supports some hypothesis as pointing to the hypothesis being more likely to be true. So taking RA to have confirmatory power means that agreement among models on a prediction counts as supporting evidence that points to the prediction as more likely to be true. The Orzack and Sober challenge, however, highlights that this may not be legitimate. Careful examination of recent literature concerning RA reinforces this worry as it reveals a relative consensus that RA does not have confirmatory power (Orzack and Sober 1993, Woodward 2006, Houkes and Vaesen 2011, Odenbaugh 2011, Odenbaugh and Alexandrova 2011, Parker 2011, Justus 2012).

For example, Odenbaugh (2011) argues, contra Levins, that "our truth is nowhere in the lies" because we are unable to de-idealize models (1187). Models, especially global climate models, are idealized insofar as they make assumptions about the target systems that are false but necessary to grasp the complexities of these systems. RA is unable to remove these assumptions, Odenbaugh argues, so we are not justified in trusting robust predictions. In the context of climate change and climate models, Parker (2011) reiterates Orzack and Sober's argument by explaining that because we are unable to confirm the truth of any one model in a set of climate models, there is no reason to have increased confidence in robust predictions.

Recent analyses of RA have tried to remedy this problem by either seeking ways around the challenge and defending the view that RA has confirmatory power or looking at other ways in which RA may be epistemically valuable. I argue none do so in a way that fully addresses the challenge.

1.2 Attempts to Recover Confirmatory Power or Epistemic Value

Some philosophers hold firm that RA may have some confirmatory power. Let us review some of these accounts. William Wimsatt (1981), like Levins, champions robustness stating that robust predictions are confirmed to the degree that the process of RA is able to filter out the "illusory parts" of the models. "Robustness," Wimsatt says, "is a criterion of the reality of entities" (75). In one of the earliest attempts to bring RA to the foreground, he identifies "illusions" of RA that he argues, quoting social scientist Donald Campbell, "occur when confirmation is attempted and found lacking" and then points to the failure of independence of evidence as responsible for producing these illusions (75). So RA fails when models are not independent enough from one another (often via the parameters of the considered models).[4] But when models are sufficiently independent, RA helps to identify robust predictions that are likely to be true because when a prediction is robust the abstractions and idealizations of the individual

[4] Like many other philosophers and scientists (Stegenga and Menon 2017, Steele and Werndl 2013, Winsberg 2018, Pirtle, Meyer, and Hamilton 2010, Odenbaugh 2011), Wimsatt calls for further attention to be given to how independence of evidence relates to the confirmatory power of RA.

models used to arrive at the robust prediction likely do not matter. As Weisberg (2006) explains: "From Wimsatt, we learn that robustness analysis' aim is to separate the scientifically important parts and predictions of our models from the illusory ones which are accidents of representations" (3).

Some philosophers do not believe Wimsatt's view is properly justified. Odenbaugh and Alexandrova (2011), using the work of Nancy Cartwright, argue that RA fails to de-idealize models because it does not overcome the problem of independence of evidence. RA, they say, hinges on the ability to demonstrate the independence of models' assumptions. But models are not adequately independent because they are "different ways of doing the same thing," in other words, models of the same phenomena are too closely related. So while their predictions may agree, their false assumptions have not been eliminated (Odenbaugh and Alexandrova 2011, 760).

Jonah Schupbach (2018), like Wimsatt, is concerned with independence of evidence and suggests that our means for satisfying diversity of evidence in RA are not satisfactory. For this reason, under prevailing frameworks for understanding RA we cannot believe that robust predictions are more likely to be true. Schupbach instead offers an alternative account wherein RA may be a means of eliminative confirmation for explanations. In Schupbach's view, means of detection are sufficiently diverse if the explanation they provide rules out another explanation. In this way, "[the target hypothesis] is incrementally confirmed by ruling out possible ways that it could be false" (Schupbach 2018, 297). The thought is that one observes target phenomena through multiple, sufficiently diverse means of detection. The diversity of the representations, which each provide different explanations for the target phenomenon, then allows us to weed out explanations until we arrive at the best possible one.

It is important to note that Schupbach's explanatory confirmation shifts the target of confirmation from predictions to explanations. Traditionally, RA is not used to explain phenomena but to attempt to confirm predictions. This is precisely the reason RA is pertinent to climate models; we are less interested in *why* the climate is warming and more interested in *the degree* to which it will warm. So Schupbach's explanatory account of RA only applies to a subset of RAs that target explanations of hypotheses and does not fit when applied to the types of RA discussed in this paper, namely the use of RA as a tool for confirming predictions.

