Journal of Experimental Zoology, Part B: Molecular and Developmental Evolution, 322:230-239, 2014

Strategies for Coordinating Experimentation and Modeling in Integrative Systems Biology

Miles MacLeod

TINT Centre of Excellence in the Philosophy of the Social Sciences, University of Helsinki

Unioninkatu 37 PL 54

Helsinki 00014

Finland

miles.macleod@helsinki.fi

Nancy J. Nersessian

School of Interactive Computing, Georgia Institute of Technology

85 Fifth St NE

Atlanta 30308 GA

USA

nancyn@cc.gatech.edu

Figures: 0

Charts: 0

Tables: 0

Abbreviated Title: Research Strategies in Integrative Systems Biology

Correspondence to:
Miles MacLeod,
Department of Political and Economic Studies, University of Helsinki
P.O. BOX 24
Helsinki 00014
Finland
miles.macleod@helsinki.fi

Supporting Grant: US National Science Foundation (DRL097394084)

Strategies for Coordinating Experimentation and Modeling in Integrative Systems Biology

Abstract:

In this paper we provide a novel analysis of the affordances and trade-offs of different strategies for integrating model-building and experimentation in integrative systems biology. Unimodal strategies rely on collaboration between experimenters and modelers in distinct laboratories. In a bimodal strategy modelers perform their own experiments. Each option has advantages and challenges. In the case of the labs we studied, the choice of strategy often depends on preferences held by lab directors as to the strategy that best achieves certain philosophical objectives concerning what they see as the aims of modeling in systems biology and the epistemic standards to apply to it. We identify an important connection between philosophical divisions in systems biology and the structure of research in systems biology. Better knowledge of these strategies and their philosophical motivations provide insight behind the diversifying structure of the field, and can help lab directors understand the challenges their researchers face, as well as the training and lab organization options available.

Strategies for Coordinating Experimentation and Modeling in Integrative Systems Biology

One of the classic problems that science repeatedly has to face is how to manage and integrate experimentation with theory or model development efficiently and effectively. Physics for instance by-and large has settled on relatively separate divisions of labor, acknowledging the specialist skills required to do either (Galison, '97). This question is becoming an increasingly important one for integrative systems biology (ISB). ISB attempts to model large-scale biochemical and cellular systems using modern computational and mathematical techniques. It positions itself independent of molecular biology, but is nonetheless dependent on that field for biological expertise and experimentation to validate those models. (Voit, '00, Kitano, '02, Bruggeman and Westerhoff, '07, Voit, '13) Indeed ISB is far from settled on its relations to molecular biology, or on what are the best modes of research organization to further its modeling practices (Calvert, '10, Calvert and Fujimura, '11, MacLeod and Nersessian, '13, Nersessian and Newstetter, '14). The complexity of problem-solving tasks ISB takes on and their dependence on the specifics of the biological case make it difficult for the field to find best strategies for field development. In our research on ISB we have discovered that there is at present no consensus among modelers on whether or not to construct computational laboratories around *unimodal* modelers, who build models in collaboration with external experimenters, or to construct wet-labs with *bimodal* modelers, who can run their own experiments in the process of building models, or other options such as structuring labs around some combination of experimenters and unimodal or bimodal modelers. However, in the current state of affairs in ISB the unimodal modeler, who possesses little or no experimental training and performs no experiments him/herself, dominates the field. Our ethnographic research has thus far focused on unimodal modelers from dedicated systems

biological labs who collaborate with experimenters located in molecular biological labs and on bimodal modelers who conduct their own experiments in the course of building models.

We present here briefly some of the results of our four year investigation into the research practices in two ISB labs, one with a unimodal structure and external experimental collaborations (Lab G) and another which has a fully equipped wet-lab for conducting bimodal modeling research (Lab C). We do not examine and compare all the lab-organization strategies that are emerging in the field, because our focus is not so much on lab organization as it is on how these two basic *types of modeling strategies* perform in these lab environments.

Our specific aim here is to elucidate these modeling strategies and compare the research benefits that characterize each of them, and the practical trade-offs of favoring one over the other. Doing so will help generate awareness of these choices and their trade-offs amongst lab directors. Importantly we show in our two labs that these choices are in fact guided by different, deeper philosophical preferences of our lab directors about the aims and independence of systems biology. Our findings dovetail with what other commentators have observed about philosophical divisions in systems biology (O'Malley and Dupré, '05), but demonstrates further how deep this issue goes to level of lab organization and training, and hence the consequences these philosophical preferences have for the shape and structure of systems biological research. Although we have done an in-depth study of only two labs, as with all ethnographic research, we do not generalize our findings but seek to abstract significant insights from our cases as candidate findings that might transfer to other situations. We hope in this way to generate more discussion of research strategies in systems biology, and their relationships to the aims and agendas of systems biology, both inside ISB itself as it continues its process of development, and also in other systems biological fields that face similar challenges.

