Proceedings of the Cognitive Science Society 33, 2011

Building Cognition: The Construction of External Representations for Discovery

Sanjay Chandrasekharan (sanjayan@cc.gatech.edu)
Nancy J. Nersessian (nancyn@cc.gatech.edu)
School of Interactive Computing, Georgia Institute of Technology,
85 Fifth Street NW, Atlanta, GA 30308, USA

Abstract

The analysis of the cognitive role played by external representations – particularly within the distributed cognition (DC) framework – has focused on the *use* of such representations in cognitive tasks. In this paper, we argue that the processes of *building* such representations require close attention as well, especially when extending the DC framework to ill-structured domains such as scientific laboratories, where building novel representations is crucial for making discoveries. Based on an ethnographic study of the building of computational models in a systems biology laboratory, we examine the complex cognitive roles played by the external representations built by the lab (pathway diagrams and models), and the building process itself.

Keywords: Distributed Cognition, External Representations, Scientific Cognition, Discovery, Creativity, Ethnography

Traveler, there is no path,
paths are made by walking.
– Antonio Machado

The role of external representations in cognitive tasks has received a lot of attention, particularly within the distributed cognition (DC) framework (Hutchins, 1995; 1995a; Kirsh, 2010). Much of the work on external representations within DC focuses on capturing detailed descriptions of the way external representations are used in highly structured task environments, such as ship navigation and landing of aircraft, and the way these representations change the nature of cognitive tasks. Less understood are the processes of building external representations to alter task environments (Kirsh, 1996; Chandrasekharan & Stewart, 2007), and the role played by this building process in cognition and problem-solving. In this paper, we focus on the building of a complex external representation – a computational model – and examine the role this external representation, and its building process, plays in structuring, as well as altering, the discovery task in a systems biology laboratory.

Only a handful of studies have examined problem solving in scientific research from a DC perspective (e.g, Nersessian et al, 2003; Alac & Hutchins, 2004; Hall, Wieckert & Wright, 2010; Goodwin, 1997; Giere, 2002). These studies do not consider how external representations are built, largely because the development of a novel external representation, and the changes it makes to the scientific task environment, are complex events that occur over long periods of time, and therefore not easily captured. Even when such data are available (Chandrasekharan, 2009; Nersessian & Chandrasekharan, 2009), it is not easy to understand building using the current DC framework, which is derived using studies of well-structured tasks, and therefore do not transfer well to the ill-structured and open-ended task environment of a scientific laboratory. Further, the DC framework, as it stands now, focuses on the *use* of existing external representations, not on the processes of generating representations, which play a significant role in scientific practice. Building representations is part of the activity of what Hall et al. (2010) call "distributing" cognition, that is, "how cognition … is produced … out of human activity (p.2)." The DC framework therefore needs to be extended to understand scientific practices that require building novel representations for problem solving. In this paper, we outline some aspects of this extension, using a case study of how computational models are built in a system biology lab.

Lab G as a Distributed Cognitive System

In our current project we are studying problem-solving practices in two integrative systems biology labs. We focus here on one lab that does only computational modeling ("Lab G"). The modelers come mainly from engineering fields, but work on building computational models of biochemical pathways, to simulate and understand phenomena as varied as Parkinson's disease, plant systems for bio-fuels, and atherosclerosis. The problems Lab G modelers work on are provided by outside experimental collaborators, who see modeling primarily as a method for identifying key experiments of scientific or commercial importance. The collaborators provide experimental data for modeling, and also generate data needed for developing or validating the model. In broad terms, the Lab G modeling processes can be understood as occurring within a distributed socio-technical system, which is the primary unit of analysis in DC. This system comprises people working together (modelers, experimentalists) to accomplish a task (discover fruitful changes to biological pathways), and the artifacts they use (models, pathways, diagrams, graphs, papers, databases, search engines) in the process.