Michael Weisberg (2006), on the other hand, argues what RA confirms is the relationship between the causal structure and the prediction. He claims that Orzack and Sober have reduced RA to nonempirical confirmation while missing two crucial steps in the process. First, Weisberg notes that the structure of a robust theorem is that of a conditional proposition linking a common causal structure and a common property among the agreeing models. Weisberg (2006) uses an example in population ecology to highlight this: "*Ceteris paribus*, if the abundance of predators is controlled mostly by the growth rate of the prey and the abundance of the prey controlled mostly by the death rate of predators, then a general pesticide will increase the abundance of the prey and decrease the abundance of predators" (737). Second, the theorist must demonstrate that the evidence for this proposition is adequately independent. If a set of models is sufficiently heterogeneous, "then it is very likely that the real-world phenomenon has a corresponding causal structure" (Weisberg 2006, 739). Then, RA itself does not have confirmatory power but rather robust predictions "are confirmed via low-level confirmation, the sort of confirmation that licenses the use of a framework to construct models of phenomena in the first place" (742). Returning once more to the population ecology example: "we are confident that ecological relationships can be represented with the models described by coupled differential equations.

Thus when we discover the consequences of these models, we are confident that most of these consequences are true of any system described by the model" (741).

Weisberg's view faces two important issues. First, RA only seems able to confirm formal relationships between model properties. What is confirmed is a conditional proposition of the form: 'if [causal structure within model] then [property of model]'. Second, any increased level of confirmation for the prediction itself does not come from RA (the commonality of prediction among various models) but from the low-level, external, and independent confirmation of each model. Were the models not externally confirmed, RA does not lend extra confirmation to the prediction. So Orzack and Sober's challenge still remains unanswered: if we do not know those representations to accurately represent the real world, then RA is still unable to offer increased confirmation of robust predictions.

1.3 RA in Climate Science

Orzack and Sober's challenge is even more prominent in the context of climate science and climate modeling not only because the earth processes being modelled are incredibly complex, necessitating abstractions and idealizations, but also because climate models are the best tools we have to inform us about the future effects of climate change. While RA is often used to bolster predictive claims, we have seen that RA may not be capable of doing so. So, if RA cannot confirm predictions produced by climate models, what is it able to do? Wendy Parker and Elisabeth Lloyd offer two different views on RA in this context and demonstrate that it is not readily apparent how capable RA's confirmatory power is in the practice of climate science.

Parker (2011) argues that in the context of climate modeling the uncertainty of individual models cannot be overlooked so RA is not necessarily fit to tell scientists that predictions are more likely to be true. She looks closely at the adequacy conditions that sets of models must meet for model agreement to have special epistemic significance (Parker 2011). When looking at how we construct sets of models we still have no way of knowing if one of the models in our model set is true, in other words, we have failed to overcome Orzack and Sober's original critique. When looking at how we evaluate the performance of model sets, we do not do our due diligence to acknowledge repeated instances of idealization and abstraction. If our means of evaluation hold the same idealizations and abstractions as the models themselves that may artificially inflate the "frequency with which ensembles are found to capture truth" (Parker 2011, 587). Ultimately, she identifies some conditions under which robust predictions may have special epistemic significance, but concludes that they are not met in present climate modeling. We are left with "goals for the construction and evaluation of ensembles" that, if met, will allow robust predictions to "have desired epistemic significance" (Parker 2011, 598). While we may be able to identify attributes that model sets must have for RA to help with confirmation, current iterations of RA in climate science are incapable of providing increased confirmation/confidence/security to a hypothesis.

Elisabeth Lloyd (2010), on the other hand, argues that through Weisberg's framework RA may have some level of confirmatory power. First, we identify a set of models covering an adequately "wide range of assumptions and conditions" (Lloyd 2010, 981). Then, as we saw with Weisberg's account, what is confirmed is a conditional proposition, 'if [causal structure within model] then [property of model]'. Then, according to Lloyd and Weisberg, if the models within the model set share a common structure we may assume that the structure is present in the real world. Lloyd (2010) uses the example of greenhouse gases causing global warming: if a set of sufficiently diverse models have the "common structure of greenhouse gas causation . . . it is very likely that the real world phenomenon has a corresponding causal structure. Therefore, we could infer that greenhouse gas concentration increases cause global warming in the real world, as the attribution studies have also shown" (981-982). Lloyd's distinctive move is to argue that RA may offer an increased level of confirmation not because of the low-level confirmation that Weisberg suggests, but because the core structure (the conditional) is backed by a variety of evidence.

