MODEL-BASED REASONING IN INTERDISCIPLINARY ENGINEERING

Nancy J. Nersessian and Christopher Patton

1 INTRODUCTION

Research in biomedical engineering often confronts the problem that it is both impractical and unethical to carry out experiments directly on animals or human subjects. In our studies of two pioneering biomedical engineering research laboratories we have found a common investigative practice in this interdisciplinary field is to design, build, and redesign *in vitro* systems, which parallel selected features of *in vivo* systems. The researchers refer to their *in vitro* models as "devices." When biological and engineering components are brought together in an investigation, researchers refer to this as a "model-system." As one respondent stated: "*when everything comes together I would call it a 'model-system' [. . .]I think you would be very safe to use that [notion] as the integrated nature, the biological aspect coming together with an engineering aspect, so it's a* ***multifaceted modeling system*** *I think that's very good terminology to describe that.*" Another researcher aptly referred to the processes of constructing and manipulating these model-systems as "*putting a thought into the bench top and seeing whether it works or not.*" The "bench top" refers not to the flat table surface but comprises all the locales where experimentation takes place. These instantiated "thoughts" (mental models) are physical models (devices) that represent what researchers deem to be salient properties and behaviors of biological systems. They are structural, behavioral, or functional analogs of *in vivo* phenomena of interest. The devices are also systems themselves, with engineering constraints that often impose simplifications and idealizations unrelated to the biological systems they model. In the following analysis we will examine some of these multifaceted systems in the problem-solving practices of the laboratories, especially as they figure in experimental situations. In each case we will examine how manipulating devices and model-systems enables a form of inference — "model-based reasoning" — different from logical inference through manipulating propositional representations.

Our analysis derives from a five-year investigation of the research practices of two laboratories; one conducts tissue engineering, the other, neural engineering. These are *hybrid* engineering and science environments. The hybrid nature of these laboratories is reflected in the bio-engineered model-systems developed by the laboratories and in the characteristics of the researcher-students who are part

Handbook of the Philosophy of Science. Volume 9: Philosophy of Technology and Engineering Sciences.
Volume editor: Anthonie Meijers. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

of a program aimed explicitly at producing interdisciplinary, integrative thinkers in bio-engineering. The laboratories and learning settings are designed to move beyond the traditional model of collaboration among engineers, biologists, and medical doctors to a new kind of integrative biomedical engineering that will shorten the span between laboratory research and medical application. This is the goal; in reality the research is still not advanced sufficiently for the desired medical applications, and integrative biology and engineering dominates practice.

"Lab A" seeks to design off-the-shelf vascular tissue replacements for the human cardio-vascular system. Some intermediate problems that drive the research are: producing "constructs" (models composed of living tissue that mimic properties of natural vessels); examining and enhancing their mechanical properties; and creating endothelial cell sources through mechanical manipulation of stem cells. Although the ultimate objective of the laboratory is to make an artifact, researchers create both new knowledge (e.g., of properties of cells under various conditions) and new know-how (e.g., of tissue engineering techniques) as part of the problem situation. "Lab D" seeks to understand the ways neurons learn in the brain to create aids for neurological disabilities or, more grandly, "*to make humans smarter*" (Lab D Director). Its intermediate investigations center on finding evidence of plasticity in a "dish" of multi-electrode neuron arrays, and producing controlled "muscle" activity in robots or in simulated agents (animats), which constitute their model-systems. Since most past research has focused on single neurons, and so little is known about how brains learn, the primary objective of this research is also to create new knowledge and know-how. Significant to our analysis, the frontier nature of both laboratories nature demands that they also design and build novel technologies of investigation.

The methods of our analysis are in some respects unusual for philosophy of science and technology, though in accord with a naturalist epistemology. The analysis derives from data we collected in a five-year ethnographic and cognitive-historical study of the two laboratories. One objective of the study was to extend Nersessian's previous work on model-based reasoning (see, e.g. Nersessian [2008]) with conceptual models to physical and computational models. Another was to attempt an integrative account of cognitive, social, cultural, and material factors in their development and use [Nersessian, 2005]. As with other ethnographies, that part of the study uses observations and interviews to uncover the activities, tools, and interpretive frameworks that support the research as they are situated in the on-going practices of the community. The cognitive-historical part of the study collects and analyzes data from traditional historical sources (publications, grant proposals, laboratory notebooks, and technological artifacts).[1] The aim is to capture the diachronic dimension of the research by tracing the intersecting trajectories of the human and technological components of the laboratory, con-

[1] Although some of the material we quote from comes from published sources, given the regulations governing confidentiality for human subjects research, if the authors are among subjects we are not able to provide citations to that material here. It seems that the possibility of conducting historical research in conjunction with research on human subjects was not anticipated!

ceived as an *evolving cognitive-cultural system*, from both historical records and ethnographic data. This novel combination of methods enables developing thick descriptions and analytical insights that capture the dynamics of inquiry characteristic of research laboratories.

In approaching the problem of integration, the trick is to create accounts that are neither cognitive with culture tacked on nor the reverse, and this necessitates re-thinking current interpretive categories. In our research we attempt a shift in analytical approach from regarding cognitive and socio-cultural factors as independent variables to regarding cognitive and socio-cultural processes as integral to one another. One way of making the shift towards integration would be to construe cognitive processes as comprising more than what takes place in the head of an individual scientist, and analyze scientific thinking as occurring within complex cognitive-cultural systems comprising humans and artifacts. A central claim of our analysis, then, is that inference is performed through interlocking mental and physical models and that the devices serve as *hubs* for interlocking cognitive and cultural facets of laboratory research.

Our strategy in this paper is to describe modeling practices in each laboratory, and to focus on one example of experimental investigation using a model-system from each lab. We then discuss more generally the nature of the reasoning involved in constructing, manipulating, and revising model-systems. One problem often noted with such "case study" methods is how one generalizes from a specific case or any number of cases. The notion underlying this problem is 'inductive generalization' from cases. We subscribe, rather, to the ethnographic notion of 'transference' across sites/cases. In ethnography one develops richly detailed descriptions of a site. When conducting comparative investigations of multiple sites, thick descriptions are created of each, and these are examined to see what might possibly transfer across sites — as well as what of importance does not. Although the specifics in each case differ, our insights about model-based reasoning do transfer across laboratories.

The main conclusions of our analysis derive from the observation that designing, redesigning, and experimenting with *in vitro* simulation models (devices) is a signature investigational practice. Physical simulation is an epistemic activity involving hypothesis exploration, testing, and generation, as well as prediction and explanation. The devices serve as hubs for *interlocking* biological and engineering concepts, methods, and materials, mental and external representations in model-based inference, design and history, and research and learning. For brevity we refer to this multidimensional notion as *interlocking models.*

2 TISSUE ENGINEERING LABORATORY: THE "VASCULAR CONSTRUCT MODEL-SYSTEMS"

Lab A has as its ultimate objective the development of artificial blood vessels (locally referred to as "constructs"). These are engineered out of living tissue and will need to have the appropriate characteristics to function within the human body,

such as sufficient strength to withstand the forces of blood flow and endothelial cells lining it that are able to proliferate. An *in vivo/in vitro/ex vivo* division is a significant component of the cognitive framework guiding practice in Lab A. The test bed environment for developing artificial blood vessels cannot be the human body, so the researchers have to design in vitro facsimiles of the *in vivo* environment and *ex vivo* implantation environments (so-called "animal models") where experiments can occur. In this context, '*ex vivo*' refers to an animal that has been altered such that experimentation can take place external to its body. The major challenge is to bring together biological and engineered materials with the desired properties so as to perform properly *in vivo*. As characterized by the Lab Director, the "major barriers" fall into two categories: mechanical properties of the tissue and the influence of mechanical forces on the tissue and cell source strategies that support endothelialization.

Many aspects of the *in vivo* phenomena are known and understood both in biological and mechanical terms, but many are not, such as how cells proliferate. Thus the lab makes contributions (new knowledge and know-how) to basic biology and to engineering applications. The daily research in Lab A is directed towards solving problems that are smaller pieces of the grand objective, such as proliferating endothelial cells within the constructs, which involves, for instance, gene profiling studies, and creating constructs that can withstand the powerful mechanical forces of blood flow *in vivo*, which involves advancing collagen gel technology. In the next sections we discuss some of the central devices that form the major components of model-systems in these labs and how they interlock in experimental practice.

2.1 The flow channel device

Lab A has been in existence since 1986 and was created specifically to move the research, as expressed by the Lab Director, "*from animal studies to cell culture.*" Early bioengineering experimentation on the vascular system was conducted by the Lab Director and colleagues on blood vessels that were altered while in the living animal. Through surgical interventions blood vessels were made to exhibit pathological conditions consisting in narrowing of native arteries (stenosis). After sacrificing these animals, the morphology of the cells that line the arterial walls at the pathologically altered regions was studied and specific aspects quantified (e.g., elongation and orientation of cell filaments). Simultaneously, arterial flow patterns (velocity profiles) were studied for pathological narrowing of the arteries through creating models that replicated the *geometrical dimensions* of these observed pathological regions. These structural models were achieved by filling arterial vessels of sacrificed animals with a fluid plastic. After hardening they were used as casts to manufacture replica of the narrowed vessel. These replica models were used in experimental "flow studies" where laser Doppler techniques were used to determine velocity patterns. The results gained from studying cell morphology and from studying velocity patterns in the replica models were correlated to gain

insights into the relations between variations in wall shear stress due to particular velocity patterns (gradients near the vessel wall) and the morphology of cells lining these vessels. The elaborate material and measurement practices related to the replica models were quickly abandoned, but they launched the director's program of studying the impact of flow in arterial wall shear stresses *in vitro* with engineered devices, the research agenda of Lab A.

