Cognitive Science 33 (2009) 1059–1086

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/j.1551-6709.2009.01050.x

Building to Discover: A Common Coding Model

Sanjay Chandrasekharan

School of Interactive Computing, Georgia Institute of Technology

Received 17 September 2008; received in revised form 4 March 2009; accepted 31 March 2009

Abstract

I present a case study of scientific discovery, where building two functional and behavioral approximations of neurons, one physical and the other computational, led to conceptual and implementation breakthroughs in a neural engineering laboratory. Such building of external systems that mimic target phenomena, and the use of these external systems to generate novel concepts and control structures, is a standard strategy in the new engineering sciences. I develop a model of the cognitive mechanism that connects such built external systems with internal models, and I examine how new discoveries, and consensus on discoveries, could arise from this external-internal coupling and the building process. The model is based on the emerging framework of common coding, which proposes a shared representation in the brain between the execution, perception, and imagination of movement.

Keywords: Distributed cognition; Scientific cognition; Simulative model-based reasoning; Common coding; Scientific visualization

1. Introduction

Building functional and behavioral approximations of phenomena of interest is a standard mode of inquiry in science and engineering. This is particularly true in the engineering sciences such as bioengineering and nano-engineering, where basic science research and engineering advances go hand in hand. The built approximations include physical systems (such as artificial blood vessels, robot insects, nano-motors, wind tunnels, etc.) and computational systems (such as models of weather, models of molecular structure, models of plant growth, etc.). Such mimicking of target phenomena is an established practice, and it leads to new discoveries and novel predictions. A famous instance is the physical model developed by

Correspondence should be sent to Sanjay Chandrasekharan, School of Interactive Computing, Georgia Institute of Technology, 801 Atlantic Drive, Atlanta, GA 30332-0280. E-mail: sanjayan@cc.gatech.edu

Watson and Crick, which led to the discovery of the structure of the DNA. Computational models now play a similar role in discovery (e.g., see Lenhard, 2006; Winsberg, 2003, 2006a,b; E. Winsberg, unpublished data).

How does building such facsimile models provide new understanding of systems and lead to control over target systems? How do such external systems interact with internal models and expert knowledge to generate discovery? The usual answers to these questions revolve around how analogous the built model is to the target, and how the mapping from the model to the target is computed. While such analogical reasoning could be one possible way in which external models contribute to innovation, we have outlined elsewhere (Nersessian & Chandrasekharan, 2009) how existing models of analogy are limited in explaining this ‘‘building-to-discover’’ strategy.

Distinct from the analogy approach, it has long been argued in the philosophy of science that scientists reason with external artifacts themselves, such as diagrams, built models, and instruments (Nersessian, 2002a; Rheinberger, 1997; also see Hacking, 1983), and such reasoning, using this continuum of external artifacts, leads to conceptual innovation (Nersessian, 2002b, 2008). Much of this approach to innovation is based on historical studies of conceptual innovation in science and examines the origin of innovations that are novel for humanity as a whole (termed ‘‘H-creative’’ by Margaret Boden, see Nersessian, 2008). However, the role of external artifacts in conceptual innovation at the personal level (‘‘P-creative,’’ where a person discovers a concept new to her) has also been shown, using experimental settings (Martin & Schwartz, 2005) and the case study method (Nersessian, 2008). External artifacts are also increasingly being designed to accelerate conceptual innovation, both at the personal level and the community level. Examples of such at the personal level include the UVA Virtual Lab (virlab.virginia.edu), a Web-based virtual learning system which allows students to physically manipulate 3D molecules on screen to understand the structure of the DNA, and a recent system that uses haptic feedback and visual analogy to help students understand nano-level physics (Millet, Lecuyer, Burkhardt, Haliyo, & Regnier, 2008). At the community level, FoldIt (http://fold.it/portal/), developed by the University of Washington, is a video game platform for generating innovation, where players from around the world manipulate protein structures on screen, to develop novel protein-folding possibilities.