Lloyd notably shifts the confirmatory power of RA from the robust prediction produced by a set of models to the causal structure. While Lloyd's account may provide us with increased reason to believe that causal structures found in model sets will be present in the real world, similarly to Schupbach's account, Lloyd's does not seem applicable to types of RAs commonly used in practice, that is RAs that target robust predictions. One may respond that if the causal structure is confirmed in the real world, so too is the prediction. But if that is the case, RA is not the source of confirmation, rather confirmation is a virtue of "the physics [being] sound and well-confirmed" (Lloyd 2013, 62). In other words, some previous confirmation of the models is what does the confirmatory work, not RA. So if not used to confirm predictions, what does RA do that is epistemically valuable when evaluating climate models?

2. A Modal Understanding of Robustness Analysis

Let us now examine RA under the modal understanding framework, proposed by Soazig Le Bihan (2017) and Armond Duwell (2018), and see how we may account for the epistemic value of RA without appealing to its confirmatory power. Le Bihan (2017) characterizes modal understanding as follows: "One has some modal understanding of some phenomena if and only if one knows how to navigate some of the possibility space associated with the phenomena" (112).

In other words, when we gain knowledge about how phenomena might come to be in our world, we have modal understanding.

We may come by modal understanding in a variety of ways which are rather comprehensively described by Duwell (2018). He outlines three kinds of understanding associated with three kinds of possibility spaces: modal understanding of phenomena, internal modal understanding of theory, and extensional modal understanding of theory. This paper focuses primarily on modal understanding of phenomena.[5]

In this section, I show how RA is one way to arrive at modal understanding of phenomena. First I show how to identify the possibility space. This is the space consisting of representations (models) and fundamental for the modal understanding framework. Next, I explain what it means to navigate the possibility space as navigating power provides modal understanding. Then, I show how RA affords us navigating power, providing us with modal understanding of phenomena. Finally, I explain how because RA affords navigating power it is epistemically valuable. I provide examples from the IPCC AR 5 Synthesis Report, the most recent comprehensive report from the United Nations' collaborative body charged with assessing the science of climate change, to illustrate how the value of RA as used in practice can be accounted for in terms of modal understanding.

2.1 Identifying the Possibility Space

The possibility space for understanding phenomena includes the set of representations of the target phenomena and the relations between those representations (Duwell 2018). It is comprehensive insofar as it contains all possible representations of the target phenomena and subsets of the target phenomena. Phenomena are typically represented by dependency structures

---

[5] I suspect that RA may provide other kinds of modal understanding, but that is outside the scope of this paper.

between variables and parameters. Though not always the case, we can think of dependency structures as causal relationships that give rise to the target phenomena within a representation[6]. Let us note that the possibility space is also contained by "adequacy conditions" that models must meet. This is because models are created by scientists with pragmatically defined and context-dependent goals. So what counts as an "adequate" representation will depend not only on the target phenomena but on the goals of the user.[7]

Let us take a simple example to illustrate how one identifies the possibility space before we do the same with the more complex example of climate models. Through this example, I will also examine why multiple models are necessary to represent target phenomena in the first place and I will clarify the notion of "dependency structures." In the modal view, representations show us a variety of ways in which the target phenomena could come to be. Let us consider models as an example of representations. Because models are necessarily selective, insofar as they are abstracted and idealized and therefore cannot wholly capture the conditions of the real world, it is common to work with various models of the same phenomena. The possibility space contains all possible models associated with the various possible selections.

For example, if I wanted to model the phenomena of dropping my pen and the pen then falling to the ground, one model may be a theory of gravity—e.g Classical Mechanics (CM). And while this representation would take into account laws such as Newton's law of universal gravitation to represent why and how the pen fell to the ground when I dropped it, the model would say nothing of the color of the pen and most likely would ignore the resistance of the air

---

[6] Le Bihan (2017) notes that dependency structures may also be mereological or logical relationships (114-115).

[7] See Duwell (2018), who notes that "it is certainly not the case that scientists are interested in *all* kinds of representations" (2) when determining the set of representations that make up the possibility space for the target phenomena, representations may have to meet adequacy conditions that coincide with the interests of those conducting the investigation.

as well as the particular shape of the pen. In doing so, the model does not wholly capture the conditions of the real world; it has been necessarily abstracted and idealized. If it were not necessary to do so, there would be no need for multiple models of the same phenomena. Instead, when we look at a set of models that each target the same phenomena, they show how the phenomena could come to be under different circumstances.