1. Unimodal and Bimodal Research Practices

1.1 Our research

We have been conducting a *cognitive ethnography* (Hutchins '95) of the research practices of two ISB research labs[1]. This kind of research modifies the techniques of classical anthropological field research and qualitative data analysis to fit the context of studying scientists in "in the wild," i.e., in the "native" environments of the research lab, with a focus on their knowledge-producing practices. We have been collecting qualitative data, comprising observations of bioengineering scientists at work in the lab and in research meetings ("field observations"), multiple interviews with them about their research practices, and historical artifacts such as grant proposals, paper drafts, dissertation proposals, powerpoint presentations, and posters. Our goal is to develop a detailed understanding of the current state of the problem-solving practices in ISB, with a major objective being to provide insight into how best to facilitate research and learning in this emergent transdisciplinary field.

While we recognize there are several computational approaches to the study of biological systems, our investigation has focused on a major area of research in ISB directed towards representing, understanding, and controlling intracellular metabolic and signaling pathways. Our study thus far comprises two biochemical systems biology laboratories located in a major university research institution. The researchers are primarily graduate students working on PhDs in engineering fields, though each lab often hires one or two postdoctoral researchers. Lab C conducts both modeling and experimentation and is working on establishing the role of biomolecular processes and element found in cellular environments in

[1] This research is funded by the US National Science Foundation grant DRL097394084. Human subjects restrictions require that we do not identity specific researchers, who are mainly PhD students, and so we refer to the labs with letter designations and lab members with letter plus number.

the control and regulation of cell signaling dynamics. This issue is a major theoretical concern for both molecular biology and physiology. Lab G does modeling and novel algorithm development but no experimental work itself. The specific modeling problems it works on come from collaborations with independent experimental labs and cover a range of different problems such as research on biofuels, Parkinson's disease, atherosclerosis, and heat shock in yeast. Both labs build for the most part Ordinary Differential Equation (ODE) models of biological pathways.

Integrative systems biology is a new field, being approximately 20 years old (Kitano, '02), although the idea of approaching biology by means of systems theory itself has an older history (O'Malley and Dupré, '05). ISB aims specifically to develop large scale mathematical models through the aid of contemporary computer power and algorithmic techniques to map the structure and dynamics of complex biological systems. The field is often characterized as having various streams; bottom-up and top-down (Bruggeman and Westerhoff, '07). "Top-down" systems biology derives from 'Omics methodology, using powerful statistical techniques to ''reverse engineer'' system structure. According to Bruggeman and Westerhof, "Bottom-up systems biology starts from the bottom (the constitutive parts) by formulating the interactive behavior (rate equation) of each component process (e.g. enzymatic process) of a manageable part of the system." (47). Our study participants broadly pursue "bottom-up" systems biology, although it should be noted these notions are rather in flux and lab G uses techniques that might be called top-down methods to fix parameters. Nonetheless the top-down 'Omics-led systems biology can be distinguished from what lab G and lab C do which relies upon building models as much as possible from the experimental data of component processes. In this regard bottom-up systems biology has a

much different relationship to established experimental biology than top-down systems biology.

Our lab researchers take the distinguishing feature of their bottom-up approach to be that the models allow them to describe how elements (in their case enzymes and metabolites) operate and interact within a systemic context or environment, and thus express particular system-determined properties and behaviors. They contrast this approach with traditional molecular biology that pursues these elements ''in isolation'' through *in vitro* analyses of the properties and structure of individual biomolecules. Of particular interest, the modelers in our labs are not in fact biologists, coming mostly from engineering fields, and many confess (particularly in Lab G) to having had few biological courses and no experimental training. This contrasts to some extent with Lab C, where at least some modelers have exposure to it as biomedical engineers for instance. In both labs, the primary skills and knowledge of the modelers are in computational model building (though often not of biological systems when they enter), combined with general knowledge of the properties and features of engineering systems (control theory, signal processing, etc.). Several of the Lab C researchers, including the director, conduct experimental research targeted specifically to obtain data for their models. When we started interviewing there was a mixture of bimodal and unimodal modelers, and only one pure experimentalist, but now Lab C is almost entirely bimodal. Lab G on the other hand is entirely unimodal. Researchers need to rely on experimentalists outside of their lab through collaborative relationships for experimental data and biological advice.

In both labs model-building is a complex process that involves constructing a pathway representation of a given biochemical network, choosing a mathematical representation for that network (from a range of possibilities) and fixing unknown parameters either through

experimentation or an algorithmic method. Most modelers aim for ordinary differential equation (ODE) models of network dynamics. Our investigations establish just how complex a process this can be, made difficult by various constraints, including particular constraints on data available and computational power, and the sheer complexity of the biochemical networks, which often generates large alternative structure and parameter spaces for researchers to have to search through and narrow (MacLeod and Nersessian, '13a, MacLeod and Nersessian, '13b). There is no ready-to-hand best strategy or methodology that can handle the diverse nature of data situations researchers find themselves in for particular systems, which means for the most part modelers bear a high cognitive load trying to develop their own strategies for their particular systems.