Along with the replica studies and the associated cell morphology studies, the director and other researchers started a line of research with *cell cultures* of the endothelial cells typically lining the arterial walls. This work provided the basis for the initial configuration of Lab A. Instead of inducing stenosis in living animals, and thus creating particular flow patterns resulting in particular wall shear stresses, they instead exposed the respective cell type in culture to wall shear stresses by "flowing" them in *a flow channel device* (called the "flow loop" in the Lab). These *in vitro* experiments on the response of the cells to shear stresses were based on an established fluid dynamic model, specifically, the fluid mechanics of a long channel with rectangular cross-section. Using this device, changes in cell morphology (elongation and orientation) could be related directly to the controlled wall shear stresses. Furthermore, the method of measuring velocity patterns in a replica model was replaced by an engineered model of exact geometrical specification, a flow channel. With the controlled flow channel the correspondence between the mathematical and the physical model became an issue of engineering a channel with the appropriate dimensions (in a physiologically meaningful range), rather than measuring velocity patterns using elaborate laser Doppler technology. The studies using the replica model had in fact dissociated the study of cell morphology from the study of flow patterns, correlating their results after the fact. The flow loop in action is a model-system in which the two foci of study are concentrated into a single system where cultured cells were exposed to flow and thus shear of a well-defined nature.

The move to *in vitro* solved problems related to the fact that it takes twenty four hours to see results of interventions made in animals and during that period many physiological changes take place. However, as a model, it does not represent the diachronic nature of the *in vivo* environment, even though it is dynamic. Blood flow *in vivo* changes, for instance, when eating and sleeping. So, eliminating the confounding factors leads to a simulation of the *in vivo* environment that is *"something very abstract because there are many in vivo environments and there are many in vivo conditions within that environment. Things change constantly in our bodies over our lifetimes; including physiological flow rates [....] So I don't think we are trying to mimic the exact conditions found in vivo.*"

In the cell-culture line of research, an engineered artifact, a flow channel with the accompanying flow-inducing components serves as an *in vitro* model paralleling certain *in vivo* conditions of the blood vessel, including both normal conditions and the pathology that previously was induced in living organisms. The flow loop represents a first-order approximation of shear stresses during blood flow in the artery, "*as engineers, we try to emulate that environment, but we also try to*

eliminate as many extraneous variables as possible. So we can focus on the effect of one or perhaps two, so that our conclusions can be drawn from the change of only one variable." The flow loop provides "*a way to impose a very well-defined shear stress across a very large population of cells such that their aggregate response will be due to [it] and we can base our conclusions on the general response of the entire population.*" The speed at which the flow loop pump operates reflects knowledge of how fast blood flows *in vivo*, and the pulse dampener turns the flowing liquid from pulsating flow to a smooth flow that allows control over the constancy of flow. The flow loop allows manipulating the amount of shear stress (speed of flow), duration, and height of the chamber. Using it in flow simulations also enables manipulation of cells or engineered vascular constructs. As the laboratory began to establish itself, flow loop studies on cultured cells were conducted by all members. Over time, new simulation devices have been designed not only to understand mechanical forces creating pathology, but for creating the vascular constructs that will some day repair diseased arteries, as will be discussed below.

From the outset, redesign has been a central activity within Lab A. The flow loop provides a major instance of this activity. In working with cultured cells, contamination is a constant problem and this problem was the driving factor in the redesign of the flow channel device. An interview with a former graduate student, now a successful faculty member at another institution, elaborates on this problem and the subsequent redesign:

> "*So, when I got here in 1994 uh, the flow chamber was a mess. It was a benchtop system, it had bulky tubes that looked something like some time machine from the 1950s or something [. . .]. But anyway it was quite messy and you know culture studies have to be done at 37 degrees so the way that they would do this was you know, incubators were certainly around in 1994, uh, they would wrap these coils, these heating coils around these glass reservoirs and because it had to be a set flow, they would use a hydrostatic pressure difference to derive the flow, and uh, a clamp, a regulated clamp to try and regulate the flow through the chamber and out into the — into the lower reservoir. So you had two reservoirs, one at the top, and one at the bottom, there'd be a hydrostatic difference between them, and then things would flow and then this whole thing would be sitting on the benchtop — big bulky glass reservoirs with bulky tubing [...]. And this was subject to about a 50% success rate.*"
>
> Interviewer: In terms of contamination?
>
> "*In terms of contamination. And the reason was because this whole thing had to be assembled outside of the hood [colloquial for 'the sterile workbench']. There was no way you could assemble this thing to stand up-this thing was on stands–you have to assemble this part outside of the hood, so basically they we would connect these joints here, and connect them outside of the hood. [. . .] Doing experiments longer*

than 48 hours was almost impossible, because at experiments longer than 48 hours the incidence of contamination was probably greater than 90%. [.....] I really like compact designs [. . .]. I instituted a lot of the things I saw over there [referring to the lab at which he interned] in our laboratory, and one of the things was model-revising this design to go into the incubator. And, that was really why we moved from a system that required heating coils and an upper reservoir and a lower reservoir to a system that was just flow driven with a peristaltic pump and a pulse dampener that was — and everything could be done inside the incubator with smaller tubing, little reservoirs as opposed to big reservoirs."

"Model-revising this design," as the former graduate student described his contribution to this line of research, meant redesigning the physical system that is the flow channel device, its parts (e.g., the reservoirs, the tubing), set-up (e.g. on stands in the lab vs. compact and in the incubator), and the physical principles governing its functional design (e.g. hydrostatic pressure difference vs. integration of a peristaltic pump). The actual flow channel, which is the part where the liquid flows over the cell cultures, was left untouched in this particular redesign. Even though redesign, in this case, did not involve those parts where the cells-in-culture interfaced with the mechanical device, re-engineering this design was central to its function as part of a model-system, which is totally dependent on its resistance to contamination of the cell cultures. The set-up that functions as the model-system is sufficiently decomposable to allow for the independent redesign of its various components.

Since then the flow loop has undergone minor redesign, such as related by a current Ph.D. student, because of the introduction of a new device, the construct (described in detail below). In discussing her own redesign, she started with telling of how the researcher just prior to her had modified the flow block to solve some technical problems. The modified device that she inherited had previously been used on cells. She now wanted to use the flow loop to experiment with the vascular constructs seeded with endothelial cells, cut open and placed flat for flowing. These flat constructs are thicker than the muscle cells used before, and bumpy. Because of these features, spacers need to be used between the block and the glass slides in order to improve the flow pattern around the boundary to bring the *in vitro* model more in accord with flow in the *in vivo* model. To begin this research, she, together with another new student, had to redesign the device by changing the width of the flow slit to hold the spacers. Most recently another student planned a significant redesign to enable flowing of constructs in tubular shape in order to accommodate implantation in an "*animal model*" that will be discussed in the next section. This redesign would mark a significant step in the move towards *in vivo* implantation in that the constructs would not need to be cut open in order to be flowed.

2.2 The "construct"

The endothelial cells that form the endothelium, a mono-layer of cells that make up the inner lining of the blood vessel are a major target of study in the Lab. *In vivo* these are in closest proximity to the blood flow. The culturing of cells needs to emulate the naturally occurring conditions of living tissue in an organism to the extent that cells are required to survive and perform in particular ways. This emulation requires such things as the appropriate CO_2 levels in incubators and that the incubators keep the cells in the requisite temperature range. Moreover, the types of cells identified for embedding in the vascular substitutes must readily be available and compatible with adjacent tissues. This requires a method for ensuring cell growth and proliferation and cell sources for production. So the introduction of the construct has led also to a line of stem cell research. Until the late 1990s flow loop experiments were done only on cell cultures on slides usually coated with a substance to make them adhere. After flowing they are removed and examined through various instruments such as the coulter counter and the confocal microscope to determine the effects of the mechanical properties of the flow on shape (morphology), alignment, proliferation (reproduction), or migration (locomotion). However, as one researcher observed, "*cell culture is not a physiological model; however, it is a model where biologic responses can be observed under carefully designed and well defined laboratory conditions.*" So, although many experiments are still done on cell cultures, the problem of constructing a tissue engineered vascular substitute has led to the creation of a new simulation device, the *construct*. As the lab director recounted, the current research aims

> "*to use this concept of tissue engineering to develop better models to study cells in culture. Putting cells in plastic and exposing them to flow is not a very good simulation of what is actually happening in the body. Endothelial cells, which have been my focus for thirty years, have a natural neighbor called smooth muscle cells. If you look within the vessel wall you have the smooth muscle cells and then the inside lining is the endothelial cells, but these cell types communicate with one another. So we had an idea: let's try to tissue engineer a better model-system for using cell cultures.*"

The construct marks a move towards a more physiological model — one whose function is more like the in vivo model along mechanical, physical, and biochemical dimensions. An actual blood vessel is in tubular shape and comprises several layers: the lumen where the blood flows; a first, mono-layer of endothelial cells that sit on collagen, an internal elastic lamina, a second layer of smooth muscle cells, collagen, and elastin, external elastic lamina, and a third layer of loosely connected fibroblasts. *In vivo*, the cells create an extracellular matrix which is a network of proteins and other molecules, and provides growth factors and mechanical properties. In the *in vitro* culturing process, the construct is seeded onto a tubular shaped silicon sleeve. Unlike cell cultures on slides, constructs are three-dimensional surfaces in which cells are embedded. The construct is a "wet"

device — a living "*blood vessel wall model*" that simulates in vivo processes. It is a significant component of many current research projects. The hope is that "*our construct behaves like a native artery because that's one step closer to being functional. So we are kind of doing a mimicking thing. So does it respond in the same manner?"* When eventually used as replacement systems for the human body, the biological substitutes must replicate the functions of the tissues being replaced. This means that the materials used to grow these substitutes must coalesce in a way that mimics the properties of native tissues. It also means that the cells that are embedded in the scaffolding material must replicate the capabilities and behaviors of native cells so that higher level tissue functions can be achieved. To "respond in the same manner" means, among other things, that it expresses the right proteins and genetic markers.