The CM model offers one way the phenomena could come to be, but Aristotelian physics (AP) does so as well. It depicts four elements that always return to their natural place. General Relativity offers yet another alternative representation of the phenomena. These representations are all a part of the possibility space associated with the phenomena of the pen falling to the ground. The possibility space then includes dependency structures that give rise to the phenomena in the different models and the relationships between them. Remember, dependency structures are often causal relationships so they include the important factors that bring about the phenomena as well as the functional relationships between such factors. In the CM model, gravity is a physical force where the force between two objects is expressed as dependent on the product of their masses divided by the square of the distance between them. So the mass and the positions of masses (the pen and the Earth's center) combined with Newton's law of universal gravity explain how and why the pen fell to the ground when dropped. In this case, movement is a result of a gravitational force acting on the pen. On the other hand, in the AP model, instead of mass and position, what is important are the elemental makeups of the pen and Earth (both made mostly of earth, the heaviest element) and why elements move to their natural places (light ones go up while heavy ones go down). Here movement is a function of the earth in the pen returning to its natural place at the center of the Earth. We can then examine the relationships between these dependency structures. Do the models share certain properties? Do some preclude others?

What factors are accounted for and not accounted for? Once we have identified some of the possible representations of the target phenomena and some ways in which representations relate to one another we have enough information to navigate the possibility space. Note that it is impractical to think we can identify all of the possible representations of the target phenomena. Modal understanding does not require we have the ability to navigate all of the possibility space, just some. Modal understanding comes in degrees.

We may now consider the more complex example of climate models. Again, the possibility space includes models of the target phenomena and their relations. First we must identify the target phenomena. The AR 5 Synthesis Report assessed 70 climate models in total including a variety of model types. These climate models aim, broadly, to provide a picture of global climate behavior so we may identify global climate behavior as the target phenomena. We may treat each model as a representation of the target phenomena in the possibility space. To identify relations between models, we must identify dependency structures within the models. Climate models contain a vast number of dependency structures as they represent incredibly complex systems. Note that fundamental laws of physics serve as the foundation of climate models. For example: the atmosphere is a thin layer of air across the surface of the Earth held in place by gravity. So atmospheric models utilize equations of gravity, combined with equations of fluid dynamics and thermodynamics to predict future states of the atmosphere, or in other words, to predict the climate. Often, models are then "tuned" where modelers "adjust parameter values (possibly chosen from some prior distribution) in order to optimise model simulation of particular variables or to improve global heat balance" (Randall, et al. 2007, 596). Let us look at one example of a dependency structure: the causal relationship between greenhouse gas (GHG) emissions and global surface temperature.

The AR 5 Synthesis Report includes important findings from the Coupled Model Intercomparison Project Phase 5 (CMIP5) which is a protocol developed by the World Climate Research Program Working Group on Coupled Modelling that aims to "provide climate scientists with a database of coupled [general circulation models] simulations under standardized boundary conditions" to facilitate "the study of intrinsic model differences at the price of idealizing the forcing scenario" (Covey, et al. 2003). In other words, CMIP5 allows climate scientists to more easily investigate the differences between models, model predictions, and model retrodictions. For example, models are run under different Representative Concentration Pathways (RCPs). These pathways represent scenarios of different levels of greenhouse gas emissions such as RCP2.6 which is a stringent mitigation scenario in which we immediately and drastically slow GHG emissions or RCP8.5 which accounts for a scenario with very high GHG emissions. Then, based on analyses of the various scenarios, "multiple lines of evidence indicate a strong, consistent, almost linear relationship between cumulative CO2 emissions and projected global temperature change" (IPCC 2014, 8). The causal relationship between GHG emissions and global temperature change is a dependency structure.

Importantly, the possibility space includes relations between dependency structures. Let us examine some findings from CMIP5 models reported in the AR 5 Synthesis Report.