Our analysis of these two labs has been developed and triangulated from several data sources. We employ the standard ethnographic data collection methods of participant observation, informant interviewing, and artifact collection. In Lab C we observed researchers as they conducted their work on the bench tops and as they wrestled with constructing pathways; we attended lab meetings, which were audio-taped to complement the field notes; we sat in on the weekly journal club and PhD proposals and defenses. We determined after preliminary field observation that it would not be possible to collect this kind of data for Lab G since researchers work only on computers and often do their work from home. They do not have lab meetings or a journal club. In the first year of our study, the Lab G director did organize several lab meetings specifically to introduce us to their research and two meetings joint with Lab C. In both labs we conducted unstructured (open) interviews with the lab members and some collaborators outside of the lab. We collected relevant artifacts including powerpoint presentations, paper drafts, published papers, grant proposals, dissertation proposals, and completed dissertations.

1.2 The Unimodal Modeler – Lab G

Lab G follows a unimodal mode of modeling practice and lab organization. Researchers arrive largely without biological training. They are engineers favored principally by the director for their mathematical and problem-solving abilities, given the mathematically intensive nature of the lab's model-building; particularly parameter-fixing and model analysis. The problems they work on take two forms. The first are strictly mathematical tasks provided by the director, such as developing new parameter-fixing algorithms for a given class of data. The second come usually from molecular biologists who approach the lab with data they have collected, looking to have models constructed that will hopefully help them develop insight into their systems and direct their own experimental work fruitfully. Ideally our lab G graduate researchers should work back and forth with their collaborators while constructing the desired models, although it rarely works out this way. Modelers must themselves sift through often large amount of literature on the networks to put models together. In most cases the data provided by the experimenters or available in the literature are insufficient alone for building dynamic ODE models of the networks. The researchers spend considerable time using – and developing – mathematical techniques to build and validate models.

1.3 The Bimodal Modeler – Lab C

Lab C operates quite differently. In the first place, while lab G has no specific theoretical agenda beyond algorithm development and promoting modeling, in Lab C there is a distinct agenda connected to general questions concerning metabolic and cell signaling process, and the researchers work on a variety of projects connected with this agenda. Most participants are bimodal modelers, both building models and doing their own experimentations as part of the model-building process. The lab has a fully equipped wet-lab. Bimodal researchers

integrate experimentation into their modeling in sophisticated ways, beyond just validating models (see section 2.2). However, these researchers often do rely on a degree of collaboration or biological expertise with bioscientists. The lab C PI, who is trained in engineering and in molecular biology, is a central resource of biological advice for the particular systems and processes her researchers are dealing with. Further outside biologists who are experts on particular processes are often sought out for advice, although not for experimentation themselves. Dissertation committees usually include bioscientists or medical researchers.

2. Central Challenges of Bimodal and Unimodal Modeling Strategies

There are numerous challenges the choice of either strategy present both for research and for training researchers in ISB. Given the brevity of this article, we focus on the challenges of collaboration that are faced primarily by the unimodal strategy, but which can also impact bimodal collaborations. These would seem to indicate that the bimodal strategy is the most effective, but in that case the challenges of training researchers come to the fore. We examine some trade-off regarding each strategy and offer suggestions for moving beyond the current situation including alternative lab structures.

2.1 Challenges of Collaboration in Systems Biology

As noted, the unimodal strategy relies significantly on modeler – experimenter collaborations. As we have witnessed in both research labs, developing a collaboration is fraught with difficulties. Our studies have identified several principal reasons for this, which result in many challenges for those trying to build models this way. In the first place molecular biologists have their own disciplinary aims and agendas which do not necessarily align with the requirements of modeling. Indeed, as mentioned, experimenters usually provide data with

the expectation of getting something that services specifically their interests. Further, they usually lack understanding of the kind of data needed for successful model-building. This means that it is usually difficult for a modeler to get the actual data they need from an experimenter. G10 for instance, who was modeling atherosclerotic inflammation, encountered this problem.

"the biologist(s) produce the data they want. But those data are not actually what we want when we do parameter estimation….. so there….might be some gap between these two, between us…they only focus on one species. But even so they don't produce enough data. ….they don't measure the concentration for example. And they have few kinetic data…"

Models usually have specific parameter requirements for the specific models they wish to build (ODE models need kinetic concentration and rate data for instance), the measuring of which can be difficult, expensive, and time consuming. Modelers are often not aware of these difficulties. As G10's one-time experimental collaborator put it:

"the data that they [modelers] want from us is something that is not simple to generate. So if they want a Km for an enzyme we have to purify the enzyme . Then we have to create all the conditions to measure it in vitro. That's not a simple undertaking. That's probably six months of work. And none of us have a person sitting around who can do that for six months. So that's the first problem. The second problem is yeah if we are going to do that. If we are going to spend six months generating what they want then we would like we need to have something that is going to come out of it."