In building a construct, different levels of approximation are used depending on the nature of the experiment. It is possible, for instance, to use only collagen and not add elastin. Most often the cells are not human aortic endothelial and smooth muscle cells. Some experiments are conducted with only a single layer of the blood vessel wall with either endothelial cells or smooth muscle cells. Thus the construct forms a family of models, designed for different experimental purposes. In the experiment we outline below, conducted with baboon cells with a baboon animal model, the third layer was not engineered because the researcher both deemed it unnecessary for her experiment and also surmised that it would grow itself. In addition, she used a teflon scaffolding on the outside of the construct since it is not strong enough to withstand the forces of the baboon's blood flow. Most significantly, in experiments with the current flow loop, the constructs need to be cut open to lay flat within the flow chamber as designed. The fact that the flowed construct is flat due to the design of the flow loop whereas the blood vessel is curved is an approximation to the tubular surface. However, since the cells are so small with respect to it, the flatness is not an approximation for these main objects of study. Indeed, from the "*cell's perspective*" it is not an approximation since "*the cell sees basically a flat surface. You know, the curvature is maybe one over a centimeter, whereas the cell is like a micrometer — like 10 micrometers in diameter. It's like ten thousandth the size, so to the cell it has no idea that there's actually a curve to it.*" That is, flowing liquid over these flat constructs will give accurate enough data on the responses of the endothelial cells that line the arterial wall, because the cells are so small with respect to the arterial wall that the cell's experience of the wall is as though it lives in a flat world. The medium flowing over the construct lacks a vertical component (flow is two-dimensional) and is unidirectional whereas *in vivo "it sloshes around in the blood vessel."* But again the focus is on first-order effects unless there is evidence of a need to consider higher-order effects, such as if they find *"that there's a whole different pattern of genes that are up-regulated in pulsatile shear, or something, maybe then it would be more interesting to use different constructs and stuff*," so as to be able to look at the higher-order effects.

The introduction of the construct led to a range of new devices for experimenting on it directly, such as the pulsatile bioreactor that was designed for "*mimicking the wall motions of the natural artery.*" It is used to expose the constructs to mechanical loads in order to improve their overall mechanical properties, what the researchers call "*exercising the cells.*" Preferably, this is done at an early stage of the formation of the construct, shortly after seeding onto a prepared tubular silicon sleeve. *In vivo*, arterial wall motion is conditioned upon pulsatile blood flow. With this bioreactor, though, which consists of a rectangular reservoir containing a fluid medium in which the tubular constructs are immersed and connected to inlet and outlet ports off the walls of the reservoir, the fluid does not itself move. Rather, the sleeves are inflated with pressurized culture medium, under pneumatic control (produced by an air pump). The medium functions as an incompressible fluid, similar to blood. By pressurizing the medium within the sleeves, the diameter of the silicon sleeve is changed, producing strain on the cells, similar to that experienced *in vivo*. However, as a model-system, "*the silicon sleeves add the next level of doubt. [...] The construct itself is not actually seeing the pressure that the sleeve does. And because of that you know — it doesn't actually see a — a pressure — it feels the distention but it doesn't really feel the pressure. It doesn't have to withstand the pressure. That's the whole idea of the sleeve.*" These differences between the *in vivo* and *in vitro* models arise from the nature of the devices themselves. The construct at present is not strong enough to sustain the actual forces of pulsating blood flow and the bioreactor itself has been designed only as a functional model that achieves a parallel motion through different behavior.

2.3 Preventing platelets: experimenting with a vascular construct model-system

In experimental situations models tend to be in interlocking configurations, that is, not as isolated entities but rather as standing in particular relations to other models. The diagram in Figure 1 is a schematic representation of our analysis of the *vascular construct model-system* for one proposed experiment aimed at solving the problem of platelet formation on constructs, and the resulting thrombosis.

The diagram traces the construction, manipulation, and propagation of models within the system that constitutes the experiment. In Figure 1 the models involved are highlighted by thick lines.

The experiment is significant because it constitutes a first move in the direction of *in vivo* research in that the animal model serves as a model for the human system in the context of the experiment. In this experiment an exteriorized shunt connecting the femoral vein and the femoral artery has been placed in a baboon so that a small amount of blood flow can be diverted through a construct during an experiment. The baboon's blood is injected with iridium so that platelets will be made visible through the gamma camera, a commercially available instrument. Each physical model is constructed to represent and perform as a selected aspect of the cardiovascular system, for example, media and constructs represent and per-

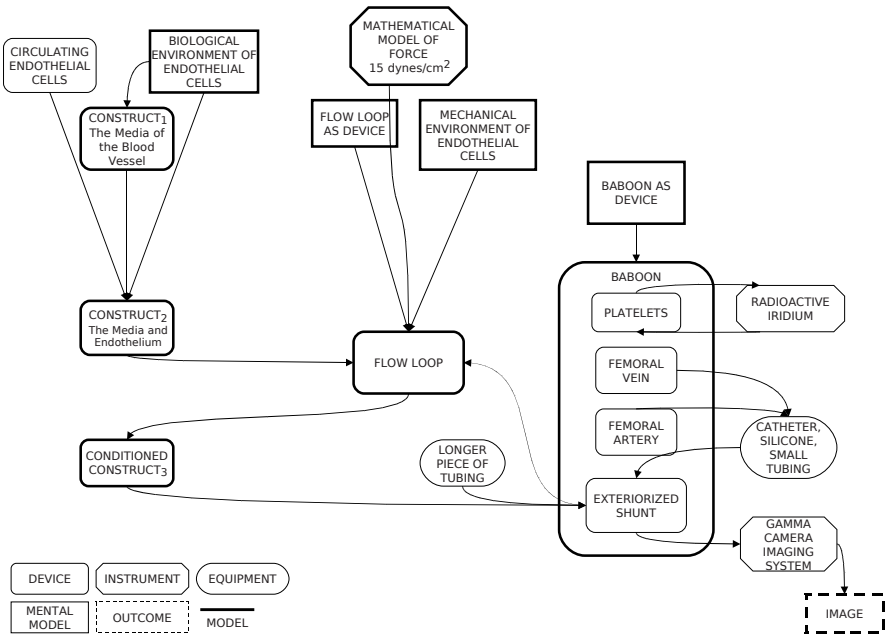

Figure 1. A vascular construct model system

form as the biological environment of the blood vessel and the flow loop represents and performs as shear stresses do on arterial walls. This experiment initially proposed the redesign of the chamber of flow loop to better approximate the in vivo model in that constructs would be flowed in tubular shape, which is necessary in order to implant them. In the end this turned out to be unnecessary because the researcher had the insight that it should be possible to design an external shunt for the flow loop — on a direct analogy with the shunt used in the baboon model — and attach the construct to it for flowing. During the several re-iterations of the experimental simulation over time, the force with which the media flowed through the construct was adjusted until platelet formation ceased.

In general, each experiment with the model-system comprises a number of interlocking models which include at least the following: community models of *in vivo* phenomena (biological, mathematical, mechanical); engineered *in vitro/ex vivo* physical models of aspects under investigation; and mental models of *in vivo and in vitro* phenomena, devices qua *in vitro/ex vivo* models, and devices qua engineered models. Each mental model is both an individual and a community achievement. Note, too, that models from mathematics, physics, and biology inform the construction of the devices. We return to this notion of interlocking models and to model-based reasoning after considering the model-systems of Lab D.

3 NEURAL ENGINEERING LABORATORY: THE "EMBODIED DISH MODEL-SYSTEMS"

Lab D is a cutting-edge neuroscience research laboratory whose focus is understanding the fundamental biological nature of learning and memory. Specifically, they are interested in "*network level plasticity, learning, and memory*" at the cellular level. Historically, neuroscience has been practiced at the two extremes on the spectrum of granularity. At one end there are neurologists who seek a gross, system-wide, psychological understanding of the brain, and at the other end are single-cell electrophysiologists who seek to understand the fundamental properties of a single neuron. This lab, however, sees itself as breaking ground in "*mezzo*" (meso) level neuroscience. The researchers are interested in the nature of networks of cortical neurons, which are too small and unorganized to be considered a functioning brain, but are at a higher level of organization that the single neuron.