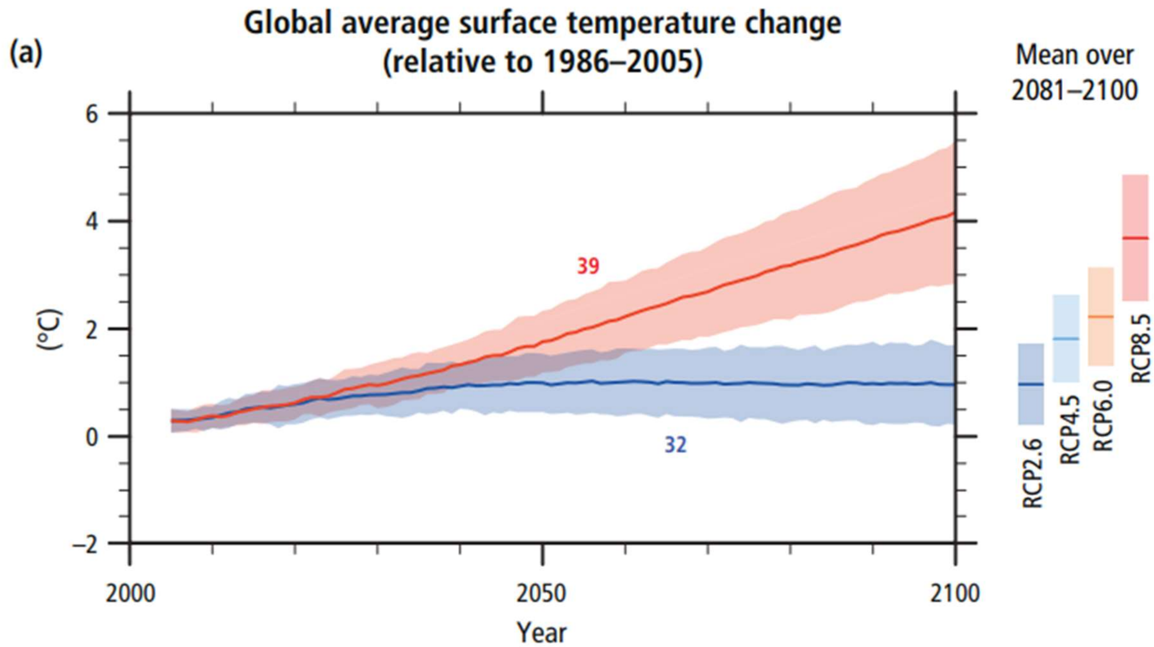

Figure 1. Figure SPM.6 from AR5 Synthesis Report Summary for Policy Makers. Global average surface temperature change based on RCPs 2.6 (stringent), 4.5 (intermediate), 6.0 (intermediate), and 8.5 (very high). (IPCC 2014, 11)

In Figure 1 we see that 39 models were run under RCP8.5 (high GHG emissions scenario) and the average global surface temperature change produced by these findings is depicted by the red line while the measure of uncertainty is depicted by the red shading. Similarly, 32 models were run under RCP2.6 (low GHG emissions scenario) and the average global surface temperature change produced by these findings is depicted by the blue line while the measure of uncertainty is depicted by the blue shading. We are looking specifically for how representations relate to one another. So when 39 CMIP5 models are run under RCP8.5, we learn how these models relate to one another: a high GHG emission scenario, on average, leads to increased global surface temperature. Again, when 32 CMIP5 models are run under RCP2.6, we learn how these models relate to one another as well: a low GHG emission scenario, on average, leads to stagnated or slightly lowered global surface temperature. We have identified several ways in which these representations relate to one another: when models are run under similar RCPs they produce similar results, high GHG scenarios tend to produce increased global surface temperature, and low GHG scenarios tend to produce decreased global surface temperature.

When we can identify some of the possible representations, such as the CMIP5 models assessed in the AR 5 Synthesis Report, as well as dependency structures within the models and how they relate to one another, such as how GHG emission scenarios affect global surface temperature, we have adequately identified some of the possibility space associated with the target phenomena of global climate behavior. Modal understanding of phenomena then hinges on the ability to navigate the possibility space, which we may do in a variety of ways.

2.2 Navigating the Possibility Space

When one knows how to navigate some of the possibility space, one has modal understanding. Knowing *how* is here understood as having certain abilities.[8] Importantly for our purposes, one knows how to navigate the possibility space for some phenomena if: 1) one knows how the models represent the target phenomena in some way and/or 2) one knows how the models relate to one another and/or 3) one knows how constraints apply to the possibility space.

To illustrate modal understanding, let us return to our simple example of the falling pen. We have a set of models (Classical Mechanics, Aristotelian physics, etc.), each targeting the phenomena of the pen falling to the ground. These models and some of their relations constitute

[8] Duwell (2018) stresses the "how" of know-how: "Notice that knowing how as opposed to knowing that plays an important role in navigating a possibility space. . . A canonical example is that knowing how to ride a bike is not reducible to knowing facts about bicycle riding. It would be odd to say that one knows how to ride a bike by stating facts about bicycle riding but not be able to ride a bike. Hence, knowing how to do something seems essentially connected to being able to do something or at least having a disposition to be able to do something in proper conditions." (3)

the possibility space. We saw earlier that in the CM model movement is a function of gravitational force acting on the pen and in the AP model movement is a function of the earth in the pen returning to its natural place at the center of the Earth. Le Bihan (2017) argues that even when models misrepresent the world, as both the CM and AP models do, they provide modal understanding because they tell us how “a particular dependency structure gives rise to [the target phenomena], or to some subset of [the target phenomena]” (117). In other words, when we look at how and why the pen falls to the ground in the models, we learn how and why the pen *could* fall to the ground in the real world. When we can “navigate the inner workings of a dependency structure” such as how movement occurs, we know how the model represents the target phenomena in some way and have modal understanding (Le Bihan 2017, 119).

