Studies in the History and Philosophy of the Biological and Biomedical Sciences (in press 2013)

Coupling Simulation and Experiment: The Bimodal Strategy in Integrative Systems Biology

Miles MacLeod

Nancy J. Nersessian

Georgia Institute of Technology

Abstract

The importation of computational methods into biology is generating novel methodological strategies for managing complexity which philosophers are only just starting to explore and elaborate. This paper aims to enrich our understanding of methodology in integrative systems biology, which is developing novel epistemic and cognitive strategies for managing complex problem- solving tasks. We illustrate this through developing case study of a *bimodal* researcher from our ethnographic investigation of two systems biology research labs. The researcher constructed models of metabolic and cell-signaling pathways by conducting her own wet-lab experimentation while building simulation models. We show how this *coupling* of experiment and simulation enabled her to build and validate her models and also triangulate and localize errors and uncertainties in them. This method can be contrasted with the *unimodal* modeling strategy in systems biology which relies more on mathematical or algorithmic methods to reduce complexity. We discuss the relative affordances and limitations of these strategies, which represent distinct opinions in the field about how to handle the investigation of complex biological systems.

Keywords: simulation; experimentation; model-building; systems biology

Coupling Simulation and Experiment: The Bimodal Strategy in Integrative Systems Biology

1. Introduction

Integrative systems biology (ISB)[1] represents a new mode of biological research. It applies and develops mathematical and computational methods and new data collecting technologies to model large scale biological systems. These systems are extremely complex and research is often characterized by many other constraints, such as the lack of adequate data and the lack of reliable theoretical frameworks for building models. The large multidimensional space of possibility covering unknown network structure and unknown parameters, often demands innovative methodological strategies to constrain and search through. These strategies, in turn, rely heavily on the affordances of computational simulation.

Most philosophical attention to systems biology so far has been directed towards its philosophical foundations, and topics of antireductionism (including emergence) and explanation in systems biology, with less attention to the methodological practices systems biologists use to manage the complexity of their systems and construct reliable computational models. The cognitive demands of investigating such complex systems require, as Simon and Wimsatt have argued more broadly, complex model-building tasks which rely for instance on heuristics that help search through and constrain possibility spaces. (N.J. Nersessian, 2002, 2008; Newell & Simon, 1972; Simon, 1979, 1983; Wimsatt, 1975, 1987, 2007b; Winsberg, 2009), The widespread incorporation of simulation into model-building practices is now enhancing the ability of researchers to investigate the dynamics of complex systems (Lenhard, 2006; Winsberg, 1999, 2003, 2010). Systems biology however has characteristics that ramp up its cognitive difficulty. As an interdisciplinary or transdisciplinary field it depends on integration of concepts, methods, and expertise from multiple disciplinary sources (O'Malley & Soyer, 2012), but lacks reliable model-building theory that operates across the different data types and quality that researchers in physics-based fields typically have (XXXX) .

Our aim in this paper is to use a detailed case study to examine how the growth of computational modeling of complex multi-level phenomena in biology is generating new methodological approaches in response to the increased cognitive demands on researchers. Specifically, we are interested in how integrative systems biologists manage the specific complexity of the systems they are investigating, and the constraints on model-building,

[1] Systems biology researchers refer to their field as either "systems biology" or "integrative systems biology." We use ISB because that is what the researchers in our lab and the institute they participate in uses. As the director of one of our labs (and of the institute) expressed it, integrative "really means you put all the insights back together into one structure that you can analyze, as opposed to looking here a little bit and looking there a little bit… how do all these pieces interact with each other and how do they form a system. Like an integrated circuit."

through the affordances of model-based reasoning (Nersessian 2002, 2008), the model-building process, computational simulation, mathematical inference and biological experimentation. We want to highlight that the methodological flexibility available to the developing field affords systems biologists the opportunity to tailor methods to specific problems. This is leading to a diversity of methods that suggests the need for more intensive case studies of different model-building practices in ISB in order to better evaluate the epistemological and representational characteristics of these models. Different strategies make different assumptions and abstractions, and work at different degrees of closeness to the biological system.

In this paper we outline the results of an ethnographic study of a particular researcher C9 working in a systems biology lab that specializes in metabolic and cell-signaling systems, which we have designated "Lab C." This lab conducts both modeling and experimentation. We tracked C9's research over the last three years of her doctoral studies. She pursued what we call a *bimodal* strategy, simultaneously developing simulation models of metabolic and signaling pathways and conducting experiments on those pathways. This is a novel and difficult strategy, requiring both high-level modeling and experimental skills, pursued by a small but significant number of researchers in ISB (not just in Lab C). We document how this strategy served the purpose of constructing models of a set of systems given both the complexity of these systems and the data and other constraints under which she was operating. She tightly integrated these two modes into a system for generating and validating information about her systems. This *coupled methodological system* enabled her to handle cognitively the complexity and uncertainty of her biological systems, by using computational simulation to help direct and focus her experimental manipulation and investigation of the physical system and wet-lab experimentation to build her models. In –particular, coordinating experiment and simulation allowed her to triangulate uncertainties and missing elements in her models without having to search through complex spaces of alternatives. [2] We contrast this strategy and its affordances with respect to other more mathematically and computationally driven strategies in systems biology.

2. Systems Biology in our Labs

ISB is an emerging field tracing back 15 years or so which presents itself as a "new wave" of biological research at the intersection of biology, engineering, and computation (Kitano, 2002).

[2] There has been some recognition of the system-like nature of modeling, simulation, and experimentation in computational biology. In particular, Carusi et al (2012) argue these should be viewed as forming an "MSE" system since each interacts with the other in the discovery process. While they are correct in asserting this, theirs is a generic methodological account that could be used in describing either the unimodal or bimodal strategy. They do not provide detailed examination of their interaction in case studies and do not discuss the novel bimodal strategy we examine here. We thank Sara Green for directing our attention to this reference after reading an earlier version of our paper.

Most research in ISB focuses on gene regulatory pathways, and metabolic systems and cell signaling pathways. Although "systems biology" emerged as a philosophical position in the mid-20th century, the prime methodological innovation of ISB has been to integrate engineering concepts of "system" and that field's mathematical tools for analyzing systems with modern-day computational processing and high-throughput technologies for data collection (Kitano, 2002; Krohs & Callebaut, 2007; O'Malley & Dupré, 2005). Doing so facilitates understanding and control of biological systems of a much larger scale than can ordinarily be obtained through the experimental techniques of classical biology, such as molecular biology which focuses on *in vitro* experimentation on small numbers of system components. One of the central claims of systems biology is that properties and biological functions of components are dependent on their part within systems (Kitano, 2002; Hans V. Westerhoff & Kell, 2007). Because of the complexity of these networks (due to the many interacting components and nonlinearities in the form of feedback relations) they need to be investigated *in silico* as well as in the laboratory.

There is already some work attempting to formulate an account of the philosophical foundations of systems biology and the epistemic and representational characteristics of systems biological models. (see for instance Bruggeman & Westerhoff, 2007; Fagan, 2012; Krohs & Callebaut, 2007; O'Malley & Dupré, 2005; Richardson & Stephan, 2007; Hans V. Westerhoff & Kell, 2007; H.V. Westerhoff et al., 2009). This has necessitated characterizing the methodology of the field. Researchers have for instance identified two main streams of systems biology – top-down and bottom-up (Bruggeman & Westerhoff, 2007; Krohs & Callebaut, 2007; Hans V. Westerhoff & Kell, 2007). Top-down systems biology reverse engineers the structure and components of very large-scale systems with the aid of high-throughput data and sophisticated computational algorithms. Bottom–up assembles models from data on the network structure of a system and properties of the network elements. In both cases the understanding of how models get built in systems biology often presumes that systems biologists have access to high-throughput data collection methods and massive quantities of date for their model building. After all, it is the combination of mathematical modeling with these 'omics and other data methods that is thought to have enabled systems approaches.(Ideker, Galitski, & Hood, 2001; Krohs & Callebaut, 2007; O'Malley & Dupré, 2005; Hans V. Westerhoff & Kell, 2007)

However in the cases we have studied systems biology rarely works like this (see also Bruggeman & Westerhoff, 2007; Wimsatt, 2007a). The picture above represents in fact only one strand of systems biology. As Voit et al. (2012) point out the predominant mode of modeling in systems biology, which they call "mesoscopic modeling", operates without rich high-throughput data. Our own research has been directed towards this strand. Our claims in this respect are the result of a 4-year ethnographic study of two biochemical systems biology

laboratories located in a major research university in the United States. In one lab ("Lab G")[3] the researchers do only modeling in collaboration with experimental researchers outside the lab. What we call the *unimodal strategy* is the predominant mode of working in this developing field. Modelers and experimentalists undertake complimentary but different tasks in the biological integration project. In Lab C, as we noted above, researchers conduct their own experiments in the course of building models – what we are calling the *bimodal strategy*. Both labs are populated largely by graduate students, who in both cases come from engineering or applied mathematics backgrounds. We will say more about our data collection and analysis methods in section 3 where we develop a representative illustration of the coupling of experiment and simulation in the bimodal strategy.