Another piece of the problem here too was that G10 did not particularly understand that as a vascular biologist, producing this kind of data was not something her collaborator ordinarily did and was not of value for her research group.

On the other side biologists, particularly those that have been approached for collaboration, often do not see good reasons to trust that cooperating with a modeling process will produce results useful for them. There is something of a catch-22 here for modelers. Modelers are usually forced to rely on limited data sets to build models, which experimenters, having intrinsic knowledge of the unreliability of data, look upon unfavorably, which discourages their cooperation. In G1's words,

"The experimentalists are extremely, extremely skeptical about our model extremely skeptical …as skeptical as they are about somebody else's data [laughs]. So for instance I spoke to my collaborator last week and looking at one dataset. She says if you will work on that piece of data I will never buy a model from you. You have to look at larger datasets. I have to otherwise she will not buy my theories."

Yet experimenters could help the models become more accurate by supplying new data sets – or even by sharing their existing unpublished data excel spread sheets, which they are often reluctant to do.

On the other hand modelers feel that experimenters simply do not understand the capabilities of models, and the power of mathematical techniques to derive network structure and generate predictions from limited data through mathematical methods of simplification and approximation. Nor do they automatically accept the argument from modelers for the necessity of models for capturing the control structure of systems as imperative for their effective manipulation and control. Modelers assert that the connectivity, nonlinear dependency, and large-scale of many biological systems cannot be computed independently of mathematical models combined with computational simulation. Instead experimenters often interpret models as just reproductions of old data – or as "tautologies" that can offer no new insight. As C7 saw it:

"They think of [models] as something that's… just hooked up to – to, you know, match figures…. I had somebody who was… data fitting, so , for them, it's just like, you're using your data and then, you know, you're plugging in some numbers to fit the output of your model to that. And then they would not pose a lot of faith in those models or what they predict."

Hence for modelers, according to G2:

"I think you need biologists who…..either understand or see the benefit in doing some computation – and they shouldn't need to understand everything we are doing basically but they need to understand how we interpret results and things like that."

Many of the researchers in Lab G attempting to model in collaboration have struggled with these relationships. Often it has resulted in a loss of interest or lack of timely response on the part of the experimenter, leaving the modeler to simply do the best he or she can on their own. Of course it is understandable why these problems exist. Modelers and experimenters come from widely different backgrounds. It is very difficult for modelers to have any insight without training into the constraints of experimental practice, and almost impossible for experimenters to really understand why the modelers need what they do without having had some hands-on modeling experience themselves. The collaborative research strategy is thus set-up for considerable difficulty at the outset. However bottom-up systems biology is dependent on molecular biology and thus dependent on its acceptance and appreciation by molecular biology as something experimentalists should invest their time in. Yet, given the difference in time scale at which wet-lab experimentation and simulation operate, the divisions of labor here are by no means favorable to the bottom-up modeler. As the Lab G director has framed it, "*and I still maintain, I've said it for 20 years, you need 10 experimentalists for every modeler*" in order to avoid the consequences of not having timely

experimental results in which modelers are often stuck waiting months for critical data before they can progress.

The best instances of collaboration we have observed are precisely those in which the experimentalists have a background in modeling, even if only a limited one, and have some understanding of the needs of large-scale systems modeling. But these are still comparatively rare situations.

2.2 Advantages of the Bimodal Strategy

It is obvious that the bimodal strategy solves many of the difficulties of collaboration. For the most part bimodal researchers can control their experimentation to suit their needs. Unlike unimodal modelers, bimodal modelers can relatively easily adapt their model- building strategies and choices to fit what is possible for them to carry out experimentally, which gives them a greater ability to direct their model-building in more reliable and productive directions. One bimodal modeler we followed, C9, expressed it this way.

"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."

Bimodal researchers are more directly in control of their experimental protocols. This is often a significant issue for unimodal modelers who are often forced to rely on data generated under circumstances that do not quite fit what they are ostensibly modeling. As we saw above this is a serious issue given that data quality and relevance is one of chief reasons molecular biologists reject large-scale models.

Further, bimodal researchers are more directly aware of the reliability of the data they obtain. We observed one particular case in lab C in which collaboration between C7 (at that time,

unimodal) and the one pure experimentalist in the lab (C11) broke down over this issue. C7 initially planned to do no experimentation himself. C11, who had no mathematical training, had been tasked to get data needed C7's modeling, but the result turned out badly, with C7 complaining about the quality of the data, and C11 complaining that she did not understand what C7 wanted. As a result C7 decided to do his own experimentation, noting "*I feel like I have more confidence in the data if I'm doing it on my own rather than having someone else do it.*"

Time-lags are also of course no longer such an issue as C9 stressed:

"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."