Then, “one reaches a *second level* of modal understanding if one *also knows how different dependency structures for* [the target phenomena] *relate to one another”* (Le Bihan 2017, 119). So modal understanding suggests that we gain some understanding of how the phenomena of dropping the pen and it hitting the ground may come to be by examining the relationship between the CM and AP models. For instance, one might know in the CM model how to account for the specific mass of the pen whereas the AP model does not offer a satisfactory way to handle the specific mass of the pen as different from a much larger or smaller mass. When looking at the models independently, we do not know what we do not know, so understanding is gained by examining how the dependency structures relate to one another.

2.3 How RA Provides Navigating Power

As it turns out, RA examines how models relate to one another thus providing us with modal understanding. Below I explain how RA may tell us 1 (how the models represent the target phenomena in some way), often tells us 2 (how the models relate to one another), and sometimes tells us 3 (how constraints apply to the possibility space).

1) As we saw in Section 1, RA does seem capable of identifying a possible representation of the world. While RA may not have the power to confirm that a causal structure identified in a set of models will be present in the real world, when diverse models produce similar results, RA may reveal a common structure that *can* give rise to the target phenomena (Lloyd 2010, Weisberg 2006). By illuminating this common structure, RA may teach us how models represent the target phenomena. In other words, because a set of models produces similar predictions, we may then investigate possible causes for those predictions which may then, in turn, reveal a common causal structure throughout the models. That common causal structure, whether we are able to confirm its existence in the real world or not, is one way that the target phenomena (or some subset of the phenomena) could come to be in the real world. So RA may tell us how models represent the target phenomena in some way.

   For example, if 21 CMIP5 climate models predict that, on average, under RCP8.5 the equatorial Pacific will experience increased annual precipitation, RA tells us to pay special attention to these models because their predictions agree. It may not be the case that because 21 models generally agree that higher GHG emission scenarios will lead to increased precipitation in the equatorial Pacific that the same causal structure exists in the real world. But it is apparent that RA, in this case, has illuminated one way in which models may represent the target phenomena.

2) RA reveals how multiple representations of the target phenomena possess a common property. Instead of being a detriment, which we largely saw it to be in Section 1, that a

given set of possible representations have a common property, it is instead a *feature* of the possibility space. One gains modal understanding whenever one learns about the features of the possibility space.

For example, what is interesting about the CMIP5 climate models assessed in the AR 5 Synthesis Report, for our purposes, is how under certain GHG emission pathways the models agree that by 2081-2100 global surface temperature will have increased by 2 degrees celsius compared 1850-1900. So these models largely agree that if we continue to emit GHGs at high rates, the Earth will warm. We saw in Section 1 that it may be a mistake to think that because models produce a robust prediction we have increased reason to believe that the prediction will come to fruition in the real world. Instead, under the modal understanding framework robust predictions tell us how the models relate to one another, namely that they agree. But as we know, RA may reveal common causal structures present in the models as well. So when 39 CMIP5 models agree that RCP8.5 will lead to a global surface temperature increase of 2 degrees celsius by 2100, we do not have increased reason to think that will be the case in the real world but instead we learn about how global surface temperature increase may come to be: through increased GHG emissions. RA in this way provides modal understanding of the target phenomena without relying on a confirmatory virtue.

3) One may imagine an extreme case in which RA reveals that all models representing the target phenomena or theory share the same property. In other words, all representations of the possibility space have the property in common. In this case, one learns about the constraints on the possibility space: models that do not possess the property would

necessarily be outside of the possibility space. This is, of course, quite unlikely but not an impossible case.

2.2 Accounting for the Value of RA

Now that we have laid out how to identify and navigate the possibility space associated with modal understanding of phenomena, how does modal understanding account for the epistemic value of RA? Le Bihan (2017) states that "the epistemic value of scientific theories and models is generally taken as coming in three kinds: 1 predictive power, 2 explanatory power, 3 heuristic power" (111). Some models, even false models, may, in fact, have predictive power. Newton's theory, for instance, may still be used to send rockets to space but does not wholly capture the conditions of the world. But, as Le Bihan (2017) argues: "Most philosophers, however, agree that this is not always the whole story, for two main reasons. The first is that not all [models and theories that misrepresent the world] are good at making predictions. The second is that most philosophers reject a purely instrumentalist view of science, and hope that [models and theories that misrepresent the world] can be conceived as affording some epistemic value beyond their mere predictive power" (111). Additionally, we have seen that it is unlikely that robust predictions have more predictive power than non-robust ones. What about explanatory power? As models are known to misrepresent the world, and are therefore unlikely to be true, they cannot be taken to have explanatory power because they fail to meet the conditions of an adequate explanation, at least on the traditional view of explanation, which requires that the explanans be true for an explanation to be adequate. We saw in Section 1 that it is unlikely RA is capable of "de-idealizing" models so we cannot be confident that robust findings have increased explanatory power (Odenbaugh 2011). Finally, models may have heuristic power insofar as they can contribute to the discovery of better models but, as Le Bihan

(2017) explains, this too may be unsatisfying as heuristic power, like predictive power, is purely instrumental and says nothing about "what it is about [models and theories that misrepresent the world] that may or may not promote scientific progress" (111).