Early attempts at investigating the nature of networks of neurons depended on cutting slices of a brain, fixing the neurons (in formaldehyde), and then looking at the physical structure of the neurons underneath a microscope. The Lab D Director was not happy with looking at dead neurons and instead wanted "*to look at living things, while the interesting parts would happen.*" To this end, he spent many years of postdoctoral research designing and perfecting what has become Lab D's most prominent model: "The Dish," and also developing technology for imaging it.

3.1 "The Dish"

Simply put, "The Dish" is a network of cortical neurons living, not as part of a larger brain, but as a small network in a Petri dish. Embedded in the bottom of this Petri dish are 64 electrodes capable of recording and injecting electrical activity in the network. The basic construction of the Dish is as follows:

- *The MEA*: the Microelectrode Array (MEA) is a small, glass, Petri-style dish which has an 8x8 grid of micro-electrodes embedded in the bottom of it. The electrodes poke upwards from the bottom of the dish into the neurons.

- *Cells*: Lab D uses parts of a rat's cortex as their network. The two basic types of neural cells, neurons and glia (support cells), are harvested from the rat and then cultured in this dish.

- *Medium*: A sugary cocktail of biological chemicals used to feed the cells.

- *The Lid*: A thin film made out of Teflon is stretched over the MEA and held in place by a thick "O" shaped piece of Teflon (like a slice of pipe). Teflon is used specifically because it is non-toxic and allows O_2 and CO_2 through so that the cells can breathe. However, it is not porous, so it keeps everything else (such as bacteria or fungus) in or out, even when it is stretched.

There are no established models of neuron communication. In fact, the endeavor of Lab D has been to build these models from the ground, up, as expressed by the Lab Director:

> *"well the level of modeling that I'm very interested in is the network modeling. So there are a lot of people out there who are doing very detailed modeling of single neurons, in fact even [in] this lab X and Y are doing that, but I'm much more interested in what happens when you take a bunch of cells and have them interact with each other, what sort of emergent properties pop up. So, it probably isn't necessary to include all the details that you would find in a [single neuron] model, with ions and stuff, but it may be, so that's part of our job is to find out which of the details of biology are important in these sort of network properties and network phenomena and which are sort of incidental."*

What is growing in the dish, then, is not a brain, nor is it even a slice of a rat's brains. Instead, the neurons are disassociated (their connections are broken apart to the point that they free from each other), plated (processed and placed on the MEA), and then cultured so that their connections are allowed to grow back together. As, one researcher described their reasoning for using The Dish as opposed to a brain slice:

> *"Yeah, one single-layer of neurons. We try to get them down in a monolayer. That's the whole idea. It's a simpler system to study then. That's all. And also, I mean, you could do slices that are not disassociated cultures but the problem in slices is it is difficult to maintain them. You have to kind of get the fluid [in] right. Sometimes only the outer cells — you know the medium does not go into the inner layers. The inner layer dies off but the outer layers are fine, stuff like that. So, it's more difficult to keep them alive. And we want to study over a long term, so we want to keep them alive over months, years."*

The Dish as a physical model of the brain of a rat that is constrained by both the state of Lab D's current understanding of the dish, as well as by technical hurdles having nothing to do with the biological model. In response, Lab D has chosen to simplify their model to a single layer of neurons to reduce the number of possible variables in the system. Doing so gives them a smaller set of questions to ask in order to build their understanding of network processing. However, they believe that this simplified model provides a close enough approximation to yield valid information.

The Dish is the lab's central model-system: the physically constructed embodiment of the Lab's selective model of the brain. It is an *in vitro* model of basic *in vivo* neurological processes. With an *in vivo* neural system there is specific structure to the pathways and connections in the networks, created by the developmental processes of the animal. Here, though, the neural structure is destroyed

so that The Dish does not directly reflect any structure within a living rat. Consequently, there is some contention over what The Dish actually models. Some maintain that The Dish is "*a model of cortical columns*" while others think it *"is a model of development [of the brain],*" and when pressed, some even admit that "*it may just be a model of itself.*" However, they all agree that studying The Dish will yield understanding of the basic workings of network-level cortical neural processing, as explained:

> *"First of all, it's a simplified model, I say that the model is not, it's artificial, it's not how it is in the brain, but I still think that the model would answer some basic questions, because the way the neurons interact should be the same whether it's inside or outside the brain. . . . You know, because we are in an artificial environment, it's not the same, you know, it's not the same concentration as it's in the brain; nothing is the same. I'm growing in an external environment. But, I think that the same rules apply."*

The Dish is designed to be a generic model of cortical processing behavior and function. The researchers are not seeking to understand the processing of a specific construction of a particular rat's brain. Instead, they are interested in how networks of neurons — in general — communicate and process information. There is clearly an intentional lack of specificity in this model in order to be able to build their understanding of neural processing in general. This generic understanding of the networks is not the end goal, however. After building the generic basis of understanding, they plan on building a more refined understanding of neural communication. As stated by the Lab Director:

> *"Clearly it's missing a lot of other brain parts that are important in what brains do. I'm assuming they are important. And at some point we might be studying cultures with different brains parts mixed together, or specific 3-dimensional pieces that are put together."*

The physical construction of the model-system not only frames the inquiry, but also is an integral part in the research progress of the lab as a whole. Only once the researchers understand their current physical model and construct an accurate abstract model (mental, mathematical, etc.) can they progress in the construction of a new model that more closely approximates the *in vivo* situation.

The construction of The Dish itself stretches across biology, chemistry, and electrical engineering, and requires the entire endeavor within Lab D to be interdisciplinary. Though Lab D's research focuses on a mostly biological entity, it is largely populated by members with backgrounds in electrical engineering and mechanical engineering — and not biology or neuroscience. Unlike Lab A, there are no telltale signs of biology: no flasks, no pipettes, no hazardous waste containers. In fact, aside from a conspicuously covered microscope and an incubator that could be mistaken as a mini-fridge, Lab D's space looks more like a computer lab than anything else. The most striking features of the Lab are the copious wires

that crisscross the space carrying electrical signals produced by the neurons to the researchers and their computers.

Although all of the researchers use The Dish, there are several different avenues pursued in this lab including pharmacological studies, morphological (imaging) studies, and simulation with both computational models and physical models. Here we highlight the Lab's electrophysiological investigations. The main goal of the research is to understand how to "communicate" electrically with the network of neurons, and to try to see — indeed to figure out what counts as — evidence of network plasticity (learning). The core of the method involves electrically stimulating a biological neural network and recording the electrical response of the dish. To this end, the electrophysiology technology in Lab D centers on finding new ways to "talk" to The Dish, and then, in turn, trying to understand what the dish is "saying" back to them in response. We will outline three ways in which experiments are conducted with Dish model-systems.

3.2 "Talking" to The Dish: electrophysiology

The objective of the experimentation in Lab D's electrophysiology research is to understand how information is encoded and processed in The Dish networks. It would be ideal to have direct readings from every single neuron in The Dish. However, with current technology this is impossible. Consequently, access to the Dish is mediated by a comparatively small set of electrodes (the MEA). The signals ultimately received are a *representation* of neural activity filtered through several models which will be discussed below. It is currently impossible to know the actual neural activity in The Dish. What they study are the "spike" data recorded from the electrodes of the MEA.

Historically, the term "spike" refers to the electrical trace left behind when an individual neuron fires. In single-cell electrophysiology, it is possible to read electrical activity directly from the neuron, and as a consequence, the model for neural firing is well known: there is a steep jump in voltage potential as the neuron de-polarizes (fires), and then a proportional drop in potential as the neuron recovers. Multi-cellular neuroscience has borrowed the notion of a 'spike', but has modified it to suit their situation. The researchers of Lab D estimate that the electrical activity recorded on a single electrode can come from on average three to five different neurons. When dealing with the electrical traces of many neurons, it is possible (and is often the case) that several neurons around a single electrode will fire simultaneously. As a result, a spike can actually represent the firing of more than one neuron. However due to the fact that they are all on one trace, it is impossible to tell the difference between the firing of a single neuron and the firings of multiple neurons. This altered conception of what a spike represents is shared among the Lab members. Historically, spikes were tagged by hand. However, Lab D has created an automated process of identifying spikes by having developed a piece of software they call the "Spike Detector." The Spike Detector embodies the Lab's model of a spike which includes the spike's "height" (difference

from the average voltage) relative to the noise on the electrode and the "width" (duration) of the change, along with a few other more subtle characteristics. The Spike Detector checks for jumps in the voltage that match this model of what a spike should look like, tags the spike, and keeps a snapshot of the electrical activity immediately around the spike. The researchers begin their analysis with the filtered spike data.

The data that the researchers ultimately use do not come directly from the neural processes in The Dish. There are several pieces of software like the Spike Detector that are used during electrophysiological experiments. These pieces of software are collectively referred to as "filters" and perform a number of different transformations on the raw neural data before the information ever reaches the researcher. Each of these filters embodies a model of an aspect of the neural data. It is possible that the filters miss actual neural firings, or detect a jump in the readings that does not correspond to any action potential (i.e., provide a false positive). Thus the meaning of an individual spike needs to be understood in terms of the filter algorithms that created it, that is to say, in terms of the series of transformations used to produce the data, each one of these transformations being built on a model of the selected electrical signals. The researchers are intimately aware of these processes and perform their analyses in light of this understanding.