The task environment of the lab, and the external representations used there, differ drastically from those usually examined in DC, such as the standard example of the cockpit and the speed-bug (Hutchins, 1995a). The main differences can be classified as follows:

Actors and Goals: The lab does not have a structured task environment, with synchronous actions connecting

individuals or groups. The objective of the lab is to make discoveries, so the lab task environment is one where the *specific goal* is not known in advance. There are very general goals, such as "discover interesting reactions", and less general goals, such as "fit model". These general goals are spread across people who share a resource (experimental data), but do not share a tightly integrated task environment. The actors have different goals; they work in different settings, at different times, and using different instruments.

Conflicts: The community sharing the data has conflicting interests. Even though the modelers are trying to help the experimentalists, it is very hard to get data from experimentalists. One reason is that the experimental labs have their own projects, and the modeler's requests are not a high priority for lab members. The two communities also work at conflicting time-scales. For instance, once developed, the models run blazingly fast, and can produce interesting predictions in a few days. But experimenters take weeks and months to gather data based on these predictions, and this 'phase-lag' frustrates the modeler. Conflicts also arise over the different levels of control afforded by models (more control) and experimental techniques (less control).

Artifacts: The lab researchers do not simply *use* external representations to reach a goal. The task of the lab is to *build* representations (pathways, models) and use them to make discoveries. These representations are themselves built from other representations (papers, data files, online databases, code), which provide information in a scattered fashion. There are significant judgments involved in assessing this scattered information (Is this database curated? Is this cell line compatible with my problem?), and integrating the information into a coherent representation (Should this reaction be included in my pathway? Are there other regulations missing here?). The engineers involved in building the models are novices in making these judgments, and they gain knowledge by discussing these judgments with the experimentalists, who, in turn, are not clear on how the model works and what the modeler is doing.

These differences suggest that understanding the lab as a distributed cognitive system requires extending the current DC framework – to task environments where goals are not clearly specified, where many kinds of conflicts exist, and where building representations is the central component of the task. Such an extension requires an understanding of the cognitive roles played by external representations in such environments, and how the features of these representations meet the demands of the task. First we consider why systems biology requires building models, and provide a brief outline of the lab G building processes.

The Need for Building Models

Computational models play a complex set of roles in the process of discovery, and very little is known about the cognitive mechanisms that underlie discovery based on such models (Nersessian & Chandrasekharan, 2009; Chandrasekharan, 2009). Recently, a combination of four factors has made the practice of building models more widespread, particularly in the bio-sciences and engineering.

1) The complex, non-linear, and dynamic nature of the problems investigated in contemporary biology (and recent science in general), requires building such models. This is because it is almost impossible to develop detailed conceptual models of cellular and molecular-level interactions in your head, or using pencil and paper, as these processes involve many levels of structure, occur simultaneously, across different time scales, and with complex feedback loops.
2) Massive amounts of data are now generated by experimental work in many areas of biology, such as high-throughput data in genetics, where the interactions among different variables are extremely complex, and cannot be understood without modeling. Further, the technology that generates the data relies heavily on embedded statistical models of the distribution of the data. These data usually require complex visualizations based on models, as they are difficult to represent, and comprehend, using traditional structures such as graphs.
3) Data in biology are closely tied to their context (e.g., specific cell lines, animals, diseases), and there is no theory that helps structure all these disparate and scattered data. Building computational models helps bring these data together in a structured fashion.
4) The development and easy availability of new technology that supports modeling and rapid prototyping has made modeling more widespread.

These factors, together with the technological resource environment of contemporary science, are driving a rapid expansion in the practice of building models. Earlier resource environments, where the only cognitive tools available were pencil and paper and the brain, saw the development of methods such as thought experiments (Nersessian, 1991), and cognitive walk-through simulations based on models drawn on paper (Nersessian, 2008). These methods have served science well, but they required idealizing the problems to a high degree, as both the cognitive and data-collection resources were limited. Finer methods and representational modes are needed to provide insight into the complex, dynamic and non-linear phenomena investigated by contemporary science, where massive amounts of data are available, and the details are critical, so idealizing them away is not an option.