The role of external artifacts in generating conceptual innovation has thus moved from a philosophical argument based on historical studies, to controlled experiments and case studies showing such an effect, and finally to development of ‘‘tinker-media’’ applications that seek to accelerate innovation by using external artifacts. However, the cognitive mechanisms that support this build-to-discover process are far from clear. Within cognitive research, the role of external artifacts in problem solving has been emphasized mainly by researchers in distributed cognition (Hutchins, 1995a,b), and recent work has extended this approach to problem solving in science (Alac & Hutchins, 2004; Becvar, Hollan, & Hutchins, 2007). Most of this research is focused on descriptively capturing how cognition is distributed in technological and scientific environments, and it does not address the more complex problems of how novel concepts are generated by building and manipulating external artifacts, and what cognitive and neural mechanisms support the distribution of cognition.

Simulative model-based reasoning, proposed by Nersessian (2002a,b, 2008), provides an account of the cognitive mechanisms underlying innovation via the model building process. The leading framework in this area, this model argues for an ''internal-external coupling'' and brings together a wide range of cognitive mechanisms, including analogical modeling, mental modeling, visual reasoning, and discourse models, to develop a comprehensive account that explains how models are used to derive novel concepts. However, at its current state of development, this framework does not address the following three central issues involved in innovation by such model building:

1. What cognitive mechanisms allow internal and external models to be knit together seamlessly?
2. How could novel concepts arise from the building process, and what cognitive mechanisms support such generation by building?
3. How do groups of scientists come to a consensus on a novel concept generated from building external models, and what cognitive mechanisms support such consensus?

In this paper, I address these three specific questions. The objective is to build on the simulative model-based reasoning framework, by providing an account of the internal-external coupling mechanism and how it supports the innovation process and consensus building.

Roughly, I will argue that the above questions could be answered by the theoretical model known as common coding, which postulates that the execution, perception, and imagination of movements share a common representation in the brain. This coding leads to any one of these three (say perception of an external movement), automatically triggering the other two (imagination and execution of movement). One effect of this mechanism is that it allows any perceived external movement to be instantaneously replicated in body coordinates, generating a dynamic movement trace that can be used to generate an action response. The trace can also be used later for cognitive operations involving movement (action simulations). In this view, movement crosses the internal/external boundary *as movement*, and thus movement could be seen as a ''lingua franca'' that is shared across internal and external models, if both have movement components, as they tend to do in science and engineering. Based on this framework, I propose the following answers to the above three questions:

1. Models in science and engineering characterize phenomena in terms of movement of bodies or particles. Hence, internal models in science and engineering have movement properties that could be understood in terms of motor simulations (see Hegarty, 2004; Nersessian, 2002a). Building and running external models involve imagining and generating fine-grained movements, which could also be understood in motor simulation terms. Movement, and its instantiation using the motor system, is thus a common element between internal and external models. This ''lingua franca'' allows seamless integration between internal and external models.
2. Building and running external models leads to the generation of new and fine-grained movement patterns. But these patterns arise in a constrained fashion, as building an accurate facsimile limits the possibilities that can be considered. These

novel-but-constrained external movement patterns ''perturb'' movement-based internal models in a systematic fashion, helping overcome conceptual grooves such as functional fixedness. Such constrained perturbation leads to the generation of nonstandard, but plausible, movement patterns in internal models, which, in combination with mathematical and logical reasoning, leads to novel concepts.

3. Movements in external models are replicated in a similar fashion across team members. These movements couple with internal models in a similar fashion across people, as long as the members share a common internal model of the phenomena under investigation. This shared representation allows team members to reach a consensus on novel concepts generated by an external model.

I will argue for this threefold proposal using a case study that illustrates the build-to-discover strategy. The next section of the paper reports this study, where two facsimile models (a physical one and a computational one) were built in a systems neuroscience lab. The building of the computational model led to new insights and control of the physical model of the system. Some of the theoretical issues raised by this case of discovery are then outlined. In the third section, I present the common coding framework of cognition. In the fourth section I discuss the theoretical possibilities offered by this framework and use it to develop answers to the questions raised in the introduction. A few questions about the framework itself are also addressed. In the fifth section, I examine the limitations of this theoretical approach and outline some of the unresolved issues.