In the simplest form of analysis, the spike data are transformed into a visualization rendered on a computer screen. There are a number of different types of visualizations the Lab uses for these basic data, and far more for their higher-level analytical transformations. The simplest visualizations come from the data-capture software they have developed called "MEABench." As the computer is capturing neural data, MEABench can display the neural activity in real-time. The MEABench screen is divided into 64 rectangles (one for each recording electrode) and arranged to topographically match the layout of the electrodes in the dish. Each rectangle, then, displays a trace of the electrical signals captured on the electrode. When the Spike Detector is turned on, the visualization tool places little red dots at the peaks of the spikes for easy visual identification.

The researchers' ability to "see" the electrical activity of the dish, then, involves a number of interconnected mental, physical, and algorithmic models. The visualization is a representation of the information produced by the filters. The filters are instantiated models of Lab D's understanding about neural signals. The signals themselves are a model of neural communication (as they abstract away other factors such as chemical signaling). Finally, The Dish itself is a model of neural functioning in rat cortex. This series of interlocking models is the base model-system for any electrophysiological experiment run within Lab D. This experimental set-up, however, is one of the most basic model-systems used in the lab. In the next section we will describe two far more complex "embodied dish" model-systems: the "Animat" and the "Hybrot."

3.3 Computational embodiment: animat model-systems

> *"[In] the traditional way to do in vitro physiology . . . the closest thing to behavior is little waves on the oscilloscope screen. [It] has nothing to do with any behavior, other than light on the screen there. And there is not any sensory input other than electrical pulses through a couple of electrodes. You know, it is very disconnected, and one of the things that was really shaping my thinking at that time was this book here. This is the first of a proceedings from this conference — simulation of adaptive behavior, here. And, I think it was 1990 that they had this conference, yeah. [A]ll of the people in that book are simulating animals or what they call "Animats." The term was coined around that time by the guys who were at this conference. They were simulating these things on the computer or they were building robots that were animal simulations. They were continually reemphasizing the importance of embodiment and being situated."*

Originally, then, the idea for animat model-systems developed in Lab D came from the domain of computational modeling. There the goal was to create a very simple model "world" and a very simple model "animal" and then simulate the activity of the "animal" in the "world." The Lab Director borrowed this idea of modeling animals, but decided that he could improve on it. While others were using a completely computationally simulated animal, he set out to make a more realistic model of an animal by using The Dish as the "brain."

The term "animat" in Lab D refers to a computationally simulated entity that is controlled by the activity in a Dish. This model-system mimics the fact that learning *in vivo* is embodied, that is, comes about from sensory stimulation and feedback. An animat consists of The Dish (the "brain" of the computational animal) and two translation programs: one designed to be the simulated sensory input apparatus and one to simulate motor output that exists in a simulated environment. In short, animats are used to simulate the embodiment of the neural networks and are used as a primitive model of an animal functioning in the world.

In the description of the experimental set-up we developed in the last section, the electrical signals produced by the neurons were simply recorded and then analyzed. This is typically referred to as "open-loop" electrophysiology. Animats, on the other hand, are part of "closed-loop electrophysiology." Closing the loop simulates embodied learning. The electrical signals produced by the neurons are not simply recorded. Instead, they are transformed in some meaningful experimental way and then fed back into The Dish. The "sensory" information is then run through a translation program that converts it into a pattern of stimulation to be administered by a stimulation board. The stimulation produces an electrical response in the dish which is recorded and run through a separate translation program which converts the signals into "motor" commands for the animat. The motor commands move the animat in its "world" which, in turn, produces a change

in the "sensory" information. The change in sensory information is again read from the electrodes in the dish, and the loop continues.

> *"There are lots of different arguments for it. I think probably one of the best arguments is... if you look at what neurons do they ... they learn things. And that's what we're trying to figure out... how learning works, how memory works. And so if we have neurons in a dish, we want to see them learn something... make associations... it's a bit more obvious to see ... the learning involved in a closed loop [situation].'Cause you define what it is it's going to learn based on ... the body you give it and the environment you allow it to work in."*

As expressed by the researcher, the animats provide a model-system for studying the relationship between the fundamental nature of information processing in the neurons and the visible behavior of animals. The animat model-system serves to demonstrate network plasticity more convincingly than open-loop experiments, and will, in turn, lead to a better understanding of how to interpret and control the activity in the dish. Just as there are many different types of creatures in the real world, Lab D has created an entire family of model animats, one of which is a simulated "moth." Here we briefly consider D2's construction and use of the moth model-system.

As conceived by the researcher,

> *"[t]he original model is basically, I have a circle and then the center of the circle, which this is the environment. The center of the circle, where you can find, what you can think of as a light... and I wanted, the moth, which would be the animat would move towards the light like moths do."*

In the case we analyzed, the screen visualization comprises a circle that delimits the entirety of the moth's world, a dot in the center represents the simulated "light" and lines across the circle represent the paths that the moth has followed during the duration of the experiment. The moth was given a simulated "eye" to "see" the light and the simulated motor ability to move around in the "world."

The two most interesting parts of any animat experiment are the programs that interpret the neural signals. The sensory translation program enables the moth's sensory system to "see" where the light was and "see" where it is, itself, and turns this information into a series of neural signals just like an animal sensory system. Once The Dish is stimulated by this sensory input, it responds with its own electrical activity. This activity is recorded and processed by the method described in the electrophysiology section, and then translated into motor commands for the moth. For instance, the neural activity can be treated as a population vector; in other words, each electrode is taken to represent a possible direction of motion. Activity on the left of the dish would indicate that the Moth wanted to move left; activity on the right would mean the Moth wanted to move right; and so on. In essence, the network as a whole determines which direction to go in based on the

sensory input. Importantly, the translator programs arise from a number of different models. First and foremost, the programs are a model of how D2 understands neural communication, and thus the translation algorithm is an exemplification of his idea about how neurons communicate. Also, put together, the translator programs form a basic approximation of how an animal's sensory-motor system operates.

In sum, an animat, such as the moth model-system is an *in vitro* simulation of an *in vivo* model of neural processing that requires "embodiment," environment, and continual sensory-motor feedback. The sensory translation program derives from a model of the researcher's understanding of how animal sensory systems translate raw sensory information into meaningful information for the cortex to process. The Dish component, itself, is a model-system that represents basic neural processing. The motor translation program embodies a model of how an animal's motor system converts neural output into motor function. Finally, taking the model-system together as a whole, the experimental animat is a simplified model of a real-world animal.

Running this model-system and analyzing the outcome of the experiment is not the end of the story though. As with most model-systems in these types of cutting-edge communities, the animat system is constantly evolving. Using the results of the first moth, the researcher revised his understanding of neural processing and, in turn, revised his model of moth. After running the moth a number of times, he updated the sensory control system, in particular, he "*simplified that even further just to have frequencies only dependent on positive X production... Just to make it, easier to analyze in data analysis... Just to try and simplify everything. I could just say, 'Okay, this part of the stimulation does this.'*" Much like with the Lab's decision to use monolayer cultures, here the researcher chose to design a less realistic, but more easily understandable model. This limiting case abstraction of an animal gives the researchers less degrees of freedom, but provides for more clearly defined questions and experiments that provide results that are less complex and more easily interpreted.

3.4 Physical embodiment: Hybrot model-systems

> *"So the point is, these cultures are basically the model of the brain, right. The point is to have an embodiment for these cultures which is basically a robot. So, you... We have a robot in the other room where you can mount this culture on top of it and you can basically program the robot to do stuff, and see how it behavior changes. So basically, the behavior of the robot controls the input to the dish ... and how it interacts. So we have the robot running around in some kind of environment doing its stuff according to some algorithm we set, which hopefully does some learning, and then we see how the behavior of the robot changes over time."*

A "hybrot" (a contraction of "hybrid robots") provides another kind of embodied experimental model-system for studying neuron learning in. Like an animat, a hybrot provides a generic model of animals. Hybrots take this idea one step further, however, and place the "body" for the network of neurons in the real world as opposed to in a computational simulation. Instead of having a simulated environment with simulated sensory input and simulated motor output, The Dish is given a physical sensory system in the real world with real motor capabilities, and thus provides a physical simulation of *in vivo* processes. There are several robot "bodies" used in Lab D, some purchased off the shelf such as the "Koala," and others that are built from scratch in the lab such as a series of "dancing feathers." The predominant and most interesting hybrot model-system used by the Lab is a mechanical arm that is physically distributed in that part of it resides outside of the physical space of the lab. It is capable of drawing (primitive) pictures that goes by the name "MEArt."

MEArt was designed as both an art project and a research project with the goal of bringing the idea of a neural-mechanical hybrid system to the public. It is "*a geographically detached, bio-cybernetic research and development project exploring aspects of creativity and artistry in the age of new biological technologies from both artistic and scientific perspectives.*" As an exhibition, MEArt is touted as a "semi-living artist" that uses neurons to control a mechanical arm which draws scribbles on a piece of paper. As a research object, it is a closed-loop electrophysiology model-system which, unlike the animat, has a physical embodiment. The model-system is constructed as follows: a webcam is used as MEArt's eye, the information from the camera is run through a sensory translation program, the program stimulates The Dish, The Dish responds, those signals are run through a motor translator program, the translator issues motor commands to the arm, and then finally the arm moves and draws a line on a piece of paper.