Constructing the Pathway and the Model

Lab G researchers mostly build ordinary differential equation (ODE) models of metabolic systems, which capture how the concentration levels of different metabolites in a given biological pathway change over time. The first step in this building process is the development of a pathway diagram, which shows the main reactions involved. The pathway diagram also captures positive and negative regulation effects, which specify how the presence of different metabolites has a positive or negative influence on

different reactions (Figure 1). A rough diagram of the pathway is usually provided by the experimental collaborators. But the modelers, who mostly come from engineering backgrounds, have to estimate the details of the pathway by themselves, particularly values of parameters

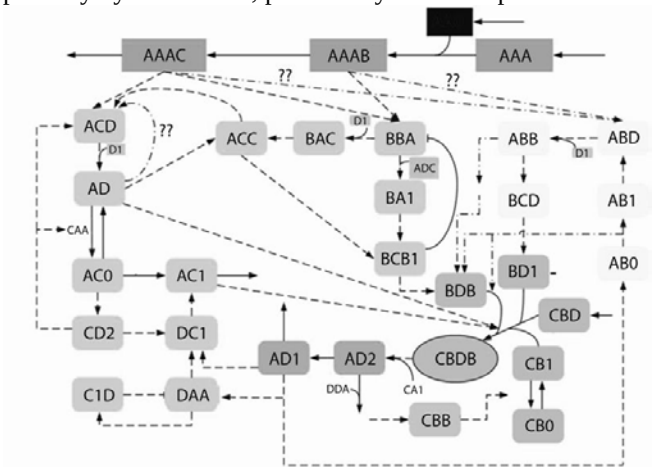

Figure 1: A sample pathway diagram. Metabolite names are replaced with alphabets. The dark lines indicate connections where material moves across nodes, the dotted lines indicate regulatory connections. Note the question marks over some connections that are postulated by the modeler.

related to metabolites, such as speed of change (kinetic order) and concentration level (rate constant), which are usually not measured by experimenters. Some of this information is available in rough form (with varying degrees of reliability) from online databases, but most often these values need to be estimated, usually through iterative testing of the model, using a range of numbers as parameter values.

Modelers also add some components to the pathway, usually metabolites that are known to interact with the network provided by the experimenters. These components are found by reading and searching biology journal articles and databases related to the problem being modeled, and also based on results from preliminary models.

Even though much of the pathway is provided by experimentalists, these kinds of additions based on literature searches are required, because the provided pathway does not identify all the components, and the regulatory influences they have on the reaction.

The pathway developed by the modeler thus brings together pieces of information that are spread over a wide set of papers, databases, and unreported data from the experimentalists. This pathway is usually trimmed, based on some simplifying assumptions, mostly to lower the mathematical and computational complexity involved in numerically solving the differential equations. After the trimming, differential equations are generated to capture the trimmed pathway. A variable is used to represent the metabolite, while the speed of its change (kinetic order) and its concentration level (rate constant) are represented by parameters, which can take many values. The next step involves estimating values for these parameters, and these values are then used to initialize simulations of the models. The simulation results are then compared to actual experimental results, to judge the 'fit' of the model.

Usually, modelers split available experimental data into two, one set is used to develop the model (training data), and the other set is used to validate/test the completed model (test data). When the model data do not fit the test data, the parameters are "tuned" to get model results that fit. Once the model fits the test data, it is run through a series of diagnostic tests, such as stability (e.g. does not crash for a range of values), sensitivity (e.g. input is proportional to output) and consistency (e.g. reactant material is not lost or added). If the diagnostic tests fail, the parameters are tuned again, and in some cases, the pathway changed, until the model meets both the fit and diagnostic tests. Figure 2 provides a broad outline of the modeling process.