This quote indicates a measure of coordination between experiment and model building. The systematic and organized manner in which C9, for instance, *coupled* model-building with experimentation had significant advantages over collaboration (MacLeod and Nersessian, '13b). For example, wet-lab experimentation allowed her to validate model elements, which were used by her as a cognitive platform for interpreting and understanding the dynamics of the networks she was studying through running repeated simulations and perturbing those simulations. When 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. Thus the limitations of the models proved to be 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. The coupling of experimentation and simulation provided continual epistemic validation for the model and reinforcement of its role as a platform for both making hypotheses and for building understanding of network dynamics (Nersessian '08).

Validation is not the only advantage of coupling experiment and simulation we observed with C9 (MacLeod and Nersessian '13). Specifically, she developed sophisticated strategies to localize inaccuracies in her model, and thus localize quite precisely where experiment was required. She would do this by running sets of controlled "simulation" experiments (fixing particular sets of parameters and varying others) against controlled biological experiments that replicated the same controls physically. In this way she could work through the model to pick out the particular problematic relationships and investigate them further. This fine-tuned operation gave her a particular ability to handle the complexity of her problem-solving task, not available to unimodal modelers. Unimodal modelers are usually forced to explore complex large parameter and structure spaces algorithmically or with Monte Carlo methods to discover approximately accurate structure and parameter values that fit the whole model to the data, while C9 could cut down this space significantly with well-targeted experimentation on just particular relationships. Once errors were localized to particular structural hypotheses or parameter values, experiments could be done to run through sets of hypotheses about the interactions of those components, and measure their parameter values. When it came to parameter-fixing C9 declared that she had always at most 2 to 3 parameters to fix. This compares with the upwards of 30 that most lab G members report.

2.3 Trade-offs and underlying philosophical motivations

The apparent advantages listed above seem to suggest that bimodal modeling might be the ideal or best practice for systems biology. Yet unimodal modeling is the dominant mode of modeling in current systems biology, whether it takes place within mixed or unimodal laboratories. There are in fact specific trade-offs to bimodality, as well as unimodality, some of which that depend on the philosophical objectives that lab PIs have for systems biology.

In first place G4 was keen to stress that having the modeler and experimenter as separate individuals, particularly at some arm's length (in different labs), reduced the risk of experimental bias, or the manipulation or interpretation of results to fit the model: "*if you produce the data to validate your model, implicitly or explicitly… there is a lot of room for interpretation*."

On the other hand one potential advantage we have witnessed is that bimodal modelers have is a greater capability for building models closer to the mechanistic details. Having experimental access means they can apply lower level representations of interactions more easily, such as Michaelis-Menten representations of enzymatic interactions. These are highly favored by molecular biologists, rightly or wrongly, as good mechanistic accounts.[2] The parameters can be potentially fixed through experimental investigation by a bimodal researcher. As we saw in G10's case collaborators are not always willing to find such parameters for modelers. Thus, unimodal modelers usually need to rely on canonical templates of systems and interactions, (e.g. power laws) that represent a wide variety of behaviors. These templates are fixed from above, through parameter-fixing processes, not directly (or from below) through experimental investigation of a system. They are however, for the most part, less familiar to and not well-understood by biologists. So for providing rich mechanistic details bimodal modeling does have an advantage. Further, bimodal researchers tend to have greater biological skills (by experience and background) and greater biological understanding of their systems. Thus bioscientists, as with the ones we interviewed, might

[2] Modelers often complain about this in fact, since molecular biologists often interpret model representations as abstractions and Michaelis-Menten type representations as mechanistic reality. In fact as modelers point out Michaelis-Menten is itself an abstraction based on various simplifications and assumptions applied to a mass-action representation. In C7's words, *"For example, there's a very famous equation that's just called the Michaelis-Menten equation. It's supposed to represent [enzyme] kinetics. But that is an approximation. Most biologists, you know, do not realize that. And they have it in their subconscious that, well, that this is a precise representation of the exact kinetics and if I were to use a more basic representation of which Michaelis-Menten is an approximation, they would not trust that."*

find model results produced through this kind of research work generally more convincing, since it fits more closely with their own epistemic standards and requirements.

But the singular trade-off of bimodal modeling in lab C according to the lab G PI is model sophistication. As G4 put it,

"If you do the experiment yourself, you know what the data are like; you know how reliable they are. You know the kind of assumptions you made in order to produce the data…. So you get better idea about the whole context. On the other hand, life is complicated, and to do good modeling is a full time job; to do experiment is a full time job. And if you don't want to do two full time job, then something will suffer from it…"

As a result ,

"….. the models that are being developed in C4's lab are…by and large, off-the-shelf type modeling approaches that are, not always, but that are often rather simplistic, I'd say."