The MEArt model-system serves as a model of an *in vivo* arm abstracted along a number of dimensions. For instance, the only kinds of information available to MEArt are positional information of its arm and sensory information from its camera. All other sensory dimensions and possible internal states have been abstracted away. Even so, the limited amount of uncertainty and complexity introduced by interaction with the real world along these two dimensions give the researchers a model that more closely approximates real neural input than animat model-systems. As one researcher expressed it,

> *"Yeah. I've done it in simulation. That's usually the first step... Just do it in simulation to get an idea... I guess one of the points is that once you stick something in the real world... and operating... you're adding things that you can't predict beforehand. And when like... Rodney Brooks did that with robots, with AI, he put AI into robots, he found out many different things that would not have been found out with just simulation.*"

There is a MEArt animat that the researchers use first to simulate the likely behavior of the arm. But then the hybrot is created to exist in the real world. The point of the prior computational simulation is to constrain the degrees of freedom tightly in their physical aspects in order to localize tests of the model. Once the general model is built from the animat, the hybrot model serves to explore the validity of the model in the real world.

Perhaps, then, the most interesting thing about the hybrots is that they blur the lines between the *in vitro* and *in vivo* worlds. Yes, they are clearly *in vitro* in the sense that the neurons exist outside their natural context and live in an artificial environment. However, the purpose of the hybrot model-system is to bring the neurons closer to their natural environment. The goal of the system as a whole is to more closely model the actual construction of an animal, and consequently the neural input. In this sense, one can view the MEArt model-system as a functioning entity in the *in vivo* world, as well as an *in vitro* model of embodied neuron learning. Through the construction of hybrots, Lab D tests not only whether or not they work merely as robots, but by creating physical models that more and more accurately approximate their understanding of a real animal, they physically explore and test their model of animal neural processing.

As with Lab A, in Lab D experimental situations models, too, tend to be in interlocking configurations, as standing in particular relations to other models. The diagram in Figure 2 is a schematic representation of our analysis of the MEArt model-system.

As can be seen in Figure 2, MEArt is a complex model-system involving several different physical and mental models (models shown with thick black outlines), instruments, and devices which converge and interlock around the MEArt hybrot. Like the vascular construct model system shown in Figure 1, here we trace the steps by which information is generated, manipulated, and propagated through the system. Instead, however, of tracing physical transformations, we trace the information transformations as they are interpreted through each of the instantiated models. Please keep in mind that even though the diagram details the MEArt system, a diagram for the Animat model-system would be identical, save for replacing "Real-World Environment" with "Simulated Environment", and "Sensory System" with "Simulated Sensory System." Since this is a closed loop system, there is no actual "start" point, but the best place to begin is with the sensory system (to the right of the figure). Here, raw sensory information about the world is gathered by a video camera and positional sensors for the arm. This information is then sent to the "Sensory to Activity Transform Program." Here, the raw data are transformed into stimulator commands which feed the information to The Dish. As mentioned earlier, this program serves as an *in silico* interlock of models. It is both an instantiation of the researcher's mental model of how sensory data are structured by an *in vivo* sensory system and an instantiation of the researcher's mental model of how information is interpreted and processed by The Dish itself. The program then directs a stimulator to inject a pattern of voltage changes into The Dish, giving it its "sensory" input. In response to the stimulation, The Dish

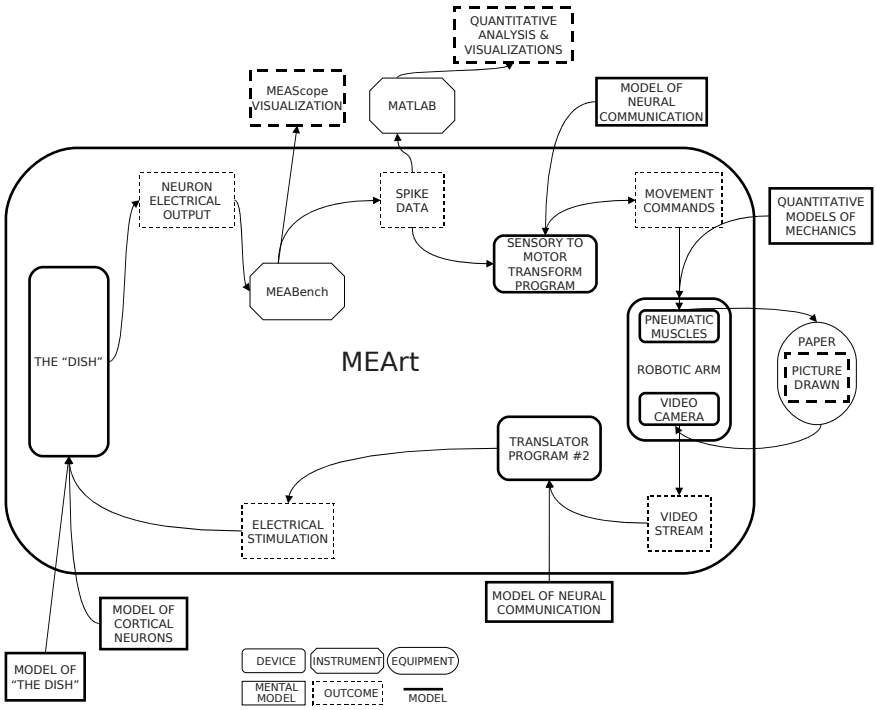

Figure 2. MEArt model system

creates a series of electrical impulses as it processes the information. This response is picked up by the data acquisition card, and is fed through the MEABench system. As previously mentioned, the raw data are subjected to a series of filters, each an instantiated model for processing neural activity. The subsequent spike data are then fed into a second "Activity to Motor Transform" program which transforms the data into commands which drive the motor output of the hybrot. Here again we have a convergence of neural communication models instantiated *in silico*, only this time the models are that of neural response and motor activity. Next, the physical hybrot (the model of an "animal") produces a physical change in the world by drawing on a sheet of paper. The drawing itself, while possibly aesthetically interesting, also serves as a representation (visualization) of the underlying processes that produced it. Finally, the physical changes in the environment are picked up by the sensory system, and the loop is closed.

Significantly, we can see a number of different levels of interlocking models in this model-system. First, at the lowest level there are the fundamental transformations that drive the hybrot. At the highest level, we have the MEArt system in its entirety, as a physically instantiated model of a real-world neural process-

ing system, and even a selective physical model of a real-world animal. Together, the researchers with the MEArt model system are able to make inferences about the relationship between plasticity (neural learning) and learning (animal behavior). As we can see here, the researchers effectively use MEArt to make inferences through the model system itself.

4 MODEL-BASED REASONING

Biomedical engineering is an interdisciplinary undertaking. The programmatic outcome of biomedical engineering is engineered artifacts built at least in part from biological components. In the cases we examined, these include an actual blood vessel substitute (as in tissue engineering Lab A) and a trained neuron culture for operating a robot (as in neural engineering Lab D). Notably, these are epistemically and ontologically *hybrid* objects, being simultaneously engineered and biological. These model-systems can be structural, behavioral, or functional (inclusive sense of 'or') analogues of real-world biological systems. To a high degree, they reflect the engineer's design perspective. The following passage from an interview with the Lab A Director is illustrative:

> LD: "*Well, it was clear to me from reading the literature that, and what was really motivating me by 1970-1971, was the fact that these characteristics of blood flow [mechanical forces] actually were influencing the biology of the wall of a blood vessel. And even more than that,* ***the way a blood vessel is designed***."
>
> Interviewer: "*So, this was influencing the characteristics of the biology*"
>
> LD: "*Yes, right, and influencing biological processes that were leading to disease.* ***The way a blood vessel is designed is****, it has an inner lining, called the endothelium. It's a monolayer; we call it a monolayer of cells because it's one cell thick. But it's the cell layer that is in direct contact with flowing blood. So, it made sense to me that if there was this influence of flow [mechanical forces] on the underlying biology of the vessel wall, that somehow that cell type had to be involved, the endothelium."* [stress added]

From the outset, he conceptualized the *in vivo* vessel from the design perspective of an engineer. He went on to note how difficult it was for many years to get biologists studying the cardiovascular system even to be interested in the possibility that mechanical forces play a role in disease.

Through processes of designing, constructing, manipulating, and modifying the *in vitro* model-systems engineers reason about, make hypotheses, and achieve understanding of *in vivo* biological phenomena. Developing an initial mental model of a blood vessel in terms of mechanical forces, for instance, led to the inference that the response of endothelial cells to the flow was implicated in disease processes.

Developing the physical models of mechanical forces and of blood vessels, has led, for instance, to inferences about gene expression under specific flow conditions.

In this section we address 1) the multidimensional nature of *interlocking models*; 2) the nature of model-based inference; and 3) how it is that *in vitro* simulations support inferences about *in vivo* phenomena.

4.1 Interlocking models

Model-systems comprise configurations of hybrid bio-engineered models. In the area of tissue engineering, for instance, some of these models refer to physiological aspects, which can include mental models of the structures or of the functions of the vasculature and physical models that interrelate structure and function at that level. Other models refer to the cellular or to the tissue level, which can again refer to the respective morphology or to aspects of the biochemistry at that level and to their interrelations. Still others point to mathematical models and engineering specifications that can serve to further constrain design options, experimentation or understandings of the biology under investigation.