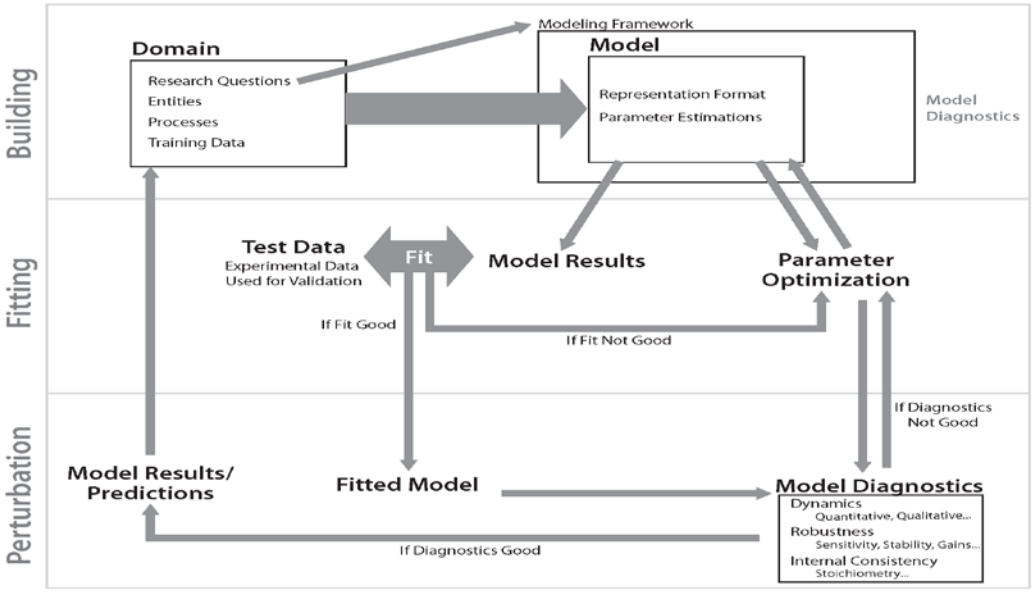

Figure 2: An outline of the modeling process in Lab G. Note that the 'building' phase incorporates both 'fitting' (of the training data) and 'perturbation' (model diagnostics) elements.

Lab G models do not have real-time dynamic visualizations. Parameter values are changed manually or using scripts. Results for different parameter values are compared using a deck of graphs, where each graph plots the concentration value of a molecule in the pathway across time. These graphs are used by the modeler while discussing the model with collaborators and other team members. A significant chunk of the parameter estimation problem is tackled using optimization algorithms (such as simulated annealing and genetic algorithms), which automatically do the 'tuning' of parameters, by comparing the output values (for different parameter inputs) against a desired value.

Importantly, the linear work flow suggested by the above description is very deceptive – the modeling process is highly iterative. For instance, to develop the pathway diagram, preliminary models are built using the network provided by the experimenters, and these are run using tentative parameter values, and the generated model data are fit to the training data. The parameter values are then revised based on this fit. If the model data do not fit after a large number of these parameter revisions – particularly if the data trends are the exact opposite of experimental data – the modeler will add some components to the network, based on elements that are known (in the literature) to be related to the pathway. There is also an instance where the modeling led to the discovery that an element (named X) in another pathway was influencing a bio-fuel pathway. Later experimental work by collaborators identified a candidate element for X. Thus, *some of the model's components – elements and values – are set by building and running the model itself.* These pathway revisions, and their justifications, are discussed with the collaborators, and if a revision is considered "reasonable" by the experimenter, it becomes a stable component of the pathway. This pathway identification process is usually "bottom-up," and creates a "composite" network, made up of parameter values and metabolites extracted from experiments in different species, different cell lines etc. This composite is usually unique, and does not exist anywhere else in the literature.