Here G4 was citing the lack of mathematical sophistication of modeling techniques or analysis in lab C and the use of rather common *"superficial"* mechanistic models of interactions such as Michaelis-Menten models. This criticism marked what G4 referred to as a "philosophical divide" between lab C and lab G that may well reflect deeper attitudes in the field. As O'Malley and Dupré (2005) have argued there are, in organization and in research strategy, divisions in systems biology over the value and aims of systems biology, even though they might not be debated openly. Some systems biologists are relatively pragmatic, aiming more immediately at using modeling as a tool for the extension of molecular biology by largely pursuing its theoretical agenda; others have a strong theoretical agenda of their own promoting the development of a theory of biological systems and advancing the role of mathematics in systems analysis. In the case of lab C and lab G these divisions are explicitly acknowledged

and are expressed in the form of laboratory organization their respective directors chosen, and the mode of modeling they favor.

As mentioned earlier lab C is driven by a particular theoretical agenda: establishing that particular biochemical systems play a key part in regulation of cell signaling and metabolic processes. The issue is a current one for molecular biology and physiology, and C4 has set herself the task of demonstrating it to these communities and with the help of them. As she stated:

"when I was trying to think about what strategy I would take a faculty position I'd been looking at originally metabolic pathways and these signaling systems… and I kinda wanted to bring them together and the reason I wanted to was I was really surprised in my PhD work I had very little exposure to signaling. So they are almost like two different camps in cell biology…. enzyme-based all people cared about was how it got into the cell all the metabolites that are involved and didn't really think all that much what was controlling the expression of these proteins or how the muscle was responding to any other cues from its environment or anything….and all these signaling people interested in what's connected to what and how this receptor is you know causing changes in gene expression but they don't think about uh basic things like the energy supply in the cell…. So there was this disconnect between the two camps so I saw my training as a way of kind of integrating those things together."

In this sense her goal is pragmatically focused on using models to assist biological investigation, which is how she herself framed her activity as assisting molecular biology by specializing in a network-level accounts. As she reported explaining to one of her lab members, "*C2 [postdoc] comes from a more traditionalist biology background. And so he gets really into the nitty gritty, he gets really excited about these 2 proteins and wants to investigate them to the nth degree…I tell him that there a lot of labs out there that can do this much easier*

than us and that's not what we really specialize in - what can we do that no one else can do - and put it in terms of the context with respect to the rest of the network."

This pragmatic focus may well explain the fact that the modeling approaches Lab C researchers undertake demonstrate C4's clear preferences for tractability and experimental accessibility in the form of smaller scale more mechanistic models that help demonstrate how particular mechanisms operate. The preference for Michaelis-Menten or simple mass action type models of interactions engages with common representative techniques of molecular biologists themselves who commonly use Michaelis-Menten as a fairly reliable easily available technique of representation for fulfilling these goals. The experimental skills of the modelers served to keep the unknown parameters of these models relatively under control, and it was certainly the case that lab C members generally had far fewer parameters to fix than lab G members.

While the attitude permeating down from the lab C PI to lab members was more pragmatic with respect to the use of models, the lab G PI was quite emphatic about the importance of advancing and enriching a sophisticated mathematical theory and mathematical analysis that could account for complex systems*. "I claim that our [rather than lab C's] models at least are going in much more depth".* For this complex parameter-fixing and mathematical work was unavoidable. Non-superficial sophisticated mathematical accounts capable of rigorous mathematical analysis require novel (not "off the shelf") mathematical analyses that can effectively capture a wide range of nonlinear behavior within tractable formalism that keeps complexity under control. Producing these and the tools for mathematical analysis however required high-levels of mathematical skill that in G4's view cannot be achieved in combination with experiment. G4 acknowledged there might well be a trade-off with the kinds of models in lab C but was skeptical that it really produced simpler outcomes *"….because they are simplistic,*

you have fewer parameters, maybe, and they think they are easier to estimate…I'm not ever sure that is true." As such the kind of systems biological projects C4 and G4 wished to advance reflected their perceptions of the aims of modeling and in turn their respective preferences for the mode of modeling and lab organization that best suited it.

2.4 The Unimodal Lab Going Forward

Unimodal lab organization in lab G encourages investment working out how to structure and control its dependency on experimental data without experimental expertise. Much method development in ISB could be construed as intended to work around the data constraint problem and thus to instantiate the field's independence of having to rely on experimentation as much as possible. At least 50% of Lab G time is in fact concerned with structure identification, data smoothing or parameter-fixing techniques rather than actual pathway or system modeling. Such techniques, like the widely used flux balance analysis, combine biological optimization constraints and conditions with parameter or structure data to produce better models.