2. A case of innovation from building

Our research group has been carrying out a 6-year study of modeling practices in three interdisciplinary research laboratories in bio-sciences and bio-engineering, conducting both ethnographic investigations of the day-to-day practices and cognitive-historical analysis of the problems, artifacts, and models used in the research. One of the labs we study is a neural engineering laboratory, where the central research problem is to develop an account of learning and plasticity in cultured neuronal networks (as opposed to single neurons). To address this problem experimentally, the lab director (D6) has developed techniques to grow an in-vitro network of cultured neurons. Building this in-vitro system involves extracting neurons from embryonic rats, dissociating them (i.e., breaking the connections between neurons), and plating them on a ''dish'' with embedded electrodes known as an MEA (multielectrode array), where the neurons regenerate connections and become a network. The system supports stimulating the neuronal network with different electrical signals (electrophysiology) and, to some extent, chemical inputs. The output of the in-vitro network is displayed as graphs using a software interface known as MEAScope (see Fig. 1).

Building this in-vitro model requires a series of systems and techniques, including techniques for dissociation, preserving neuron cultures for long periods of time, dish seals that prevent contamination, electronics-friendly incubators, etc. Besides this infrastructure for the development and upkeep of the in-vitro model, there is another range of systems that

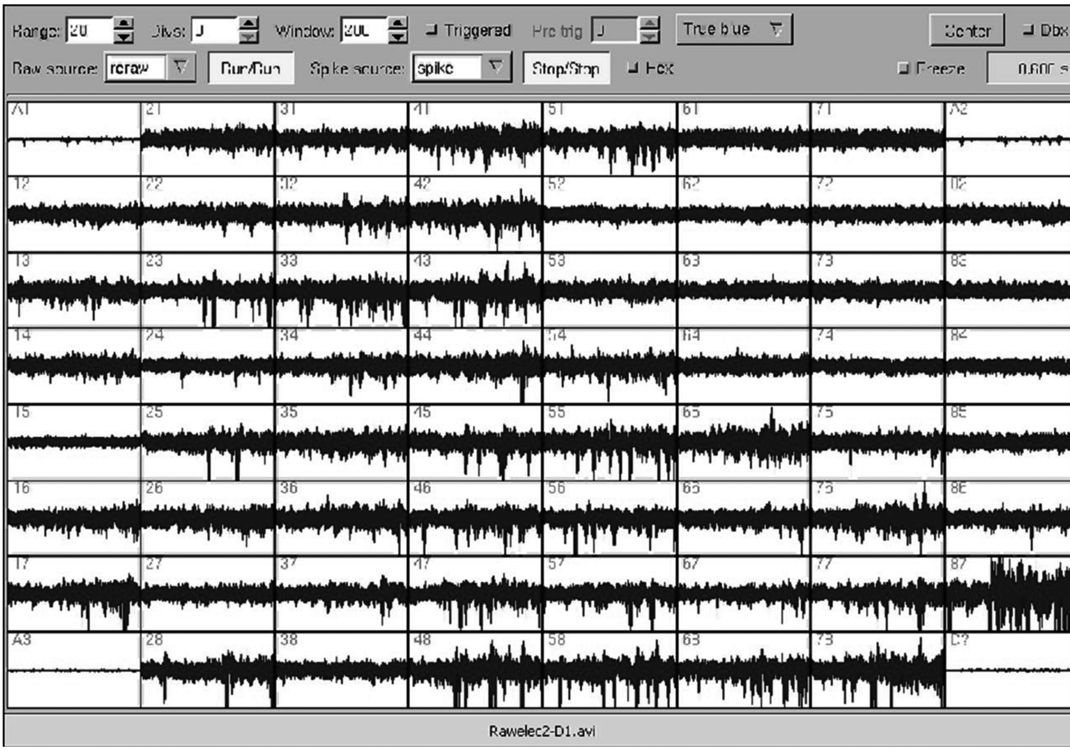

Fig. 1. MEAscope display of bursting phenomena.

support electrically stimulating the neurons, tracking their output, and feeding some of the output back into the network. Of particular interest is what the lab calls ‘‘embodiments,’’ which are robotic systems controlled by the output of the neuronal network in the dish. The primary embodiments are a virtual agent that moves around in a computer screen (animat) and a robotic arm that draws lines on paper. The embodiment and the dish can be connected in a ‘‘closed loop,’’ such that the dish’s output causes the embodiment to move (using a translator program that maps the dish signal to motor commands for the robot/agent), and this movement, in turn, is captured by video and converted back into electrical signals, which are then fed back into the dish. The stated goal of the lab is to understand the dynamics of learning in the neuronal network in such a way that it leads to the development of a *control structure,* which will allow the dish to be *trained* to control the embodiment systematically, using feedback (for a detailed discussion, see Nersessian & Patton, 2009).