The modal view explains how models and RA may be epistemically valuable: "Insofar as knowing-how has genuine, intrinsic, epistemic value, [models and theories that misrepresent the world] that have navigating power have genuine, intrinsic, epistemic value" (Le Bihan 2017, 122). As we have seen, RA may be capable of providing navigating power, thus facilitating understanding of the target phenomena. So RA, too, may be said to have genuine, intrinsic, epistemic value. While models never fully capture that which they represent, they nonetheless provide understanding of that which they represent. The modal view provides depth to this analysis by explaining how models provide understanding: they provide navigating power. RA provides us the know-how to navigate the possibility space and thus some understanding of the target phenomena and how the target phenomena may come to be in the real world.

For example, when one knows how models relate to one another, "the ways in which the models can be changed to recover a larger portion of the phenomena become clearer" (Le Bihan 2017, 122). So when examining our set of 39 climate models run under GHG emission scenario RCP8.5, we may identify the presence or absence of other dependency structures, such as the causal relationship between GHG emissions and ocean acidification, to determine how that dependency structure relates to global surface temperature. Does the ocean pH range increase with global surface temperature? Decrease? Must the ocean pH be at a certain level to reach global surface temperature predictions? If so, what are they? RA may allow us to answer these questions and, in turn, allows us to understand how we may recover a larger portion of the target phenomena.

One may also understand how to better discern different types of ways that models may relate to one another and how to apply these relations to other contexts. For instance, if a climate model utilizes one method to represent Arctic sea ice thickness and produces a similar global surface temperature prediction to another model that does not account for Arctic sea ice thickness, scientists may then be led to investigate the nature of the relationship between Arctic sea ice thickness and global surface temperature.

In these ways, and others, the modal understanding framework accounts for the value of RA without appealing to its confirmatory power. Modal understanding is a form of knowledge and, as such, has genuine, intrinsic, epistemic value. RA provides navigating power, affording us the know-how to navigate some of the possibility space, thus providing us with modal understanding. So RA is epistemically valuable not because of its ability to confirm but because it provides us with modal understanding.

3. Conclusion

We saw in Section 1 that while philosophers of science have reached the relative consensus that RA does not have confirmatory power, it is still used in practice to confirm predictions. Some philosophers push back, arguing that RA may, in fact, be of confirmatory virtue or of other epistemic/cognitive value. I argued that ultimately RA fails to provide increased confirmation. In the specific context of climate change and climate modeling, RA is often used to lend increased confirmation to predictions about climate behavior. For example, the AR 5 Synthesis Report found that climate models largely agree that increased GHG emissions lead to increased global surface temperatures. The problem, then, is that “such agreed-on or robust findings are sometimes highlighted in articles and reports on climate change" providing a level of certainty to the claim (Parker 2011, 580). So if not used to confirm, then what is RA useful for?

Instead of appealing to RA's confirmatory power, I offer an alternative framework to account for the usefulness and value of RA: Le Bihan (2017) and Duwell's (2018) modal understanding framework. This alternative framework does not answer Orzack and Sober's original challenge: it remains unclear how we bridge the gap between models and the real world. But through the modal understanding framework, RA gives us a way to learn about the relationships between models and use their predictions in a better way.

One has modal understanding when one knows how to navigate some of the possibility space associated with the target phenomena. After establishing how to identify and then navigate the possibility space, we saw how RA may provide modal understanding in three distinct ways by providing knowledge of how the models represent the target phenomena in some way, how the models relate to one another, and how constraints apply to the possibility space. Finally, I demonstrated how RA is intrinsically and epistemically valuable without appealing to its confirmatory power because it teaches us how the target phenomena may come to be.

While it would take significantly more space and time than I have here to fully explore, it is my hope that because RA provides modal understanding of phenomena we will be able to engage in some forms of reasoning that may be useful for policy decision making. We tend to think that the most useful information we get from science tells us what kinds of possible interventions could help us manipulate the world to achieve desired outcomes. Policy is just this kind of intervention. Explanatory understanding, i.e. understanding of the true causal structures underlying the phenomena, is typically taken to give us the means for such interventions. On the basis of our understanding of the world as knowledge of the underlying causal structure, we would be able to predict which consequences various possible interventions on that causal structure would have and make policy decisions accordingly. However, because of the extreme complexity of many of the systems involved in policymaking, we typically cannot trust that we understand the world and its causal structure as it actually is. This is especially true of climate change—we do not have a magical model revealing the true causal mechanisms in the world that would allow us to design the kinds of interventions warranted to combat rising temperatures and sea levels at a global scale. Instead, we possess various highly idealized models the predictions of which are difficult to ascertain. Even when multiple models agree, we have seen that it is not clear that this provides enough confirmatory power to warrant solidly grounded policy making. With the real causal mechanisms of climate change unknown to us, we must look to the possible causal mechanisms.