The systems biological model- building process in both labs cobbles together a mix of different kinetic rate and concentration data sources, different pieces of biological knowledge (often also from different sources), a variety mathematical and algorithmic techniques (many self-generated), and lots of what our researchers call "informed guesses." These are all combined in the search for good approximate pathway representations, different models of biochemical interactions, and best fitting parameter sets (XXXX) . The results are often highly approximate and abstract, but produce numerous predictive successes. The context of modeling in our labs explains why this is. Modeling in these contexts is an "art" (a description favored by our researchers) of *managing complexity* which is characterized not only by complex highly non-linear systems, but also by significant constraints on the researcher's ability to unpack and understand these systems. These constraints need to be integrated into problem solving strategies (XXXX) . A sampling of these are:

Data Constraints: No modeler we have encountered has ever had sufficient data from the experimental literature or from collaborators for what they are trying to achieve. Although time-series data for molecular concentrations, for instance, are valued for dynamical model-building these are rarely available, for the simple reason that molecular biologists do not always find such data of value within their own field and either do not produce it or do not keep or share with modelers the excel spread sheets of the raw data they are not interested in. Our modelers generally do not have independent access to high-throughput technology. What the modelers tend to have is jumble of different types and sources of data collected from their own search and analysis of the literature and, if they are lucky, from their collaborators. This lack of data generates significant uncertainty.

Collaboration Constraints: Our modelers almost always encounter difficulties trying to collaborate with experimentalists in order to get the data and the biological expertise they

[3] This research is supported by the US National Science Foundation and is subject to human subjects restrictions that prevent us from identifying the researchers.

need (since they are rarely familiar with the biological system they are modeling). For one thing, the collaborators operate at different time scales. Running a wet-lab experiment takes much longer to do than running a simulation, which often leaves modelers sitting around waiting for long periods in order to make their next step. Lacking biological backgrounds, modelers tend not to understand what experiments are technically feasible within the competence of their collaborators. They do not understand the constraints on doing experimentation such as the limits on available time and financial resources, or the fact that the experimentalist might actually "have passed the point" of that research and are already working on something else. These limitations often have to be worked around by modelers.

Computational Constraints: There are multiple computational constraints, for instance, those stemming from computational processing speeds which often mean that algorithmic methods for parameter-fixing need to be optimized for the particular system.

The complexity of the pathways plus these kinds of constraints means that researchers are confronted with complex-problem-solving tasks, and their problem-solving strategies must effectively manage and contain this complexity. The cognitive load is particularly high with such nonlinear systems because it is hard to infer the consequences of changes to a model, and thus hard to conceive a viable restricted search-space that can lead to robust solutions. In general, generic methods for dealing with the range of types of data a modeler might have and their relative incompleteness are not available, which means modelers have to fashion strategies, mathematical or otherwise, to get around this problem for their specific systems.

As such the researchers in our labs face a relatively unstructured model-building process. Problem-solving methods often need to be adapted or created for the particular system and the particular constraints with which they are operating. There are, in fact, many choices of how to go about this process, such as choices of the mathematical framework and modeling techniques to apply (such as differential vs agent-based), the computer language to use, what to include in the pathway structure, what to leave out, which parameter-fixing method to use and so on. All of the above leads to the situation that ISB is complex methodologically and philosophers do not yet have a good understanding the different methodological approaches by means of which model-building in systems biology manages these complex problem-solving tasks. However evaluating ISB, whether in terms of the explanatory nature and value of the models it produces or the epistemic principles guiding it, depends on developing a better understanding of the epistemic strategies researchers use to produce their models. Our own strategy for getting at the philosophical issues has been to conduct long-term, detailed ethnographic studies of ISB researchers "in the wild," as they go about their research.

3. Discovering the sensitivity of cancer cells: C9's investigative pathway

Here we provide an exemplar of the *bimodal strategy* through developing a case study of one Lab C graduate student researcher, C9. Lab C is situated in a biomedical engineering department and all of Lab C members have various engineering back grounds, with the exception of the lab manager (who has since transitioned to a graduate student) whose background is in biology. When we entered Lab C, only two researchers and the lab director (who was trained in a bimodal lab) were using the bimodal strategy, but during the four years all of the researchers, including some new postdocs, have adopted it as a way of managing the complexity of their research problems.

In our ethnographic investigation of Lab C we conducted open and semi-structured interviews with all lab members, which we audio-recorded and transcribed. We carried out many hours of field observations in the Lab C wet-lab, taking notes, and attended and recorded lab meetings. Additionally we collected artifacts related to the lab history and the on-going research, including grant proposals, power point presentations, posters, paper drafts, sketches, publications, dissertation proposals and completed dissertations. All of these materials are archived in an on-line data base.

We used a range of qualitative methods including grounded coding (Corbin & Strauss, 2008), thematic analysis (Howitt & Cramer, 2008), case study (Patton, 2002) and cognitive-historical analysis (Nersessian 1995, 2008). The case study we develop here represents a triangulation (meaning here a means of "validating" an interpretation through the use of multiple sources of data and/or methods (Lincoln & Guba, 1985) of insights from these analyses based on our various data.

The path C9's research took over the course of her PhD is central to understanding the role the adoption of a bimodal strategy played in directing and determining her investigations. Through this strategy she was able to leverage affordances of both simulation and experimentation as an effective means for handling her complex problem solving task. As is common in this research, her research took a circuitous or serpentine path, driven by how the dynamics of her coupled methodological system was able, in particular, to generate novel relevant phenomena.

In the descriptions below we have changed the names of the principal chemical agents and systems involved to preserve C9's anonymity. The overall aim of her research was to try to explain different sensitivities in cancer cell lines to chemotherapy drugs, specifically to what we will call drug "X." A clinical researcher at a medical school had brought the problem of differential sensitivity of specific cancer cell lines to X to the attention of C9's lab director (C4). C4 hypothesized that the sensitivities were somehow related to signaling functions of Reactive Oxygen Species (ROS) such as hydrogen peroxide within cells, on the basis of two plausible assumptions: first, that this signaling system is sensitive to drugs like X, which generate hydrogen peroxide, and second, that this signaling system modulates pathways relevant to cell

apoptosis (self-initiated cell death) and proliferation. In the course of her research C9 constructed four models (which we have labeled Model 1 to Model 4). These were constructed consecutively and form the tasks around which her research was organized. Although C9 carried out all the model-building and experimentation for her project, as the research progressed she had extensive discussions about how to interpret what she was finding and how to proceed with her advisor C4 and a senior biochemist from another institution who had become an informal mentor to C4 and later joined C9's dissertation committee. We present abbreviated descriptions of Model 1 and 2, and focus on the latter models which required her to do wet-lab experimentation.

3.1 Phase 1: From local simulation to global simulation

Her research began with the task of modeling a particular pathway thought to be an important instance of how the redox environment of a cell (the balance of oxidants and antioxidants, or of oxidized and reduced chemical agents) affects signaling within the cell. The particular signaling pathway with which her research began is that of the activation pathway Y, which has a particular transcription factor involved in programmed cell death or apoptosis. C9 framed the "working hypothesis" of the research thusly:

So our working hypothesis has always been that, some cells are preferentially …resistant to X [drug] because [the drug] does something that leads to signal transduction with the cell that leads to, you know, anti-apoptotic transcriptions or something like that. And we know in the literature also that there are certain points in the Y pathway that are ROS … regulated. So then it didn't take too much to say, ok if you have this drug that induces ROS it is a possibility that the ROS that's induced can affect this pathway within this cell that might lead it to be pro-survival.

Her first year was dedicated to researching an accurate topology of the Y pathway by searching through the literature and looking up rate constants and chemical concentrations, putting together an ODE (ordinary differential equation) model on the basis of these, and then testing the model experimentally against published data. By the end of the first year C9 and C4 reasoned that they had a "pretty good model" that simulated the interaction of the products of ROS processing with Y itself and the regulation of these processes.

On presenting this model at conferences the reaction took two forms. On the one hand there was encouragement for the basic concept this model seemed to illustrate: Y is redox regulated. On the other hand there was resistance to the fact that the model represented such a small fraction of in vivo physiological process. At this point their biochemist mentor encouraged them to shift their attention from the small Y model to the whole system of redox regulation itself; that is, from the entry of ROS, such as hydrogen peroxide, into the cell through to the processes by which ROS are processed and cycled. By situating the Y model within this larger

global system they would have a more accurate and realistic understanding of the smaller one through simulating, in particular, the environmental factors influencing the Y model's various inputs or control points. Model 2 was again based upon a process of building out the extended biological pathway from the existing literature. The pathway diagram from which her simulation model was constructed (Figure 1), appears in her dissertation defense presentation, announced as the "first ever comprehensive account of the mammalian antioxidant system."