Another solution however which has yet to really garner as much attention is to work on providing modelers and experimenters with better "interactional expertise" so that collaboration can work more efficiently (Collins and Evans, '02). Interactional expertise is the expertise that derives from knowing enough about a field's methodology and practices to be able to engage with it productively "linguistically", without necessarily being able to operate within the field. This requires of modelers, for instance, that they have some knowledge of experimental practices and the epistemic aims of experimenters. We have witnessed that small interventions can have big pay-offs for developing interactive expertise. Two lab G members, for example, were sent to an experimental lab for a summer month to experience

first-hand experimental work and the results were encouraging. As G16, one of this pair, stated:

"you are looking from far away, you just see that this person is just going into lab and just pipetting and that's not interesting and why would you do that? But then when you get it, you see that there are a lot of reasonings [sic] going on and they are involved with their own sort of culture, I guess." (138)

Leveraging from this experience she was able to develop a more complex understanding of what she as a modeler would need to know to collaborate:

"and then we need to be able to communicate, we need to have an idea of what type of experiments are done…..their area of research is very limited. They just know some sorts of experiments that they do in their lab with the equipment they have. And then you sometimes need to include someone else in the project to do some other part for us to be able to build a dynamic network of this specific pathway in this specific organism, we need this kind of data. If we don't have it, we can't. And then, I'm a modeler, you're a biologist, you don't do that type of experiment. Who do you think could do that? And then, how much do you think it will cost? I can ask a question from you, but I need to have an idea that such a thing exists to be able to, like, think about or suggest it and all."

G16 displayed after these experiences the kind of meta-level cognition central to interactional expertise, namely an ability to cognize the situation from an experimenters point of view, and thus make better judgments about what could be asked from them, increasing her own potential efficiency and effectiveness in collaborative partnerships.

On the other side of the interaction, i.e. that of the collaborating bioscientist, we worked with ISB faculty to develop an introductory biosystems modeling course for both bioscientists and

novice modelers. Over the course of the semester we discovered that even a little knowledge of model-building goes a long way. One student was a biologist who had tried unsuccessfully to collaborate with a modeler. We interviewed her during and after the course and she reported having a much better idea of what the modeler had wanted from her all along:

"I wish I had taken this class 2 years ago. I wish [name: modeler] and I had taken it together. We would have looked at each other and said 'Oh, I get it – I know what you are doing now.' It would have been very helpful for me to understand what kind of data he needed; to understand what kinds of questions he should have been asking of me."

The last statement is especially interesting from the perspective of interactional expertise because she felt she now understood and could elicit what he "should have been asking" of her. As she progressed in the formulation of her own research project, she declared that even with this –and a subsequent course – she was in no position to build systems models, but now understood the value of modeling for biology. Through this experience she had develop sufficient interactional expertise, which she recognized as essential her ability to collaborate with modelers.

2.5 Other Possible Lab Organizations

We have presented several of our significant findings from two laboratories which represent specific basic possibilities of lab organization. The field however is exploring more mixed options. G4 himself cited the Lauffenberger lab at MIT, with a large-scale mix of unimodal modelers and experimenters, as a possible lab organization with the potential to solve some of the undesirable trade-offs of his approach, although one certainly outside his funding resources. Funding constraints, of course, play a role in lab organization, but so do "philosophical" stances, such as G4's choice for a unimodal lab is an accommodation to his aim

of developing sophisticated mathematical models. C4, a new assistant professor, undertook starting up a wet-lab despite having quite limited funding because she believed it important for modelers to be able to conduct their own experiments. As another possibility, some of the collaborating biologists we talked to favored the idea of situating unimodal modelers within their labs (but had not acted on it). While they would likely not be doing their own experiments, these modelers would interact directly with experimenters in an experimental environment and hopefully develop a better understanding of experiment's affordances and constraints as a result – and vice versa for the experimentalists and modeling. Given the hurdles of the complex mathematics involved in systems biological modeling both the lab G director and some of his experimental collaborators we interviewed thought it more reasonable to expect modelers to be able to adapt better to experimental contexts than vice-versa. C4 herself trained first as a modeler and expressed her own preference in the choice of engineering trained students rather than biologically trained ones. Although not in the labs we studied, such strategies of education and placement are recognized as potentially good solutions for generating biological or modeling awareness and interactional expertise and are being explored elsewhere in the field.

3. Conclusions

While our research has focused specifically on integrative systems biology these questions about modeling strategy are just as pressing across biology wherever new highly technical computational, mathematical and engineering resources are being brought to bear. This is also the case in evolutionary biology, which is already grappling with how to integrate a diversity of distinct approaches and perspectives, such as the experimental approach of evo-devo and the theoretical model based approach of population genetics. (Wagner, '07, Brigandt and Love, '12) In fact the relatively new field of evolutionary systems biology is playing just

such a potentially integrative role (O'Malley and Soyer, '12, Soyer and O'Malley, '13). It will face the task of evaluating the cost and benefits of the unimodal and bimodal modeling strategies we've identified above while also factoring in the added expertise required for evolutionary analyses and explanation at often multiple levels of organization, such as expertise for the construction of genotype-phenotype maps and fitness landscapes, and experimental genomic comparison across organisms. Employing these methods invokes additional families of mathematical models, additional layers of experimental work and above all additional knowledge requirements to integrative systems biology in the form of evolutionary, and also developmental, theory. Evolutionary systems biology is well placed to extricate the role of non-linearity in evolutionary processes and the generation of novelty. However if the problem of developing best research strategies and most effective collaborative organization is complex in the case of integrative systems biology, then it is even more so for evolutionary systems biology, and these lessons from integrative systems biology could undoubtedly be an instructive model and source of experience for the latter.