One of central problems the lab members face is the unavailability of rich, and dependable, data. In modeling, data are used for many purposes. One central use of data is to establish that the model captures a possible biological mechanism, and this is done by showing that the model's output matches the output from experiments (fitting data). A second use of data is to tune parameter values during the training phase of building the model. The fit with the experimental data from each training simulation can indicate how the model parameters need to be changed, to generate model data that fit the training data. This use is highly dependent on the type of data available. Most of the time, the available data are 'qualitative' in nature – usually how an experimental manipulation led to a change in a metabolite level from a baseline. Mostly, this is reported as a single data point, indicating the level going up or down, and then holding steady. However, when this type of "steady-state" data fits the results of the model, this fit does not indicate that the model has captured the biological mechanism. A range of parameter values can generate model results that fit such sparse data – the fit is not unique. Further, since the pathway is an approximation, the modeler is uncertain as to whether the lack of a unique and accurate solution is due to poor estimation of parameters, or because some elements are missing from her pathway.

As an example instance of modeling in this lab, consider G12, an electrical engineer by training, who is modeling atherosclerosis. When she started modeling, she had no background on atherosclerosis. She was provided a rough outline of the pathway by her experimental collaborators, and she learned more about the pathway by reading papers.

The initial papers were from the collaborating lab, but then she spread out using the reference lists of those papers. The data available were mostly steady-state data. Once she had read a number of papers, she started building rudimentary computer models and testing these using available data. She then added some components to the model based on connections in the literature. It is worth noting here that while her problem mostly concerned endothelial cells, some of her parameters were taken from experiments with neurons, a very different cell class, and a domain of research (neuroscience) that is not usually connected to research in endothelial cells. After discussion, her collaborators endorsed some of her additions, as "reasonable".

Estimating parameter values for her model was a tough problem, since the data were sparse. To get parameter values that generated model data that fit the training data, she ran a large number of simulations and compared the results with the training data. Finally, she arrived at a set of values that generated data that *roughly* matched the training data. Once this was done, she tested her model against the test data, and got a rough fit there as well. Based on this fit, she generated a number of predictions from the model, by changing the parameter values. Some of these predictions would be tested by her experimental collaborators.

This exemplar is representative of much of the modeling in this lab, where external representations (pathways and models) are built up from scattered and unreliable information, and discussion. These representations are built by modelers (engineers with no background in biology) using an iterative building strategy, based on rough data and guidelines from domain experts. This building process leads to closer collaboration between the modelers and the experimentalists. The completed model's predictions guide experimental decisions, and potential discoveries in critical areas such as biofuel production. The data from these experiments are then incorporated into the model, leading to another cycle of experiments and discoveries.

In the following section we examine some of the different cognitive roles played by pathway diagrams and models.

Cognitive Roles of Pathways & Models

Saliency: Kirsh (1995) argues that counting small dots like these (……) using a pencil is easier than counting them with your eye, because the pencil changes the dot-counting task to counting of pencil movements, which are more salient. Further, the pencil movement prevents counting of dots previously counted. The model serves a similar cognitive role, since the model variables and their changes are more salient than the actual reaction variables, and they can be tracked separately as the reaction proceeds through time. Further, the model allows grouping different reaction components, and running the groups separately, and then bringing the different groups together and running this integrated model. This is not possible with actual reactions.

Integration: The pathway is developed by combining inputs from many sources, including rough outlines from experimenters, related papers, and databases. This construction process creates a composite structure that brings together information from disparate literatures. Each element of the pathway exists in a paper somewhere, but the composite structure created by the modeler exists only in her work. This composite structure plays a book-keeping function, bringing together information that is spread across papers and domains that may not be related as research streams. The model built using such a pathway could be seen as a "running" literature review, and the model's correct predictions provide an external, *global*, validation for the experiment results on the pathway elements.