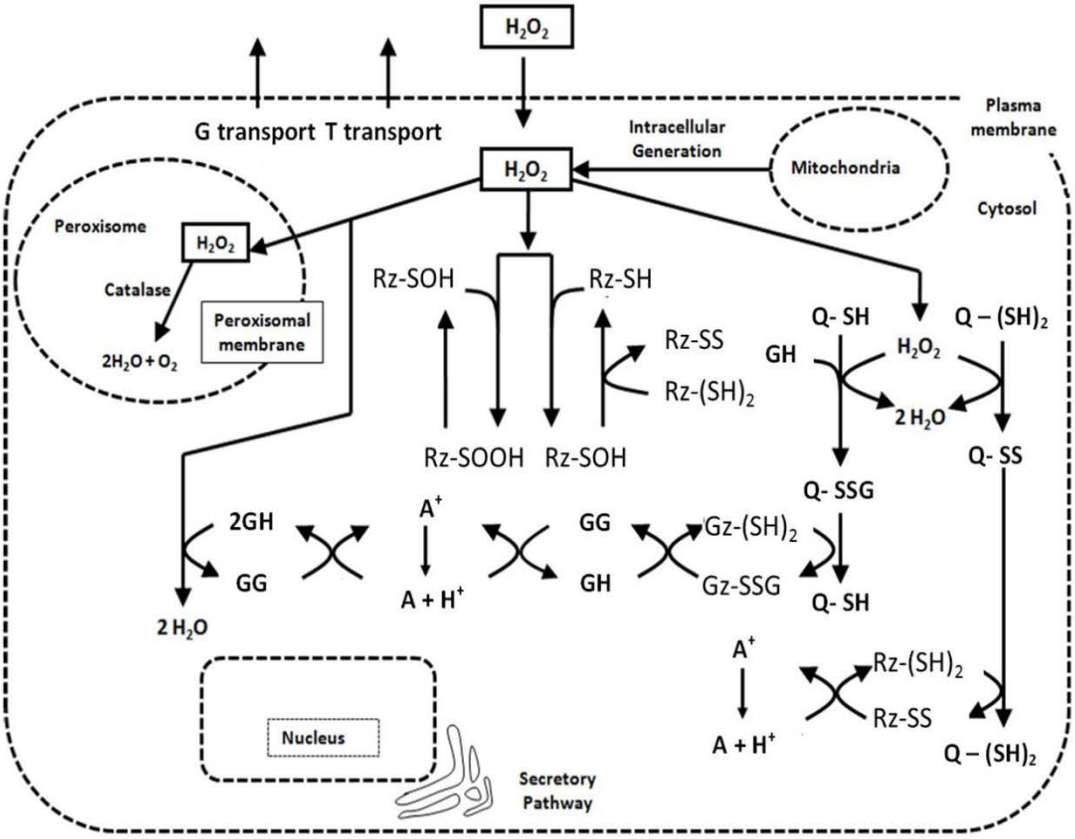


Figure 1: C9's pathway diagram for redox cycling (Model 1). The Q's are enzymes. Rz, Qz and Gz, are reductases, i.e, enzymes that catalyze reduction relations.

Model 2 was an ODE model that used a simple model of enzyme kinetics (Michaelis-Menten[4]) commonly assumed by modelers of metabolic systems, to describe the changes in

[4] Michaelis-Menten kinetics is a particular model of enzyme kinetics or the rate at which enzymes catalyze a particular reaction. Michaelis-Menten postulates that the reaction rate $v = V_{max}[S]/(K_m + [S])$ where $[S]$ is the

concentration of cellular redox buffering components. It contained 4 branches or pathways of H_20_2 elimination (see Figure 1). The model follows hydrogen peroxide's entry into the cell and the processes of redox buffering. These processes eliminate incoming hydrogen peroxide. Although the structure of the system and its participating elements were reasonably well accounted for, this move to a "whole cell" perspective and general model of redox buffering multiplied the number of components, which in turn multiplied the number of rate constants and concentrations that needed unearthing. C9 spent nearly two years foraging through the literature for these parameter values.

Model 2's output did manage to simulate the basic dynamics recorded in the experimental literature for key proteins in the network such as G and A. C9 was able to draw upon the successful simulation to make a number of inferences, including a sensitivity analysis of parameter responses.

3.2 Phase 2: Discovering the cancer drug 'X' system

In the two instances above, the structural features of the system had already been well-established in the literature and C9's task was more or less that of assembling the information in the form of a dynamic model. Shortly, as we will see, C9 was forced into the situation of having to derive, experimentally, unknown features of the system herself. In the process, she discovered that her expectations, based on the literature and her previous models, about how the system should behave were wrong and the essential mechanics of the systems were in fact unknown, and she had to localize and isolate these inaccuracies. This part of her discovery process started with Model 2, the global model of ROS.

C9's plan with Model 2 was to move on to build a mechanism that could confirm their hypothesis that it was redox regulation of drug X that was up-regulating Y-activity thus helping insensitive cells survive the drug. She planned to model this by perturbing the values of hydrogen peroxide entering the Model 2 system, given X was known to raise hydrogen peroxide levels, and then feed the outputs into the redox sensitive points of the Model 1 pathway. Knowledge of drug X's operation in causing toxicity was in fact limited, although it was thought to intercalate DNA. One of its main known side effects was though the production of ROS, making it a candidate for C9's analysis. But no one in the field knew at the time precisely how X worked, so she began experimental research to obtain the data needed for her simulation.

For the wet-lab experiments they obtained twopatient cell lines EU1 and EU3 from the clinical researcher at the medical school. EU1 was X-insensitive and was retrieved from a patient who

---

concentration of the substrate molecules (the molecules acted upon by the enzyme), V_{max} is the maximum reaction rate, and K_m is the concentration at which v is half of V_{max} measured empirically. This model's most important property is its prediction of a saturation rate for $v = V_{max.}$

had not responded to X treatment. EU3's patient, however, had had a good response leading to remission. As mentioned previously, C9 hoped to combine her global model (Model 2) with the Y pathway (Model 1). However sometime early in the course of putting together the findings from the experimental research and the model in writing what she hoped would be a paper published in an experimental journal of her work linking X to the Y pathway a problem emerged:

> *And this one [the experimental paper],the issue that I think we are having or, I don't even know if it's an issue, not really sure yet, 'cause we, we don't know what's going on, is that, the cell lines are what would be logically expected based on literature and what not, with regards to the cell that's insensitive to X should have more Y activation--- And that's what our model is sort of predicting but experimentally, we are seeing kind of the opposite and we're not really sure how reconcile that ….*

This observation from her initial experimentation led C9 to an extended novel investigation in which she went back and forth between modeling and experimentation. Her experiments with X were producing data that were the opposite of what their model was giving and of what would be expected from the existing experimental literature. Y was expected to be more active or up-regulated in the case of the insensitive EU1 cells, and down-regulated in the case of the sensitive EU3 cells. An experiment she ran however found more up-regulation in the EU3 cells. She used experiment to isolate the problem. There was more ROS in EU1 cells due to X, as expected, ruling out something in the general ROS or Y mechanisms of the cells. She found an important paper that associated the generation of the toxic form of X with an ROS reducing enzyme A. Her further experiments showed that A levels were low in EU3 and high in EU1. This suggested that both toxicity and the extra ROS were emerging from the mechanisms by which X was activated into its toxic or reduced form. As a result C4 decided they needed to model the production of ROS by X (i.e. build a new model we call Model 3) rather than simply inputting the estimated amount of ROS produced straight into the global model (Model 2). The new model-building would serve to open up an area of Model 2 otherwise black boxed-in order to look for differences between EU1 and EU3. Instead of combining

> *all these reactions into one single arrow ["black-boxing"] and then just have an estimate of what the culmination of all of these reactions would be -- we realized that there were areas where there are differences between the EU1 and EU3 cells particularly with their[enzyme] A.*

Being forced away again from her intentions to go back to the Y model, she turned to focus on the process involved in the reception of X into the cell and the production of ROS. She relied on a two pronged strategy. First she constructed this mechanism in close conjunction with these and other experiments she conducted on different interactions to "build out the model." For

instance she investigated the role of SOD, O_2 and CPR experimentally while feeding X into the different cell lines and comparing the results. As she went along, she simulated the hypothetical mechanisms she developed in the process and checked their output with controlled experiments on the different interactions. Finally she ran simulations to fix parameters. This process led her to model a new bi-modular structure for X metabolism (see Figure 2).

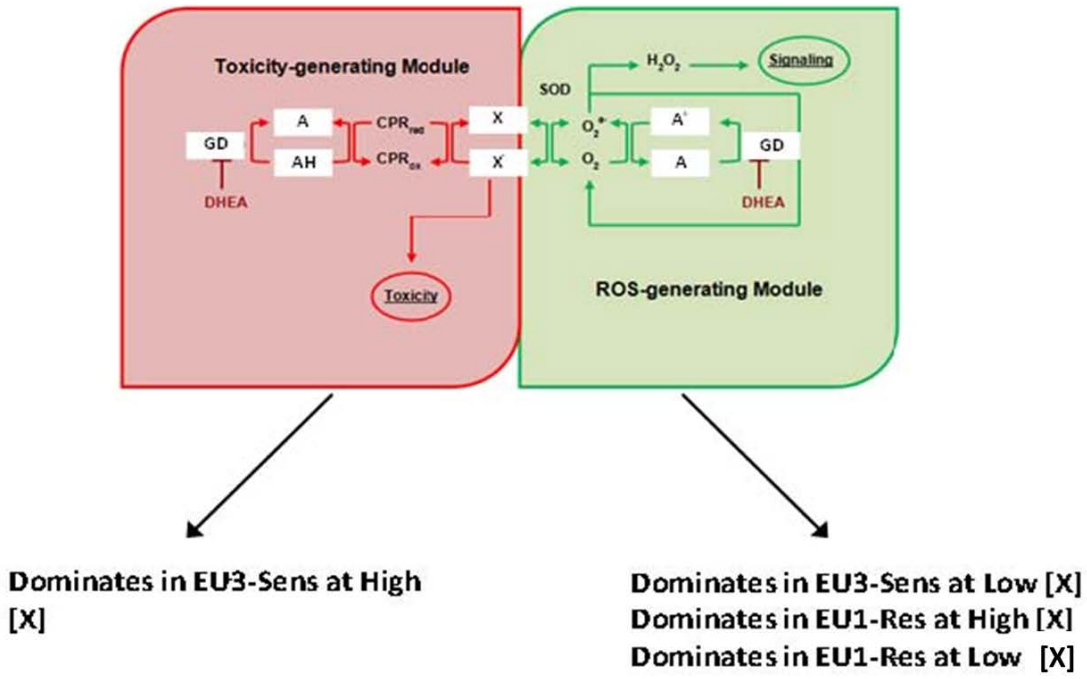


Figure 2: C9's proposed bimodular model of X metabolism in leukemia cells

In the second place once a model was up and running she checked it by perturbing the model (running simulations) and performing corresponding controlled experiments. She "physically experimented on her model" (as she called it). For instance she conducted simulation experiments to rule out other possible causes such as efflux of the drug from the cell at different rates in the different cell lines and also "experimentally perturb[ed]" the model by simultaneously inhibiting key elements of the cycles (according to the model), namely the enzyme GD and super-oxide (SOD), to test the model outside standard clinical ranges. At the

same time as running these simulations she conducted wet-lab experiments on the two cell lines using chemical agents to inhibit physically these elements and record the output.