Our research suggests that at this point integrative systems biology is in a fluid state regarding development of research and training practices. As we have tried to show, part of this fluidity traces to distinct choices over research strategies that researchers pursue and lab organization that lab directors institute with different trade-offs - particularly experimental support and molecular detail against mathematical and computational sophistication, given the real difficulties of interdisciplinary collaboration. Lab directors and researchers can themselves factor their reflections on these trade-offs and the challenges they involve for their researchers into their own decision-making about how to go about systems biological research. Our cases suggest these strategic choices are influenced by the epistemic and cognitive perspectives underlying them which favor different preferences with respect to the aims and independence of systems biology, indicating the deeper role these preferences are

playing in the evolving structure of systems biology, and their importance to any attempts to evaluate and understand methodological choices in systems biology.

Finally it is worth pointing out that systems biology has not been alone in attempting collaborations between experimentalists and modelers from different fields and disciplines with distinct expertise. In fact there have been many historical instances of biologists collaborating successfully with others outside their discipline. Structural biology for instance in the form of Nuclear Magnetic Resonance research and spectroscopy studies provides many examples of successful ongoing collaborations involving physicists, chemists and biologists. Biomedical Engineering is an even more modern instance of the successful institutionalization of collaborations between engineers and biologists. Undoubtedly there are lessons systems biology can take from these experiences. However at this point in time there has been very little consideration either by the researchers we have interviewed or by philosophers of what could be taken from these other contexts and how well their lessons might adapt to the case of systems biology. There is thus an obvious next step for this research to broaden its scope to consider not just additional laboratories with different structure but also to inform systems biologists on the relevance of other historical experiences and strategies for combining experimentation and modeling skills.

Acknowledgments:

We appreciate the support of the US National Science Foundation in conducting this research (DRL097394084). Miles MacLeod's participation was also supported by a postdoctoral fellowship at the TINT Center, University of Helsinki. We thank the directors of Lab C and Lab G for welcoming us into the lab and the lab-members of those labs for 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 also thank two anonymous reviewers and associate editor of this journal, Ingo Brigandt, for their helpful comments.

Brigandt I, Love AC. 2012. Conceptualizing evolutionary novelty: moving beyond definitional debates. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution 318:417-427.
Bruggeman FJ, Westerhoff HV. 2007. The Nature of Systems Biology. TRENDS in Microbiology 15:45-50.
Calvert J. 2010. Systems biology, interdisciplinarity and disciplinary identity. Collaboration in the New Life Sciences, Aldershot: Ashgate.
Calvert J, Fujimura JH. 2011. Calculating life? Duelling discourses in interdisciplinary systems biology. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences.
Collins HM, Evans R. 2002. The Third Wave of Science Studies Studies of Expertise and Experience. Social studies of science 32:235-296.
Galison P. 1997. Image and logic: A material culture of microphysics. University of Chicago Press.
Kitano H. 2002. Looking Beyond the Details: a Rise in System-Oriented Approaches in Genetics and Molecular Biology. Current genetics 41:1-10.
MacLeod M, Nersessian NJ. 2013. The creative industry of systems biology. Mind & Society 12:35-48.
MacLeod M, Nersessian NJ. 2013a. Building Simulations from the Ground-Up: Modeling and Theory in Systems Biology. Philosophy of Science (forthcoming).
MacLeod M, Nersessian NJ. 2013b. Coupling Simulation and Experiment: The Bimodal Strategy in Integrative Systems Biology. Studies in History and Philosophy of Biological and Biomedical Sciences http://dx.doi.org/10.1016/j.shpsc.2013.07.001,.
Nersessian NJ, Newstetter WC. 2014. Interdisciplinarity in Engineering. In: Aditya J, Olds B editor^editors. Cambridge Handbook of Engineering Education Research. Cambridge: Cambridge University Press. p. (in press).
O'Malley MA, Dupré J. 2005. Fundamental Issues in Systems Biology. BioEssays 27:1270-1276.
O'Malley MA, Soyer OS. 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 43:58-68.
Soyer OS, O'Malley MA. 2013. Evolutionary systems biology: What it is and why it matters. BioEssays.
Wagner GP. 2007. How wide and how deep is the divide between population genetics and developmental evolution? Biology and Philosophy 22:145-153.
Voit EO. 2000. Computational Analysis of Biochemical Systems: A Practical Guide for Biochemists and Molecular Biologists. Cambridge: Cambridge Univ Press.
Voit EO. 2013. A First Course in Systems Biology. New York: Garland Science.