Models are sites of serious, long-term investment of resources, and they define and delimit the research program. One device (an individual model) leads to creating new devices to form various model-systems affording potential experimental situations, such as happened following the introduction of the construct, which in turn led to the new bioreactor or The Dish, which led to building the animats and hybrots. As we can see by all our cases of model-systems, conducting experiments most often requires complex configurations of models. We attempt to capture these kinds of interconnections and more with the multidimensional notion of *interlocking models.* The devices serve as hubs for interlocking models along many dimensions. Models interlock biological and engineering concepts, methods, and materials (interdisciplinary melding). Models interlock in their design and construction (such as the flow loop modification because of the construct and the construct leading to the pulsatile bioreactor). Models interlock in experimental design. Mental and physical models interlock in model-based inference. Furthermore, models interlock research and learning, and serve as sites of cognitive and socio-cultural integration.

Two dimensions of interlocking merit special notice. First, in these engineering labs where design and redesign is an overarching agenda, a significant dimension of interlocking is historical: "history" is to be understood here as intellectually "hands on," that is meaningfully related to working with devices. Current devices interlock prior and future devices. Redesign of devices in the laboratory is embedded in an understanding of how a certain problem situation has led to the realization of certain design options. In other words, the agenda of redesign characteristic of an engineering research laboratory requires researchers appropriate some of its history as they go about their research. The current design is understood to be conditioned on the problem situation as it existed for the laboratory at a prior time, even if the situation is not fully known to the current researchers.

Thus, within the laboratory redesign is an *agenda,* and with it the historicity of the artifacts becomes a *resource* for novel design options.

Second, devices and model-systems are physical models that interlock not only with various community models but also are external representations that interlock with whatever mental representations are used by researchers in reasoning processes. As with other external representations such as diagrams and sketches, physical models constrain and provide affordances that augment the internal representations that participate in problem-solving processes. This is not the place to argue for the point, but Nersessian [2002; 2008] advances the hypothesis that the internal representational structures are also models — mental models (structural, behavioral, or functional analogs representations) — that afford simulative model-based reasoning. Given that the intended function of physical simulations with model-systems is epistemic, they are integral in the researcher's mental modeling processes.

4.2 Model-based inference

Here we outline the kinds of reasoning that can be carried out through model-systems. Characterizing a model, loosely, as a representation of a system with interactive parts with representations of those interactions, an instance of model-based reasoning: 1) involves the construction or retrieval of a model, 2) derives inferences through manipulation of the model, and 3) those inferences can be specific or generic, that is, they can either apply to the particular model or to the model understood as a model-type, representing members of a class of phenomena. The bio-engineered model-system is both a conceptual and an *in vitro* physical system representing the *in vivo* system under investigation. As such it is an abstraction — selective and schematic in nature — that represents dimensions of biological phenomena of interest to the researchers. Models are interpretations of target phenomena (e.g., forces within the human vascular system, learning in the brain) constructed to satisfy constraints drawn from the domain of the target problem (e.g., the biology and physics of the vascular system, the information processing carried out by the brain) and, often, one or more source domain (e.g. the flow loop's cardio-vascular and engineering domains, the dish's electro-physiology and engineering domains). Constraints include: spatial, temporal, topological, causal, material, categorical, logical, and mathematical. Performing a simulation can lead to new constraints — or to recognizing previously unnoticed constraints. Inferences made through simulations can provide new data and hypotheses that play a role in evaluating and revising models to comply with constraints or provide potential new understanding into *in vivo* phenomena.

For the model-systems such as we have been studying, models are structural, functional, or behavioral analogs of physical objects, entities, or processes used in experimental situations. Such models represent demonstratively (as opposed to descriptively). The primary evaluation relationship between the *in vitro* model and what it represents is goodness of fit, that is, what and how well does it cap-

ture the salient *in vivo* dimensions. Adequate models need to be of *the same kind* with respect to salient dimensions of target phenomena (often taking several iterations to achieve this objective). Inferences are derived through manipulations of the model. Operations on models require transformations consistent with the constraints of the domain. Importantly, the kinds of reasoning processes in model-based inference include, though are not limited to (not ordered): *abstraction* (limiting case, generic, approximation, idealization, generalization); *simulation* (inferring outcomes or new states via model manipulation (mental, physical, computational)); *evaluation* (goodness of fit, explanatory power, implications); and *adaptation* (constraint satisfaction, coherence, plausibility).

Selectivity in design of models enables bracketing of (potentially) irrelevant features and focuses attention on those relevant to the problem-solving context. Relevant constraints need to be determined for target and source domains. Abstractive and evaluative processes disregard irrelevant factors. Constructed models can instantiate irrelevant factors, but to reason correctly via a model requires recognizing these as scaffolding for the cognitively germane factors. As noted, different kinds of abstractive processes underlie selectivity and those that figure in model construction, in particular, include: idealization, approximation, limiting case, and generic abstraction. These provide different means of selectively focusing on features relevant to the problem solving while suppressing information that could inhibit the process. Suppression and selective highlighting of features provide ways of representing the problem in a cognitively tractable manner. Most importantly for conceptual innovation, they enable integration of information from different sources.

Idealization is a common strategy for relating mathematical representations to phenomena. Thinking of the sides of a triangle as having zero width, from the perspective of a geometrical figure, or the mass of a body as concentrated at a point, from the perspective of determining motion, allows strict application of mathematical formulae. Once mathematized, the idealization provides a point of departure from which to add in information about real-world phenomena as deemed relevant to a problem. Limiting case abstractions involve extrapolation or reduction to a minimum. Approximation provides a means of discounting the relevance of differences. A standard approximation in physics is the "first-order approximation" used when applying a mathematical representation, as with the force of the flow modeled by the flow loop device. Basically it makes the assumption that any higher-order effects are likely to be irrelevant or to be so complex as to make the analysis intractable. A laminar flow — one without currents or eddies — provides a first-order approximation sufficient for solving problems pertaining to many fluid dynamical phenomena. Some abstractions are deliberate, as we saw in choice to use steady laminar flow rather than pulsatile and in the approximation that flow is 2-D and over the flat surface of constructs. Some abstractions occur because the devices are also systems themselves; possessing engineering constraints that often require simplification and idealization not deriving specifically from the biological system they are modeling, as we saw with design and construction of The Dish. As

mentioned before, The Dish only provides a mono-layer network of neurons instead of the rich 3D connections found *in vivo*. Here, the researchers have reduced the complexity of the dish to a minimum along one physical dimension. The choice to use only mono-layer networks overcomes two hurdles. First, the recording technology is limited to a mono-layer 8x8 grid embedded in the bottom of the dish. Any neurons that were not in the bottom layer of the dish would not be accessible for recording. Also, from the perspective of data analysis, the mono-layer provides a reasonable reduction of information, while still maintaining the salient qualities of inter-neuron communication. This limiting case makes the interpretation of the electrical response tractable, both the physical and informational. But then, the researchers need always to be aware of how the abstractions might influence the other models with which they interlock in experimental situations.

Although abstractive processes often occur in tandem, differentiating them serves to call attention to a kind of abstraction that is especially productive in merging constraints from multiple sources. For instance, in considering the behavior of a physical system, such as a spring, scientists often draw a specific instance, but then reason with it as without specificity as to the number and width of coils. To reason about it as a generic simple harmonic oscillator requires, further, suppressing features of its spring-like structure and behaviors. We call this process *abstraction via generic modeling* or, simply, generic abstraction. In model-based reasoning, constraints drawn from different sources need to be understood at a level of generality sufficient for retrieval, integration, and transfer. Further, generic abstraction gives generality to inferences made from the specific models that are constructed. As Berkeley noted, one cannot imagine a *triangle-in-general* but only some specific instance. However, in considering what it has in common with all triangles, we are able to view the specific triangle as lacking specificity in the lengths of the sides and degrees of the angles.

The same concrete representation can be interpreted and understood as generic or specific depending on the demands of the reasoning context. Researchers are able to understand which inferences made from the behavior of a construct apply to blood vessels *in vivo* and what apply to the specific *in vitro* construct model. In the experiment within the vascular construct model-system outlined in Figure 1, it is possible to identify inferences that are specific to the flow loop and construct models and inferences made from them that apply to *in vivo* vascular systems, generally. One is as follows. The hypothesis of the experiment is that conditioning the constructs by first subjecting them to the kinds of forces experienced *in vivo* will prevent platelet formation (thus reducing the risk of thrombosis). Inferences about the responses of the endothelial cells lining the construct to the forces exerted on them by the fluid flowing in the flow loop are generic to members of a class of cardiovascular systems (to a first-order approximation). But inferences about the strength of the construct based on the collagen scaffold are specific to that model. Another example comes from the physical structure of the network of neurons plated on each dish. The current process for plating neurons onto the MEA calls for breaking apart the connections between the neurons, and placing

the free-floating cells in a single layer on the bottom of the dish. The cells are then left to re-grow connections to each other. Clearly, the structure of the resulting network has no specific relation to the original cortex. In effect, The Dish is a specific instantiation of a network of neurons that can be used by researchers as and reasoned with as a generic network. In framing The Dish in this manner, it allows the researchers to infer network processing characteristics of a specific network and transfer their findings to other cortical networks.