The Model 3 simulation predicted that in the case of low A, redox-cycling took place. When A is high however, A will absorb the oxidative species that would otherwise be reduced by toxic X (X^- in Figure 2) leaving toxic X to go free. The reductive conversion/toxic X production model is represented by the left module in Figure 2, reproduced in several of her presentations. The redox recycling module is represented by the right module.

The success of this close coupling of model simulation and experiment gave her the license to assert confidently that they had discovered the mechanism behind the relative sensitivity of these cell lines represented by Model 3. She traced back the cause of sensitivity to the levels of GD each line possessed and thus whether or not the line could replenish stocks of A quickly to keep cycling X. The result C9 stated had immediate clinical relevance in so far as GD is regularly measured clinically in that its level could be used as a signal whether X treatment is likely to kill a patient's cancer cells or not.

This coupling of experimental and modeling work formed the subject matter for the second paper they published. There is a final twist in this story however. Having sent it to a well-known computational biology journal they were surprised that the reviewers complained that the levels of X they were using to study their cells were higher than those used clinically (even though they had seen plenty of papers using their levels). So they went back and used lower values with the cell lines. Experiment produced radically different behavior, which -- somewhat surprisingly -- their model reproduced. The differences in behavior can be seen in these two experimental measurement graphs. The first measures X production when X treatment levels were high and the latter when low.

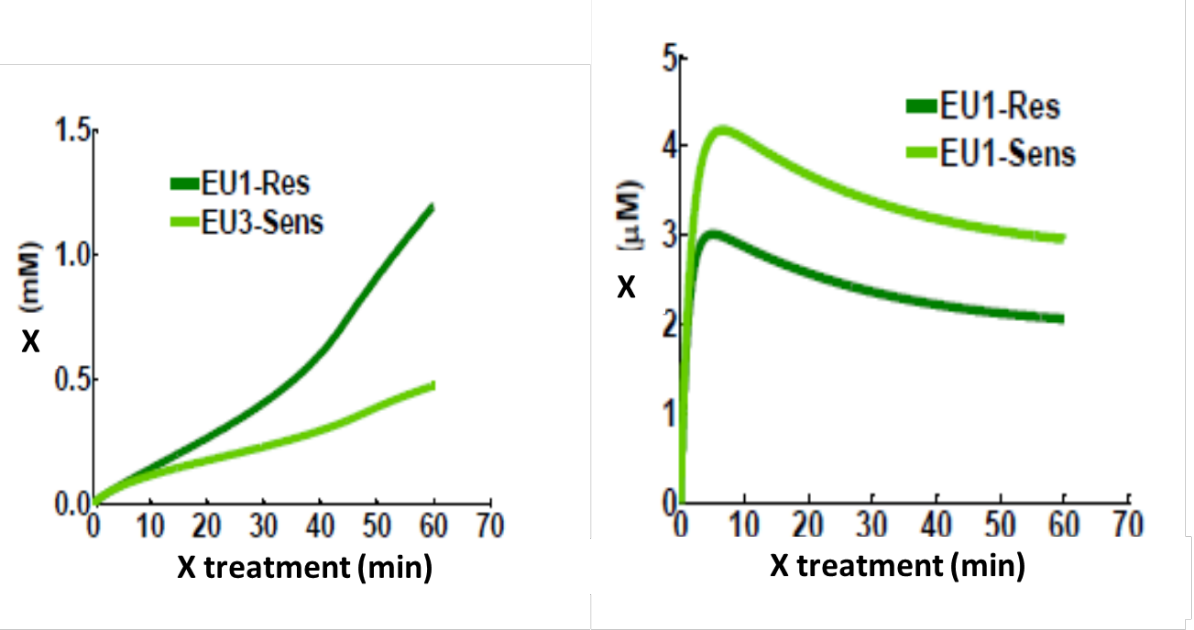


Figure 3: Experimental results of X treatment.

With low X treatment both cell lines exhibit essentially the same behavior, although at different levels. That her model was able to reproduce these experimental observations was considered a powerful validation for her discovery of a new mechanism.

3.3 Phase 3: Wrapping up, more surprises, more discoveries

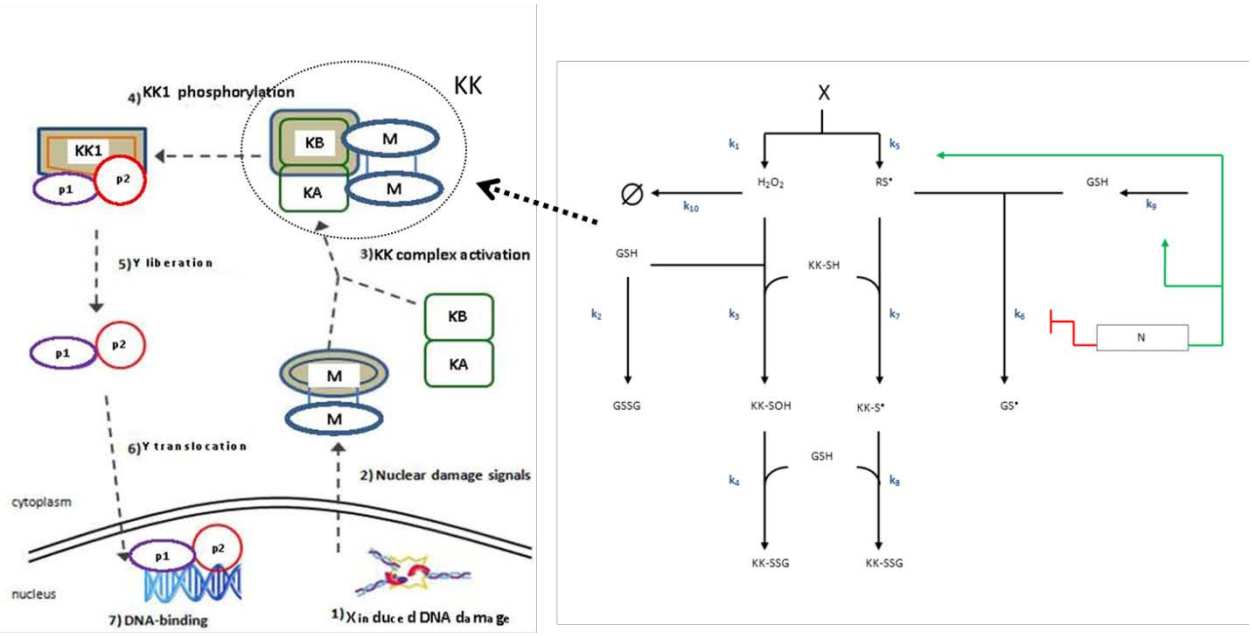
Although she had discovered a novel mechanism to explain drug X sensitivity, with some help of course from the literature, even in the last phase of her research she still found herself dealing with unexpected and unknown behaviors that required simultaneous experimental and modeling work. In this phase she had planned to go back to her original question, namely the redox regulation of transcription factor Y. There was still the question of how EU1 cells were handling both the extra oxidative stress these cells were generating under X treatment and how they were surviving whatever toxic X they were generating. The answer had they had though all along lay in the redox regulation of the Y pathway.

She already had a model for Y (Model 1), so the issue was to establish a connection between it and the other models. To do this she started a new line of experimentation directed at confirming Model 1's predictions. First she established that X treatment in EU1 cells created higher levels of hydrogen peroxide, using dyes. Second, she showed that X was correlated in these cells with increased Y production. When antioxidants were introduced at the same time Y

went down again, establishing the relation between X-induced ROS and Y levels. C9 thought this kind of experimental detail or "fine resolution" necessary to convince other researchers of these causal relations. It enabled her to show that the causes she was hypothesizing in her models were robustly responsible:

it's kind of like there's nothing written in stone about the steps you take. You need to sort of say to yourself, ok, how fine of a resolution am I comfortable with, or how fine of a resolution to actually need to… to get other people to believe this is actually what's happening.

Having established these relations experimentally it was time to get into the "nitty gritty" details of the Y-pathway itself. Going back to Model 1 the question was which pathway elements would be modulated by the increase in ROS or what could be the potential points in the Y network that govern redox regulation. Using the model as the basis, she planned to work through wet lab modulations of the different components using a particular antioxidant we label N. Her initial purpose was to confirm by soaking up oxidants which pathway modulations must be due to excess oxidants caused by X and which were not.