Linking: As we saw in G12 case, the information used to build the pathway/model can be from disparate domains of research (neurons, endothelial cells). Other work in the lab connects biophysics and neuroscience, and also tribology (the study of interacting surfaces in relative motion), paper science, and plant cell-wall growth. Such composite models connect disparate literatures, and searches on their keywords provide entry points (Kirsh, 2001) into systems biology for people with different backgrounds. Over time, the model also links together a range of results in a domain, and thus prevents the dissipation of data and concepts. This 'ratchet effect' also allows novices in the lab to start at a more complex level than if they start from scratch.

Mangrove: Once built, the model generates predictions that are used by experimental collaborators to develop new experiments, and the results from these experiments are fed back into the model. This process, over time, creates a new collaboration space, and brings the modeling and experimental communities together, leading to a research domain that is distinct from the backgrounds of researchers in both the streams. The building process also leads to new shared mental representations. For instance, each reaction occurs in a specific location in the cell (nucleus, organelles, cytoplasm), and every reaction is determined by the structural properties of the molecules involved. The experimentalist's judgments are based on this spatial complexity. But the ODE models do not take into account any of this spatial complexity, and the modeler with an engineering background is largely unaware of this complexity when she starts. Over time, the building of the model, and the discussion with experimentalists about possible additions, leads to her developing more awareness about the spatial complexity, and sometimes new modeling strategies that take into account this complexity. In the other direction, discussions about the mathematical advantages provided by time-series data could influence experimentalists to report data across time, even if the results are not statistically significant. The building process thus leads to overlapping problem representations, and approaches that fit the other community's task better.

We term this growth over time of shared collaboration space and mental representations the "mangrove function" of external representations, after Clark's (1997) example of the growth of a mangrove tree to illustrate how writing can generate thought. A mangrove tree germinates from a seed floating in shallow water. It then sends out a complex web

of roots to the ground, creating a "plant on stilts." This structure traps floating debris, and over time, sand accumulates around this debris, creating a little island around the plant. The tree thus generates its own land to grow. This is similar to how the building of the model generates its own task environment, collaboration space, research domain, and shared representations.

Stop-and-Poke: One of the central reasons for building a model is to have a more controlled environment, where almost any variable can be controlled in highly accurate and specific ways. The model also allows many variables to be changed at the same time, which is not possible with current experimental techniques. Further, the model output can be tracked visually over time, and this tracking can instantly suggest the type of change that needs to be made to variables to get a desired output. Also, the model can be stopped at any time-point, and the state of the different variables can be examined at that point. This ability to stop-and-poke the simulation is crucial for understanding the dynamic interplay among different components. It allows identification of global patterns in variable states, and their relation to the output. Over time, comparisons of such patterns lead to identification of reliable mechanisms. In contrast, experimental states cannot be stopped in between and each variable examined in detail.

Coagulation: Models are built by systematically replicating experimental data. Each replication adds complexity to the model, until the model "fits" all available experimental data well. At this point the model can "enact" the behavior of the pathway that is being examined, and can be used to make predictions, where variables are changed in ways that generate desirable results. The notion of fit is complex, as it is not a data point-by-point replication of all experimental data for all variables. Rather, 'fit' usually means the model replicates the *trends* (metabolite production going up/down) in the experimental data, for *most* of the major variables. In other words, fit is a global pattern, and it is approximate. While estimating unknown parameters, the modeler uses the fit with the experimental data as an anchor. For each change in a parameter, the way the model's output map to the experimental results (the "fit landscape") changes. But only parameter values that improve fit, or keep fit at an acceptable level, are considered. The building process proceeds by using the global behavior of the model (fit) as an anchor to specify the local structure (parameter values), which are involved in generating the fit itself. Since the fit is also used to add/delete components in the pathway, the model-building process can be thought of as a coagulation process, where each of the elements (pathway-structure, parameters, fit) are fluid in the beginning, but get more and more constrained by their interactions. This process is very complex and almost impossible without building the model.