In sum, the kinds of abstractive processes we have been considering provide ways of selectively representing and reasoning with models. Idealization, limiting case, approximation, and generic abstraction provide a means of representing and integrating constraints stemming from different sources into models. To reason correctly by means of the models requires recognizing which features are germane and which are cognitively inert within that problem-solving context. This point requires consideration of how *in vitro* models represent *in vivo* phenomena.

4.3 Parallelism and exemplification

In vitro models are virtual worlds which bio-engineers design and construct to selectively "*parallel*" or "*mimic*" or "*approximate*" or "*simulate*" (to use their words) the *in vivo* phenomena. In some cases, as with the mechanical forces in the blood vessel and the morphology and alignment of endothelial cells, the phenomena are well-understood. In other cases, the models are the only means they have from which to determine the nature of the *in vivo* phenomena. They are hypotheses based on incomplete understanding, such as The Dish as a model of learning, and the objective is to develop understanding of the brain or the endothelial cells from the parallel world of the model-system. The physical and computational models are designed and created with the intent of performing simulations with them. Model-systems are sites of systematic experimentation and the nature of the parallelism between *in vivo* phenomena and their corresponding model-systems is continually evolving, if only in minor ways, as they become better or different kinds of approximations. Thus parallelism is an historical process, both because of the overarching redesign agenda and because of the fact the model-systems are not fast-lived set ups but painstakingly constructed sites of serious investment over considerable time spans.

What is the nature of the parallelism or mimicry, that is, of the relationships between the models and what they represent, especially in relation to their status as epistemic artifacts? The goal in creating these virtual worlds is for the models to be of the *same kind* as the real-world phenomena along particular dimensions, such as first-order processes of blood flow in arteries or type of information processing in the brain. In such respects, the referential relations are built into the design and building of the models. Several accounts of models propose that "similarity" is the basis of the representational relation. Models are similar to and different from what they model along relevant aspects [Cartwright, 1983; Giere, 1988; Nersessian, 1992; 2002]. In the limit, models are isomorphic to relevant

aspects of the phenomena they model; indeed, if they are successful with the tissue engineered vascular wall model it will be identical to an artery in structure, behavior, and function — that is, itself an artery. In considering how they represent, however, the pertinent word here is "*relevant.*" Not just any similarities and differences matter. For the model-systems, the relevant aspects are those which *exemplify* the phenomena being investigated. As introduced by Nelson Goodman in *Languages of Art* [Goodman, 1968] a representation exemplifies certain features or properties if it "both is and refers to" something which has those feature or properties; that is "[e]xemplification is possession plus reference [p. 53]. One of his examples is that of a tailor's fabric swatch, which "is a sample of color weave, texture, and pattern; but not of size, shape, or absolute weight or value" [*ibid.*]. In our cases, the notion of exemplification captures the idea that a model needs to be of the same kind along relevant aspects. The relevant aspects are those which exemplify the objects, entities, or processes under investigation. For instance, the hybrot not only is a model-system used to represent neural processing *in vitro*, but it is also a system that is, in actuality, doing neural processing. It both refers to selected processing in a brain and is processing in a brain, simultaneously, in the context of the research in Lab D.

Catherine Elgin [1996; 2004] has built on Goodman's notion to address the epistemic problem that science makes extensive use of practices that, if we insist on equating truth and epistemic acceptability, lead clearly to falsehoods, such as limiting case demonstrations, idealizations, curve smoothing, and computational modeling. Yet science accepts the outcomes of such practices as advancing scientific goals, for example, deriving a mathematical representation or predicting future real-world states on the basis of a computational simulation. We can conclude from this either that science is cognitively defective or allow that scientific understanding "is often couched in and conveyed by symbols that are not and do not purport to be true" [Elgin, 2004, p. 116]. She advocates that the epistemic value of modeling, as well as other strategies and symbols (propositional and iconic) used by scientists, lies in their ability to exemplify that is, to "at once refer to and instantiate a feature" [Elgin, 1996, p. 171]. Models thus afford the researcher with epistemic access to selected features of phenomena. As exemplifications, models are devised so as to focus attention on features relevant to epistemic goals. As Elgin points out, physical exemplifications are routinely created for experimental purposes such as when a lump of ore is refined to consist only of iron. The physical abstraction, then, affords epistemic access to the properties of iron in the context of real world experimentation.

The bio-engineers we have been studying design and create hybrid exemplifications paralleling the *in vivo* phenomena they wish to study to the degree of specificity that they believe sufficient or to the degree that lack of knowledge of the phenomena or the nature of the design materials constrains them. The models abstract away irrelevant (or potentially relevant) features, thereby focusing attention on those features salient to the problem-solving context. Understanding what inferences can be made in general about the behavior of endothelial cells in the

artery or neurons in the brain on the basis of a simulation through a model-system derives not from making inductive generalizations but through understanding how the one model exemplifies selected features of all such phenomena. A properly designed model-system warrants the researchers in pursing where the experimental outcomes of the *in vitro* world might lead, in lieu of being able to carry these out in the *in vivo* world.

Finally, drawing on the notion of exemplification to explicate how constructed models are selective representations serves as a reminder of the deeply socio-cultural nature of representation. As Goodman observed, what a model exemplifies depends on goals, purposes, and context. A paint chip usually exemplifies a color, but might in certain circumstances be taken to exemplify a geometrical shape, such as a rectangle. It exemplifies color within a particular set of social norms surrounding the practice of picking paint for one's walls or house. In their model-based reasoning practices, the researchers in Lab A and Lab D are drawing on a repertoire of representational practices and the conventions of specific communities surrounding these.

5 CONCLUSION

We have discussed the nature of the model-based reasoning carried out by means of constructing and manipulating physical simulation models in the interdisciplinary context of biomedical engineering. In reasoning processes, models have two faces: mental and physical. That is, inference involves co-constructing and manipulating physical and mental models. The physical models are hybrid designs that merge biological and engineering constraints and represent selective understandings. The mental models, too, interlock biological and engineering concepts and understandings. Problem solving through simulation with model-systems is an epistemic activity that enables inference through selectively creating objects, situations, events, and processes that exemplify those of interest in selective ways. It is also a socio-cultural activity in that, even when considering reasoning activities, the models and the practices of using them are both cognitive and cultural achievements. So, though the model-systems as represented in Figures 1 and 2 convey only a pared-down representations, the model-systems and experiments are to be understood as embedded in rich cognitive-cultural systems distributed in space and time, themselves designed to enable and support such experimentation. Experimentation and inference are conditioned on a fabric of interlocking models — across space, time, people, and artifacts, connecting mental and physical representations. Our meta-analysis has separated some of the threads, but in practice they are intricately woven and inseparable.

Devices and model-systems are what socio-cultural studies of science refer to as the "material culture" of the community, but they also function as what cognitive studies of science refer to as "cognitive artifacts" participating in the reasoning, representational, and problem-solving processes of a distributed system. Our point is that within the research of the laboratories, they are both, and it is not possible

to fathom how they produce knowledge claims by focusing exclusively on one or the other aspect. They are representations of current understandings and thus play a role in model-based reasoning; they are central to social practices related to community membership; they are sites of learning; they provide ties that bind one generation of researchers (around five years) to another; they perform as cultural "ratchets" that enable one generation to build upon the results of the previous, and thus move the problem solving forward. In sum, they are hubs that interlock the various dimensions of the cognitive-cultural fabric in which problem solving takes place in these interdisciplinary engineering research laboratories.

ACKNOWLEDGEMENTS

We would like to thank other members of our group who have contributed to this research, especially Wendy Newstetter (co-PI), Lisa Osbeck, Elke Kurz-Milcke, Ellie Harmon, and Arvind Venkataramani. We thank the members of Labs A and D for allowing us into their work environment, granting us numerous interviews, and sharing their knowledge with us. We gratefully acknowledge the support of the National Science Foundation ROLE Grants REC0106773 and DRL0411825, and of the Radcliffe Institute for Advanced Study and the National Endowment for the Humanities (Nersessian).

BIBLIOGRAPHY

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

[Elgin, 1996] C. Z. Elgin. *Considered Judgment*. Princeton: Princeton University Press, 1996.

[Elgin, 2004] C. Z. Elgin. "True Enough." *Philosophical Issues* 14 (2004): 113-31.

[Giere, 1988] R. N. Giere. *Explaining Science: A Cognitive Approach*. Chicago: University of Chicago Press, 1988.

[Nersessian, 1992] N. J. Nersessian. "How Do Scientists Think? Capturing the Dynamics of Conceptual Change in Science." In *Minnesota Studies in the Philosophy of Science*, edited by R. Giere, 3-45. Minneapolis: University of Minnesota Press, 1992.

[Nersessian, 2002] N. J. Nersessian. "The Cognitive Basis of Model-Based Reasoning in Science." In *The Cognitive Basis of Science*, edited by P. Carruthers, S. Stich and M. Siegal, 133-53. Cambridge: Cambridge University Press, 2002.

[Nersessian, 2005] N. J. Nersessian. "Interpreting Scientific and Engineering Practices: Integrating the Cognitive, Social, and Cultural Dimensions." In *Scientific and Technological Thinking*, edited by M. Gorman, R. D. Tweney, D. Gooding and A. Kincannon, 17-56. Hillsdale, N. J.: Lawrence Erlbaum, 2005.

[Nersessian, 2008] N. J. Nersessian. *Creating Scientific Concepts*. Cambridge, MA: MIT Press, 2008.