Y- Pathway (Model 1)

KK s-gluthionylation pathway (Model 4) with N interaction

Figure 4: The left pathway is C9's pathway model of pathway-Y (Model 1). The right pathway represents the pathway for the KK s-glutathionylation system with the action of N built in, where KB is the active component of the KK complex. The KK complex is circled on the left pathway. On the right the values of fixed parameters k_5, k_6 and k_9 were used to help reveal the action of N. The values of KK predicted by the right pathway in response to X and anti-oxidation could be used to modulate KK in the Y-pathway.

Going through elements of the Y-pathway with a more detailed wet lab examination (Figure 4) adding X she discovered one protein KB whose s- glutathionylation levels changed when X was added. S-glutathionylation modification is caused by oxidation of the protein and thus provides an indication of sensitivity of that protein to oxidative agents. In this way she isolated experimentally the KK complex (which binds KB) as an ROS sensitive component, circled in the Y-pathway above. C9 confirmed this in the wet-lab by adding N and seeing KB levels drop accordingly. However when testing with N more expansively she discovered that the effects of KB varied non-monotonically with the levels of N in the system. This was non-intuitive for them and entirely unexpected, and put them again in the situation of having discovered complexity which they had not expected, and needing to draw on both her experimental and modeling resources to untangle the knot. C9 and C4 decided they needed to build out Model 1 by adding a model of KK s-glutathionylation (Model 4) to explain the wet lab findings. This model, which C9 referred to in her dissertation as "mechanistic" because it worked quite closely to the chemical details, targeted the "mechanisms by which a protein such as KB can be s-glutathionylated in the presence of an ROS promoting agent like drug X."

C9's literature search yielded up 10 possible biochemical reaction mechanisms by which KB s-glutathionylation could occur. She used both the conditions under which those mechanisms were observed in comparison to the mechanism she postulated and also those conditions that were was possible under the treatment conditions she was employing in the wet-lab experiments to narrow it down to three possible mechanisms. These candidate mechanisms were all then modeled via simulations of different concentrations of N, and fit to the experimental conditions to choose the best of them. The fitting process revealed certain modified parameters that suggested N was having both pro- and antioxidant effects (see Figure 4). With these, more literature searches, simulations and fittings she was able to postulate a mechanism by which N operated on glutathionylation process. In this way she managed to reproduce the nonlinear response of the KB protein. Above certain concentrations it has pro-oxidant effects. In clinical terms this means that N might actually be counterproductive in treating cancer in certain instances. This work proved to be the final research of the thesis, and the center of her third paper. However, the primary aim of discovering how manipulation of Y through ROS signaling affected cell viability remained an open and unresolved question.

The result of her PhD research, as she saw it, when looked at as a whole, was to reinforce the idea that "redox mechanisms do play a role in chemotherapy administration and more attention should be paid to those mechanisms…." She characterized her work as having opened up the discussion without necessarily coming to any precise conclusions that could be used clinically. In the end, they never managed to quite make the final connection between the

Y pathway and the insensitivity of EU1 cells. Time constraints in the end ruled this out, and C9 had to settle for showing that drug X was interacting with certain parts of the system without building a theory of how it produces its ultimate effects. But all she had accomplished in modeling and wet lab research most definitely merited a PhD.

4. Case discussion: Experiment and simulation as a "coupled system"

This first-hand account of C9's investigative pathway offers valuable insights into the ways in which bimodal researchers in systems biology, like C9, manage the complexity of their biological systems by coupling modeling and experimentation. The description we have provided tells the tale of an intensive research process that ranged over different systems and that took turns and detours on the way, meeting unexpected obstacles. Most importantly C9's model-building demanded new experimental information and experimental testing to succeed.

As we will see, the ability to do her own experiments when and how she saw fit expanded significantly what she could do with models, when compared with the majority of modelers that have to rely on collaboration alone. As she described it:

> *I like the idea that I'm building my model things are popping up in my head oh wow this would be a good experiment. I plan out the experiment myself and then go into the lab and I do it.*

For one thing, her bimodal strategy solved the considerable collaborative problems pure modelers face when trying to convey their experimental needs to pure experimentalists:

> *I personally think [my approach] is better only because… I could tell someone what I needed. But, if they, I think not really understanding the modeling aspect, they can't accurately come up with a good enough experimental protocol to get what it is I need.*

In collaborative relationships experimenters and modelers interact on a model but as we have documented elsewhere (XXXX) there are typically many inefficiencies involved in such relationships. Experimentalists often do not understand what modelers require because they do not understand modeling, and modelers do not know how to frame a request appropriate to the affordances and constraints of experimentation or, more precisely, of their specific collaborating experimenters. C9 of course was not confined by such problems. As a result she was able to efficiently coordinate her modeling and experimental activities, and more importantly make sure that her experimentation was well-adapted to obtaining the information she required within the constraints of her experimental abilities. The result of this coordination was that it gave her the ability to run experimentation and modeling as an effective *coupled system*. This coupling gave her the capacity to discover and extract the relevant information out

of a complex jumble of biochemical interactions, and to control and direct the information flow in the course of building her models.

4.1 The coupled system

To account for the nature and operation of this coupling we start from an overview of C9's use of models. As with other model-based reasoning (Nersessian, 2002, 2008), once an initial " model was constructed it became the main cognitive apparatus through which C9 interpreted the experimental data and advanced her research through choosing what to investigate, experiment on, and model further. She relied on the behavior of the model itself to inform on the properties and functions of its parts, particularly to reveal the existence of missing parts, while providing information only available at the system level, such as the relationships between indirectly related variables. Through running simulations and producing visualizations of the resulting variable and parameter relations, her model could provide information otherwise inaccessible to wet lab experimental practice, such as revealing which pathways were bearing most of the burden in a system or which points were relevant to controlling a network.

As we saw, in many instances C9 was running experimentation while in the process of constructing her models. When asked what was driving her results – models or experiments – she repeatedly replied that she was doing them nearly simultaneously:

> *You can say that I did my model first, but I don't, I don't see me as, I don't see it as I finished my model and then I did my experiments. I see it as kind of like I did them at the same time.*

She referred to their interaction as "synergistic" in her dissertation proposal. We detailed the aspects of this tight coupling by looking at the reciprocal roles model building played in directing her experimentation, and conversely the role experimentation played in her model building. We have outlined the way her model building and experiment interacted in these cases in Figure 5 below.

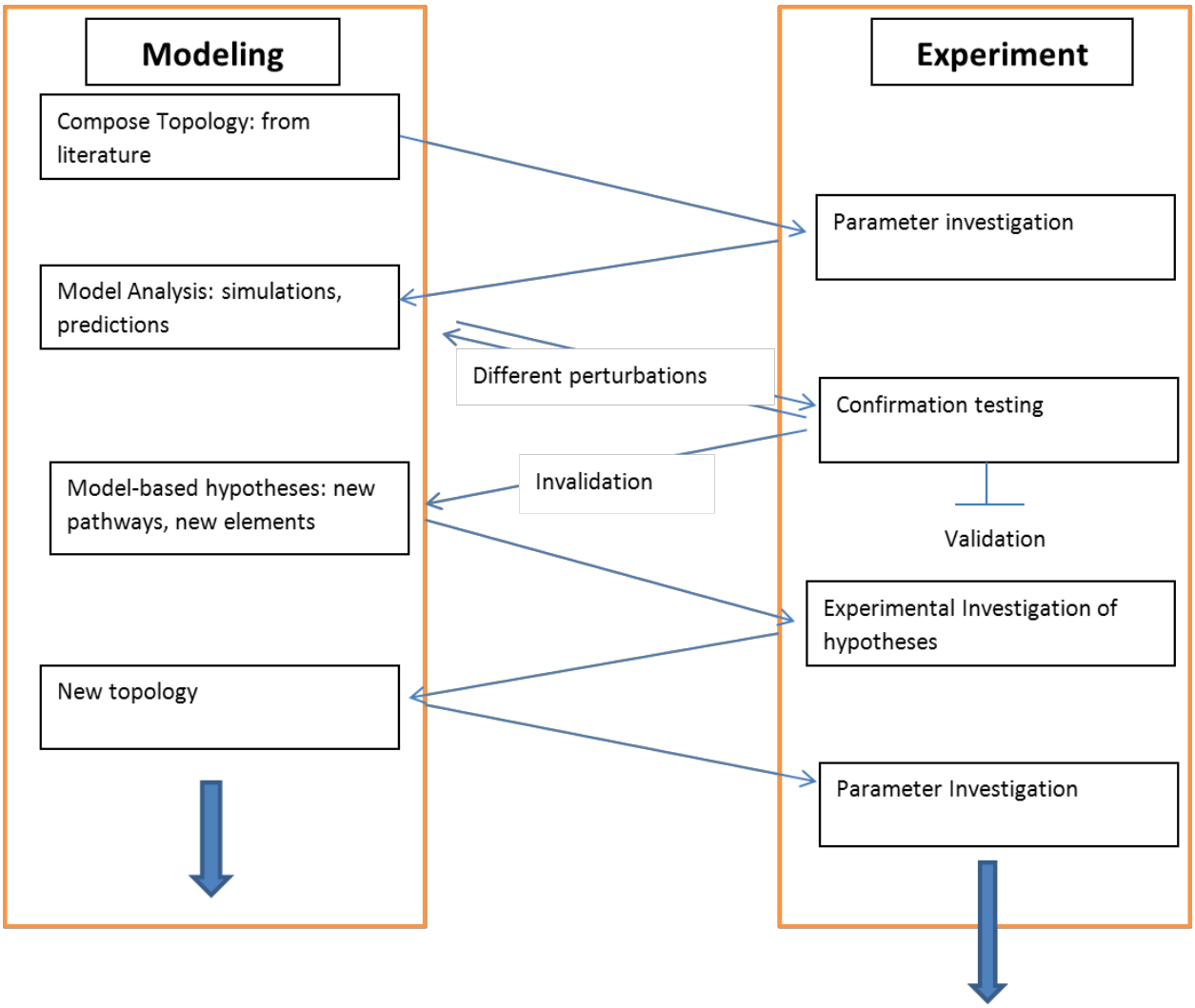


Figure 5: This diagram illustrates the coupling of modeling and experimental procedures in the course of C9's model-building as an iterative interactive process of discovering the correct topology and parameter values for her networks. The process ends when validation is achieved. Experimental results departing from the model simulations ("invalidation"), required the generation of new hypotheses, derived from the existing models.