In situations where understanding the *world* as possessing a clear picture of its true causal structure is out of reach, perhaps there is a second best option: understanding of *phenomena* and its possible causal structures. Through the modal understanding framework RA is valuable precisely because it allows us to address "*what—if—things—had—been—different"* questions. When one knows how to navigate some of the possibility space associated with the target phenomena, one understands how the target phenomena may come to be and may thus be better situated to design appropriate interventions. Again, I do not expect the modal understanding framework to solve the Orzack and Sober problem. But were there enough empirical evidence given to us, modal understanding may give us the kind of modal reasoning useful for decision making, especially in the case of combating climate change.

Works Cited

Covey, C. et al. (2003). "An Overview of Results from the Coupled Model Intercomparison Project (CMIP)." *Global and Planetary Change,* 37, 103-133.

Duwell, A. (2018). "Understanding quantum phenomena and quantum theories, Studies in History and Philosophy of Modern Physics." *Studies in History and Philosophy of Modern Physics*, 1-14.

Houkes, W., & Vaesen, K. (2012). "Robust! Handle with Care." *Philosophy of Science*, *79*(3), 345–364.

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland.

Justus, J. (2012). "The Elusive Basis of Inferential Robustness." *Philosophy of Science,* 79(5), 795-807.

Le Bihan, S. (2017). "Enlightening Falsehoods: A Modal View of Scientific Understanding." In *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*, edited by Grimm, S., Baumberger C., and Sabine Ammon, Routledge, 111-35.

Levins, R. (1966). "The strategy of model building in population biology." *American Scientist,* 54(4), 421–431.

Levins, R. (1993). "A Response to Orzack and Sober: Former Analysis and Fluidity of Science." *The Quarterly Review of Biology,* 68(4), 547-555.

Lloyd, E. (2010). "Confirmation and Robustness of Climate Models." *Philosophy of Science,* 77(5), 971-984.

Lloyd, E. (2015). "Model robustness as a confirmatory virtue: The case of climate science." *Studies in History and Philosophy of Science,* 49, 58-68.

Odenbaugh, J. (2011). "True Lies: Realism, Robustness, and Models." *Philosophy of Science,* 78(5), 1177-1188.

Odenbaugh, J. (2019) *Ecological Models: Elements in the Philosophy of Biology.* Cambridge University Press.

Odenbaugh, Jay, and Anna Alexandrova. (2011). "Buyer Beware: Robustness Analyses in Economics and Biology." *Biology & Philosophy,* 26(5), 757–771.

Orzack, S. and Sober E. (1993). “A Critical Assessment of Levins’s The Strategy of Model Building in Population Biology (1966).” *The Quarterly Review of Biology,* 68(4), 533-546.

Parker, W. (2011). “When climate models agree: The significance of robust model predictions.” *Philosophy of Science,* 78(4), 579–600.

Pirtle Z., Meyer R, Hamilton A. (2010) “What does it mean when climate models agree? A case for assessing independence among general circulation models.” *Environmental Science & Policy,* 13, 351–361.

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, R.J. Stouffer, A. Sumi and K.E. Taylor. (2007). “Climate Models and Their Evaluation.” In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Schupbach, J. (2018). “Robustness Analysis and Explanatory Reasoning.” *British Journal for the Philosophy of Science,* 69, 275-300.

Steele, K., & Werndl, C. (2013). “Climate Models, Calibration, and Confirmation.” *The British Journal for the Philosophy of Science*, *64*(3), 609–635.

Stegenga, J., & Menon, T. (2017). “Robustness and Independent Evidence.” *Philosophy of Science*, *84*(3), 414–435.

Weisberg, M. (2006). “Robustness Analysis.” *Philosophy of Science*, *73*(5), 730–742.

Wimsatt, W. C. (1981). “Robustness, reliability, and overdetermination.” In *Characterizing the Robustness of Science*, edited by Léna Soler, Emiliano Trizio, Thomas Nickles, and William Wimsatt. Springer, 61–87.

Winsberg, E. (2018). “What does robustness teach us in climate science: a re-appraisal.” *Synthese*.

Woodward, J. (2006). “Some varieties of robustness.” *Journal of Economic Methodology*, *13*(2), 219–240.

Woodward, J. (2010). Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, *25*(3), 287–318.