Mutation: The model-building process begins by capturing a reaction using variables, and then proceeds by identifying ideal combinations of numbers for the variables – combinations that generate data close to experimental data. Variables are a way of getting the building process going by representing the unknown using place-holders. But this representation has an interesting side effect. The variable representation provides the modeler with a more flexible way of thinking about the reaction, compared to the experimentalist, who works only with one set of values (the experimental results), which are privileged values, arising from a set of spatial/structural properties of the molecules. For the modeler, the variables can take any set of values, as long as they generate a fit with experimental data. The variable representation allows the modeler to *naturally* think of the experimental value as *one possible scenario*, and also examine why this scenario is common. This allows her to naturally think of broader design patterns, and principles, that generate the natural order, such as thermodynamic principles. This is an ongoing effort in the lab. The "variable thinking" also supports the modeler's objective of altering the structure of the reaction, in a way that patterns commonly seen in nature (such as the thickness of lignin in plant cell walls) can be redesigned. This objective requires: 1) not fixating on the given natural order, and 2) thinking of design principles underlying this natural order. The variable representation facilitates both these cognitive steps.

Conclusion

The study of scientific laboratories as distributed cognitive systems is in its infancy. We believe that such analyses could provide insights into how discovery and innovation happens in science and engineering. But this analysis requires moving away from directly applying the DC framework as it exists to such discovery environments. Instead, we need to extend the DC framework in new ways, particularly taking into account ill-structured task environments, and the way building new representations changes the task environment. In this paper, we have outlined some ways of understanding this process, using the building of a computational model in a system biology lab.

Acknowledgments

We gratefully acknowledge the support of National Science Foundation REESE grant DRL097394084. Our analysis has benefited from discussions with co-PI Wendy Newstetter, Lisa Osbeck and Vrishali Subramanian. We thank the members of Lab G for granting us interviews. Thanks to Joshua Aurigemma for fine-tuning the modeling graphic.

References

Alac, M. & Hutchins, E. (2004). I See What You are Saying: Action as Cognition in fMRI Brain Mapping Practice, *Journal of Cognition and Culture*, 4, 629-661.

Chandrasekharan, S., & Stewart, T. (2007). The origin of epistemic structures and proto-representations. *Adaptive Behavior*, 15(3), 329–353.

Chandrasekharan, S. (2009). Building to Discover: A Common Coding Model. *Cognitive Science*, 33(6), 1059-1086.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.

Giere, R. (2002). Scientific Cognition as Distributed Cognition. In P. Carruthers, S. Stitch and M. Siegal, (Eds.) *The Cognitive Bases of Science*, Cambridge: Cambridge University Press

Goodwin, C. (1997). The Blackness of Black: Color Categories as Situated Practice. In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, Tools and Reasoning: Essays on Situated Cognition* (pp. 111-140). Berlin, Heidelberg, New York: Springer.

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

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

Hutchins, E. (1995a). How a cockpit remembers its speeds. *Cognitive Science*, 19, 265–288.

Kirsh, D. (1995). Complementary strategies: Why we use our hands when we think. In J. D. Moore and J. F. Lehman (Eds.) *Proceedings of the Seventheenth Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum. pp. 212-217.

Kirsh, D. (1996). Adapting the Environment instead of Oneself. *Adaptive Behavior*, 4,415-452

Kirsh, D. (2001). The Context of Work, *Human computer Interaction*, 16(2-4), 305-322

Kirsh, D. (2010). Thinking With External Representations. *AI & Society*, 25 (4) 441–454

Nersessian, N. J. (1991). Why do Thought Experiments Work? *Proceedings of the 13th Annual Meeting of the Cognitive Science Society*, pp.480-486.

Nersessian, N. J., Kurz-Milcke, E., Newstetter, W. C., & Davies, J. (2003). Research laboratories as evolving distributed cognitive systems. *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. pp.857-862.

Nersessian, N. J. (2008). *Creating Scientific Concepts*. Cambridge, MA: MIT Press.

Nersessian, N. J., & Chandrasekharan, S. (2009). Hybrid analogies in conceptual innovation in science. *Cognitive Systems Research*, 10, 178-188.