C9's usual strategy, as evidenced particularly in the case of the Y (Model 1) and global redox (Model 2) models, was in the first place to construct a topological visualization of the biological pathway based on her extensive review of the literature. This involved tracing out and piecing together parts of the network from different sources. It also involved collecting as much parameter information as she could. Her second move then was to rely on her own wet lab experimentation to obtain pathway and parameter information that was not available in the literature. Having then put together this information in the form of a dynamic mathematical model she could begin the process of running simulations to get a feel for the model's dynamics

and also to produce predictions that she could try to verify experimentally. Because of her skills she was able to run simulations and experiments nearly simultaneously, as we saw in the drug X (Model 3) case, by replicating interventions and perturbations in her models in the physical systems, for instance, with the aid of known blocking or suppressing agents.

The systemization and organization with which she integrated experimentation and model-building had significant epistemological and cognitive affordances. Epistemically this process was generally a back and forth one for testing the model. In the case that the simulations did not replicate the experimental data she had to rely on the elements of the model she was relatively confident about, and probe those parts she was not. Hence her models and their limitations provided cognitive resources that directed further experimental activity on her part to track down solutions in the form of new pathways and new functional metabolic elements that could fill out the structure of the models. These elements would not have been apparent from an experimental program of investigation alone and could only have been identified within the context of a model and the system-level perspective it provided. For instance, when building her model of the Y pathway (Model 1) she discovered the role of cysteine by virtue of her model.

> *I sort of reached that KK, was um, needed to be disulfide bonded in order to be active. I reached that through a conjecture because I was like something is missing. When I was drawing my pictures I was like okay, there is a big gap here. How is it that KK has to oxidized to be active? And I was like "There must be a redox sensitive Cysteine."*

Conversely, when things went unexpectedly in the case of antioxidant levels in EU cells, the question for C9 was what experimental information was needed in order to uncover the links (and active chemical elements) that would get the model producing the right results

Through this coupling, the topology of these complex systems could be constructed by going back and forth between simulation and experimentation, measuring parameters, and making new hypotheses. Experimentation provided constant epistemic validation for the model and reinforcement of its role as a platform for hypotheses, and both provided a source of mental models (Nersessian 2008) for reasoning about the dynamics of the phenomena. The possibility of doing constant validation in this way created what she saw as a significant distinction between her research and "labs that do straight modeling" -- that have to go somewhere else to get validation (i.e. to an experimental collaborative partner) during the model-building processes. In comparison with these "theoretical modelers," she explained: "we don't just come up with ideas and then just shoot them out there and wait for people to do them."

We can sum up the reciprocal epistemic roles and coupled nature of modeling and experimentation in C9's research in the following table. Modeling steps or simulation were

supported in C9's practice by corresponding experimental steps, and vice-versa, which served to guide her investigation towards a representation of her systems.

Models	Experiment
Providing testable statements for experiment	Validation of model through confirmed predictions
Providing relevant classes of parameters for testing or discovery	Measurement of model parameters *in vivo* (rate constants, concentrations)
Providing modular or black box elements for experimental investigation in case of model failure. i.e 'arrows'	Discovery of disconfirming elements or unexpected behaviors.
Providing a cognitive resource for the inference of new pathways or molecular elements	Discovery or confirmation of new pathways or new molecular elements

Table 1: The coupling of modeling (simulation) and experimentation. A coupled system conceives of the actions of one both dependent on and prompting actions of the other.

4.2 Managing Complexity through Triangulation

This coupled system provided C9 a strategy for handling the complexity of the problem-solving tasks she faced, as a means for efficient and effective exploration of the large possibility space through triangulation. The concept of triangulation, defined by Denzin (2009; 297) as "the combination of methodologies in the study of the same phenomenon," is, as noted with our ethnographic research, usually understood as a particular method of validation where multiple methods are used to establish the robustness of a result through independent means. Campbell and Fiske (1959) for instance describe it as a validation process that balances the weakness of any one method against one another. As Wimsatt (2007b) perceptively notes however "processes of validation often shade into processes of discovery - since both involve a winnowing of the generalizable and the reliable from the specific and artifactual" (56) The discovery dimension of triangulation is highly significant in C9's case. By running controlled experiments and simulations side-by-side C9's was able precisely to discover reliable pathway structures novel biochemical mechanisms, and parameter values, and to build robust, validated models.

We mentioned earlier that systems biologists, because of limited data, face large possibility spaces which are difficult to search through not only because of the collaborative and

computational constraints they have to deal with, but also because the systems are nonlinear, which makes it much more difficult to determine what pathway structure or parameter hypotheses might be in the neighborhood of a solution without having to test a huge range of different alternatives. The bimodal strategy, however, provided C9 advantages for narrowing down the kind of searching she had to do. For example, she was able to diagnose and localize uncertainties in her models and to discover relevant sources of information for improving her models by simulating and comparing parts of them against controlled experimentations. This localization afforded her an ability to posit and experimentally test hypotheses about uncertain mechanisms and experimentally extract relevant lacking information.[5]

In the first place C9 used perturbations of the model and experimental perturbations of the biological system to derive strengths and weaknesses of the models she had constructed from the available literature. In particular she could isolate parts of the models and simulate those interactions and check them against experiments that isolated these relations physically in order to establish the accuracy of those parts or measure their parameter values. She could use her model to identify sources of biological information necessary for improving the model. For example the discovery of the ROS sensitivity of the protein KB was achieved by testing different proteins at different potential points of the Y-pathway experimentally with drug X. The result of this simultaneously probing models and the biological systems through perturbation and controlled simulations and experimentation was that she could robustly establish the parts of the model that were accurate but also localize the problematic parts or parts that were relevant to the model's development. Once localized, experimental work could be done to run through sets of hypotheses about the interactions of those components, as she did when formulating a mechanism for the KB s-glutathione system and the antioxidant N. Further perturbations to the biological system, by changing chemical inputs and using other chemicals to suppress particular reactions, could be checked against the dynamics the model was predicting and used to fix pertinent constraints on the model. This gave her for instance an ability to fine-tune the particular parameters of the particular reactions and interactions she was experimenting on without having to do large scale fitting.

This "model-guided" experimental localization served to localize the redox sensitivities in the Y-pathway to the KB s-gluthionylation which then further provided a model for localizing antioxidant N's non-linear behavior through experimental measurements of particular parameters in that pathway. This experiment also served to identify the unexpected sensitivities of EU1/EU3, which C9 traced through a combination of experiment and simulation to "a single arrow" in her pathway model. Each case led to the production of new localized

[5] Note our use of "localization" here differs from the use of the term by Bechtel and Richardson (2010), Wimsatt (2007b) and others. Our emphasis is on localization of errors or inaccuracies in a network rather than localization of function, although the former often serves to help reveal the latter, as it does for C9.

models for particular mechanisms. Rather than having to model all the possible 10 mechanisms for N however and trying to infer a best fit, C9 used controlled experimentations of the mechanisms to narrow the choice down to three. In the EU1/EU3 case she derived the role of enzyme A from the literature and was able to run experiments isolating and testing its role.

This coupling of methods by an individual researcher is a good illustration of the use triangulation to manage complexity. Triangulation is an appropriate notion here, because C9 relied on both information embodied in her models and experimentation to triangulate the location of inaccuracies and missing elements. It worked because she was able to build up confidence in the parts of her models through experimental testing and perturbation so that problems and uncertainties could be localized and uncertainty reduced. Doing so reduced problems to a smaller scale and a limited set of possibilities that could be hypothesized and tested through controlled experimentation. Triangulating in this way avoided the challenge, given limited experimental data, of having to find and algorithmically sort through pathway structure alternatives and parameter values (a major problem for the unimodal strategy), knowing these would often have to be refit with the entire model to the data. It avoided having to take steps of approximation or abstraction that could effectively black-box the alternatives. Staying close to her biological systems in this way C9 had an inside running on tracking down accurate representations.

5. General Discussion: Methodological Strategies of Systems Biology

This case study highlights the flexibility of approaches within systems biology and the diverse possible roles of simulation and experimentation within these complex problem solving contexts. As mentioned the systems biologists in our study operate under a variety of constraints, and no particular methodological strategy will be reliable in all conditions. The lack of general theoretical frameworks that can be consistently relied upon, given the general uncertainty over network structure and parameters, means that researchers bear the cognitive weight of handling the complexity of their systems when finding parameter-fixing strategies and making decisions about pathway structure. Lots of different variables and parameters need to be tied together often simultaneously. Deriving strategies usually means building up their own partial mental models of dominant system dynamics with the aid of simulation and using these as the basis from which to devise means of sorting through the space of possibilities for parameters and network structure. It further means that cognitive constraints limit the size of networks modelers can handle and forces them to make approximations about pathway structure and biochemical interactions. Indeed the majority of modeling we have encountered is in fact neither top-down nor bottom-up but, as we noted, better described as mesoscopic or middle-out (Noble, 2006; Voit, 2013; Voit et al., 2012). It focuses on middle-scale systems but

averages over lower-level information to limit their detail. These averaging methods and scale restrictions keep the model-building within cognitive limitations. Any model produced this way can in theory be further built-out by increasing the adding to the network or investigating the lower-level in more detail in order to correct the model.

C9's strategy should be understood as a particular attempt to manage the various constraints she faced by devising a workable method that gave her cognitive control over the system sufficient to build-up a model of the dominant dynamic relationships. But systems biologists have a range of possible investigative pathways they can take, many of which can and do produce desired outcomes. The most important determinant of how to proceed for the unimodal strategy is the type and availability of data for the biological system. Steady-state data, for instance, require different mathematical and algorithmic strategies compared with time-series data in order to understand the role and function of network elements and use that understanding to narrow down the search space (Voit, 2013) . C9's strategy was a strategy that could address at least to some extent the lack of data for her systems and accurate knowledge about the systems' structure. Strategies preference certain disciplinary skills and knowledge backgrounds, and require developing other skills and knowledge to varying degrees. The lab G director has described ISB as a space of methodological possibilities bound by computer science, engineering and biology – what we have been calling an *adaptive problem space* (Nersessian 2006).

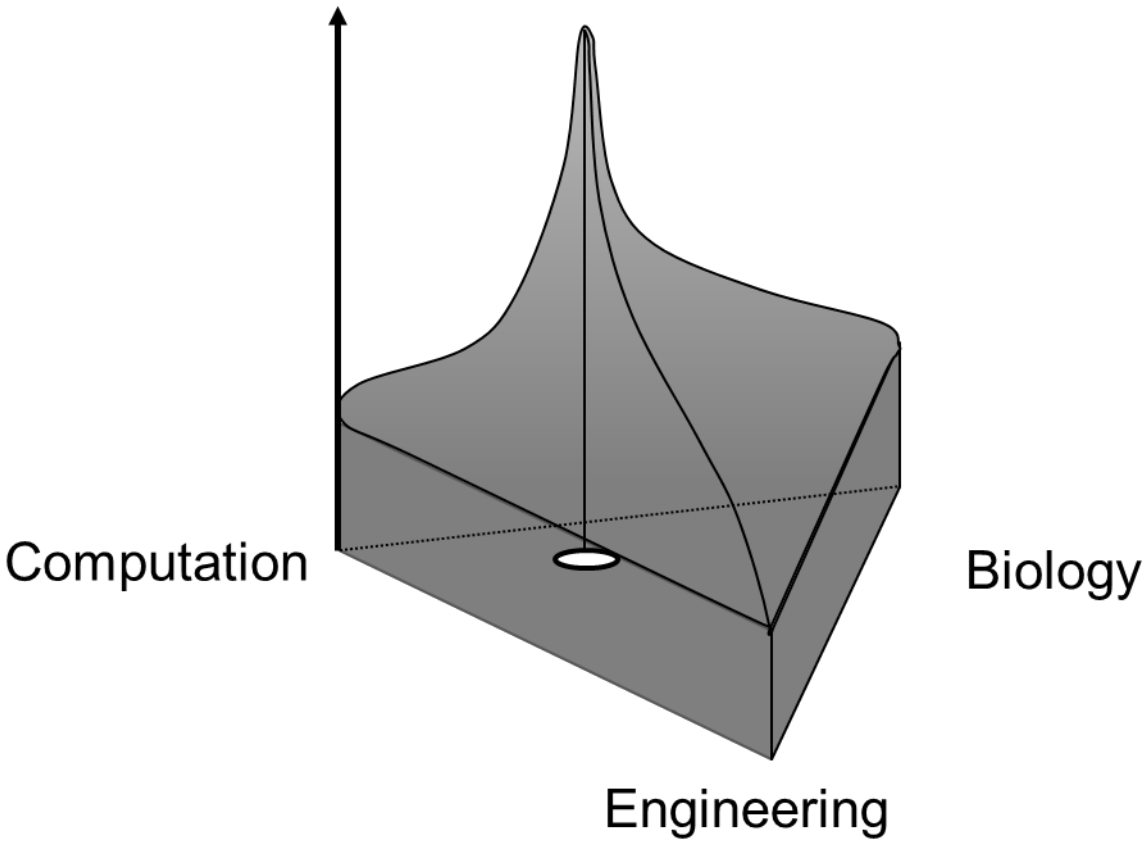


Figure 6: An adaptive problem space – researchers fashion or adapt methodologies that draw preferentially on methods from these disciplines, and thus "pull-up" at different positions within this space.

Where a researcher pulls up in this space depends to an extent on their background training, the collaborative possibilities available to them and what they themselves are willing to learn. Different researchers can choose to operate with methodologies closer to those of one participating discipline or work towards a more integrated approach that combines techniques from some or all these fields. The current methodological flexibility of ISB creates space for the diversity of backgrounds of those entering systems biology labs. Whereas C9 could use targeted experiment to investigate her network and resolve uncertainties, other researchers particularly in lab G, rely on mathematical inference to infer missing elements, and mathematical and algorithmic methods (like Monte Carlo methods) tailored to their systems to search through alternative possibilities for best fits. And while C9 could model and study individual interactions using relatively low-level enzyme (Michaelis-Menten) kinetics and measure parameters directly if required, many mathematically-inclined modelers try to fit more abstract power-law models that average out complex details of interactions and approximate

them. These more mathematical methods are strategies first and foremost that work to simplify the complexity of these networks, and turn intractable problem-solving tasks into tractable ones by simplifying the search tasks. Whatever the strategy however, understanding the dynamics for particular networks is necessary to formulating search strategies given particular data.

Of course whether the results of such estimation and inference strategies will be accepted by the biological community without experimental evidence is another vexed question. It's less of a problem for C9 however who chose to "pull up" methodologically in-between biology, mathematics and computation, with her particular hybrid strategy, and could justify her models through her experimental results.

Limitations of the bimodal research in systems biology

In ISB some kind of interaction between experimentation and simulation is required for modeling large scale systems. For the most part this interaction does not extend beyond experiment supplying adequate data for confirming model predictions in contrast to the intensive interaction in C9's research. Although the predominant mode of integration is a unimodal collaboration strategy, the collaborations we have witnessed and heard described by Lab G researchers have been fraught with difficulties, in some instances leading to the abandonment of a line of research. The bimodal modeler does have an advantage in this respect. The modeler – experimentalist team usually does not easily overcome the barriers of specialization. However we do not want to imply that a bimodal strategy is necessarily the best way to do systems biology under all circumstances.

In the first place there are limits to what can reliably be managed by such a bimodal strategy. C9's method worked well because she was dealing with relatively small scale systems, with a manageable amount of unknowns. Her experimental work was containable and she could direct her modeling to keep it within reasonable constraints. Secondly the biochemical information she had allowed her to develop detailed models of the biochemical interactions, reducing the need for mathematical averaging techniques and the potential for error they bring with them. Even so she did not get as far on her problem as she initially thought she would, and as would be needed to solve the clinical problem. A dedicated modeler, however, can handle a larger system as long as data are already available or sufficient experimenters available with whom to collaborate. The lab G director expressed in an interview that the ideal ratio is "10 experimenters for every 1 modeler." He also expressed that part of his preference for having someone else do the experimental validation research is that one might too easily see what one wants if you do both (so-called "confirmation bias").

Further, becoming skilled in both experimentation and modeling takes time and involves compromise. Many researchers in ISB come from mathematical or engineering backgrounds and have little if any knowledge of experimentation. They need to be trained as systems biology modelers, which involves not only developing new mathematical and computational skill, but skills directed towards searching the on-line data bases and other literature sources pertinent to the various systems they are modeling. As we have witnessed, learning this latter skill is no easy task, especially given that today they might be modeling a yeast system and tomorrow a cancer-producing system. And, we would add, they need to invest in developing collaboration skills much more than is currently done. Those adopting the bimodal strategy need to develop all of these (they might collaborate too), but in addition they need to learn experimental design and bench-top skills. The normal course of a PhD is 5 years. Thus, there is a "philosophical divide" (Lab G director) between those who advocate concurrent or sequential (e.g. PhD modeling, postdoc experimentation) bimodal training. Concurrent training could have the consequence of ending with "modeling lite and experimenting lite," but sequential puts the researcher in the position of basically "starting over" after several years of education. In our final, post-defense interview with C9, she did express some regret that she had not been able to work more on her modeling skills with the Lab G director, who was on her committee.

6. Conclusion

The importation of computational methods into biology is generating novel methodological strategies for managing complexity which philosophers are only just starting to explore and make sense of. ISB itself is not a homogenous enterprise, but consists of diverse approaches and strategies for handling complex and unstructured problem solving tasks, where in fact, contrary to the widespread 'omics rhetoric, the lack of data and different qualities of data are among the foremost difficulties for some researchers. Among these strategies are the bimodal and unimodal approaches. We have tried to show how the bimodal strategy can be understood as a particular response to features of the problem-solving contexts of systems biology; namely complex problems, lack of theoretical starting points, and data and other constraints. In the case of the bimodal strategy, experimentation and simulation are closely coupled in the model-building process, to validate model-building steps, but also to provide and effective means of limiting search spaces and triangulating on a good representation. It has both advantages and limitation when compared to the mathematically–based unimodal strategies we have studied. This paper is thus a contribution to building a richer picture of the epistemic practices of these new interdisciplinary biological fields and understanding the ways in which simulation is helping adapt new responses to complexity. If our goal is to be able to evaluate systems biology as a methodological system then studies like this help us understand that systems biological models can differ substantially in their connection to empirical results but also in the their degree of abstraction from the actual underlying biological system.

Although philosophers, looking mainly at physics-based fields, have made good arguments that the epistemology of simulation and experiment can, in principle, be quite similar(see Morrison, 2009; Parker, 2009; Winsberg, 2010), our illustration of the bimodal strategy provides an exemplar of how context matters when establishing their roles and relative epistemic value (see also,(Winsberg, 1999)). C9 gave much more weight epistemically to physical experiment than to simulation. She continually validated her models experimentally. She was really never in a position where she could rely on a well-validated or theoretically-supported model to conduct only simulation experiments on her systems. The uncertainty over the structure and properties of her systems prevented C9 from simply building a model by building accepted assumptions and filling numerical details into an established modeling framework. .

Finally, although it is not possible to develop this idea within the confines of the present paper, we believe it is important to note that addressing the cognitive complexity of ISB research will require moving beyond a discussion of heuristics to an analysis of how researchers "*off-load*" (Hollan et al., 2000; Hutchins, 1995) or "*distribute"* (Nersessian, 2008, 2009, 2012; Nersessian et al. 2003; Giere, 1992; Hall et al., 2010) their cognition through creating distributed problem-solving systems appropriate to their tasks. Most philosophers when discussing cognition and complexity in science have focused on heuristics that help narrow down search spaces for complex problems (Newell & Simon, 1972; Wimsatt, 2007b). Cognition is interpreted in terms of its constraints which heuristics help overcome. In the case of mesoscopic modeling it is certainly true that model-building methods are developed with cognitive limitations in mind. C9's method however was not just a matter of employing well-established heuristics, such as "false models as a means to truer theory" (Wimsatt, 1987, 2007b). The bimodal strategy also allowed her to distribute cognition through the coupling of simulation and experiment (XXXX). Running simulations provided a method for calculating and visualizing dominant network patterns, which helped C9 develop and simulate her own mental models of system dynamics. She used the understanding obtained through these to direct experimentation, manipulate the biological materials, and make inferences for interpreting the results, thus turning experimentation into a sharper investigative tool that could help her efficiently and intelligently search through the space of network and parameter possibilities. Direct experimental engagement with her systems also served to reduce the risk of error. The whole problem-solving system served to augment her cognitive ability to investigate the complex biological systems.

Acknowledgements

We appreciate the support of the US National Science Foundation in conducting this research (DRL097394084). We thank the director of Lab C and, especially, C9 for welcoming us into the lab and granting us numerous interviews. We thank the members of our research group for contributing valuable insights, especially Vrishali Subramanhian, Lisa Osbeck, Sanjay Chandrasekharan, and Wendy Newstetter. We appreciate the helpful comments of two anonymous reviewers and the editor of the journal.

References

Bechtel, W., & Richardson, R. C. (2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*: MIT Press.

Bruggeman, F. J., & Westerhoff, H. V. (2007). The Nature of Systems Biology. *TRENDS in Microbiology, 15*(1), 45-50.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin, 56*(2), 81-105.

Carusi, A., Burrage, K., & Rodríguez, B. (2012). Bridging experiments, models and simulations: an integrative approach to validation in computational cardiac electrophysiology. *American Journal of Physiology-Heart and Circulatory Physiology, 303*(2), H144-H155.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks: Sage Publications Inc.

Denzin, N. K. (2009). *The research act: A theoretical introduction to sociological methods*: Aldine De Gruyter.

Fagan, M. B. (2012). The joint account of mechanistic explanation. *Philosophy of Science, 79*(4), 448-472.

Giere, R. N. (1992). *Cognitive Models of Science. Minnesota Studies in the Philosophy of Science 15*. Minneapolis: University of Minnesota Press.

Hall, R., Wieckert, K., & Wright, K. (2010). How does cognition get distributed? Case studies of making concepts general in technical and scientific work. In M. Banich & D. Caccmise (Eds.), *Generalization of Knowledge: Multidisciplinary perspectives*. NY: Psychology Press.

Hollan, J., Hutchins, E., & Kirsch, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction, 7*(2), 174-196.

Howitt, D., & Cramer, D. (2008). *Introduction to research methods in psychology* (2nd ed.): Pearson Education.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Ideker, T., Galitski, T., & Hood, L. (2001). A new approach to decoding life: systems biology. *Annual review of genomics and human genetics, 2*(1), 343-372.

Kitano, H. (2002). Looking Beyond the Details: a Rise in System-Oriented Approaches in Genetics and Molecular Biology. *Current genetics, 41*(1), 1-10.

Krohs, U., & Callebaut, W. (2007). Data without Models merging with Models without Data. *Systems Biology: Philosophical Foundations (Boogerd FC, Bruggeman FJ, Hofmeyer J-HS, Westerhoff HV, eds)*, 181-213.

Lenhard, J. (2006). Surprised by a nanowire: Simulation, control, and understanding. *Philosophy of Science, 73*(5), 605-616.
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks: Sage Publications Inc.
Morrison, M. (2009). Models, measurement and computer simulation: the changing face of experimentation. *Philosophical Studies, 143*(1), 33-57.
Nersessian, N. J. (1995). Opening the black box: cognitive science and history of science. *Osiris, 10*, 194-211.
Nersessian, N. J. (2002). The cognitive basis of model-based reasoning in science. *The cognitive basis of science*, 133-153.
Nersessian, N. J. (2006) Boundary objects, trading zones, and adaptive spaces: How to create interdisciplinary emergence? Address to US National Science Foundation Sciences of Learning Centers PI meeting, 10/2006
Nersessian, N. J. (2008). *Creating scientific concepts*: The MIT Press.
Nersessian, N. J. (2009). How do engineering scientists think? Model-based simulation in biomedical engineering research laboratories. *Topics in Cognitive Science, 1*, 730-757.
Nersessian, N. J. (2012). Engineering Concepts: The Interplay between Concept Formation and Modeling Practices in Bioengineering Sciences. *Mind, Culture, and Activity, 19*(3), 222-239.
Nersessian, N. J., Kurz-Milcke, E., Newstetter, W. C., & Davies, J. (2003). Research laboratories as evolving distributed cognitive systems. *Proceedings of the 25th annual conference of the cognitive science society*, 857-862.
Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 14): Prentice-Hall Englewood Cliffs, NJ.
Noble, D. (2006). *The music of life: biology beyond the genome*: OUP Oxford.
O'Malley, M. A., & Dupré, J. (2005). Fundamental Issues in Systems Biology. *BioEssays, 27*(12), 1270-1276.
O'Malley, M. A., & Soyer, O. S. (2012). The roles of integration in molecular systems biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43*(1), 58-68.
Parker, W. S. (2009). Does matter really matter? Computer simulations, experiments, and materiality. *Synthese, 169*(3), 483-496.
Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods*. Thousand Oaks: SAGE Publications.
Richardson, R. C., & Stephan, A. (2007). Mechanism and mechanical explanation in systems biology. *Boogerd et al*, 123-144.
Simon, H. A. (1979). *Models of thought* (Vol. 1): Yale University Press.
Simon, H. A. (1983). *Reason in human affairs*: Stanford University Press.
Voit, E. O. (2013). Mesoscopic modeling as a starting point for computational analyses of cystic fibrosis as a systemic disease. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*.
Voit, E. O., Qi, Z., & Kikuchi, S. (2012). Mesoscopic Models of Neurotransmission as Intermediates between Disease Simulators and Tools for Discovering Design Principles. *Pharmacopsychiatry, 45*(1), 22.
Westerhoff, H. V., & Kell, D. B. (2007). The Methodologies of Systems Biology. In F. C. Boogerd, F. J. Bruggeman, J.-H. S. Hofmeyr & H. V. Westerhoff (Eds.), *Systems Biology: Philosophical Foundations* (pp. 23-70). Amsterdam: Elsevier.
Westerhoff, H. V., Winder, C., Messiha, H., Simeonidis, E., Adamczyk, M., Verma, M., . . . Dunn, W. (2009). Systems biology: the elements and principles of life. *FEBS letters, 583*(24), 3882-3890.
Wimsatt, W. C. (1975). Complexity and organization *Topics in the philosophy of biology* (pp. 174-193): Springer.
Wimsatt, W. C. (1987). False models as means to truer theories. *Neutral models in biology*, 23-55.

Wimsatt, W. C. (2007a). On building reliable pictures with unreliable data: An evolutionary and developmental coda for the new systems biology. *Systems biology: Philosophical foundations*, 103-120.
Wimsatt, W. C. (2007b). *Re-engineering philosophy for limited beings: Piecewise approximations to reality*: Harvard University Press.
Winsberg, E. (1999). Sanctioning models: The epistemology of simulation. *Science in Context, 12*(2), 275-292.
Winsberg, E. (2003). Simulated experiments: Methodology for a virtual world. *Philosophy of Science, 70*(1), 105-125.
Winsberg, E. (2009). A tale of two methods. *Synthese, 169*(3), 575-592.
Winsberg, E. (2010). *Science in the Age of Computer Simulation*: University of Chicago Press.