Our study began very close to the establishment of the lab (early 2003), when it had three researchers (D2, D4, D11). They were all ‘‘playing with the dish,’’ which is the lab’s term for using the MEA to stimulate the neuronal network using different electrical signals, and tracking the output. Since the lab had just started, and there existed very little work on learning and plasticity in such cultured neuronal networks, there were limited constraints on approaches to solving the problem. The only theoretical model guiding the lab’s work was the Hebbian rule (roughly stated as: neurons that fire together, wire together), originally developed for single neurons, but the implicit assumption in the lab was that it applied to neuronal networks as well.

For clarity of exposition, I will present the research activity of interest in four phases, although the research was ongoing and there were no such distinct phases of activity.

2.1. Research phase 1

The work began by trying to replicate a plasticity result reported by a Japanese group. D4 initiated this experiment, but she was not successful in replicating the reported results. One of the problems she faced was ''bursting,'' a form of networkwide electrical activity spontaneously exhibited by such cultured neuronal networks. Bursting created a problem in understanding plasticity, because it prevented the detection of any systematic change that arose due to controlled stimulation of the network. The group hypothesized that bursting arose because of deafferentation, that is, because the neurons in the dish lacked the sensory inputs they would ordinarily get if they were in a live animal's brain. Given this view, and the problem of the noise generated by bursts, D4 decided that she needed to get rid of bursts to generate and study plasticity, and began working on ''quieting'' bursts in the dish. She hypothesized that since bursting possibly arose because of deafferentation, it would be possible to lower bursting by providing the network with artificial sensory input, that is, some form of electrical stimulation. Sometime in late 2003, she achieved a breakthrough, managing to ''quiet'' bursts entirely in a dish network using a sequence of background electrical stimulation. As per the group's deafferentation hypothesis, this result was presented as a way of using electrical stimulation to ''substitute'' for natural sensory input.

D2, meanwhile, was working on translating the neuronal network's output as motor commands to control the agents. D11, however, had decided early in this phase to branch away from the cultured neuronal network entirely and develop a computational model that mimicked the in-vitro system. Notably, the model was not based on the dish's output, but on results from existing literature. This model, when developed, would be a second-order in-vitro system, an ''in-silico'' system approximating the existing in-vitro system as much as possible (see Nersessian & Chandrasekharan, 2009, on the analogical relationship between the in-vitro dish and the computational model, and how it emerged over time.)

2.2. Research phase 2

This part of the research started at the end of 2003 and involved D4 trying to induce plasticity in the burst-quieted network. However, for almost two semesters, she could make no progress, as no constant change could be detected in the network. This was mostly because of a ''drift'' phenomenon, where the activity pattern evoked by a stimulus did not stay constant across trials but ''drifted'' away to another pattern. This ''drift'' prevented her from tracking the effect of a stimulus, as the network never responded the same to a constant stimulus.

D11, meanwhile, was spending most of his time away from the lab space, working on modeling the dish. The model tried to replicate the dish as closely as possible, down to the area covered by the artificial neurons (3 mm × 3 mm), the grid of electrodes (8 × 8), and the number electrodes used for recording and stimulation (60). It was developed using a basic neuroscience modeling platform (CSIM) and used a standard simple model of neurons, known as ''leaky-integrate-fire.'' The name is based on the way the artificial neurons respond to stimulus, where each neuron is considered similar to a ''bucket'' of activation, and when it gets stimulated, the activation rises. Once it reaches a certain threshold, the neuron fires. It is called leaky because the activation level ''leaks'' out all the time—so if it does not get any stimulation, the activation level leaks to the bottom. The parameters of the model, such as type of synapses, synaptic connection distance, percentage of excitatory and inhibitory neurons, conduction delay, conduction velocity, noise levels, action potential effects, spontaneous activity, etc., were based on results reported in the literature.

By the mid-2004, D11 had developed the model, and he started getting what he calls a ''feel'' for his model network, by which he means (presumably) how the network behaves under different conditions. He then developed a visualization that captured the activity in the network (such as the distribution of the synaptic weight change and synaptic state changes, for more than 50,000 synapses) in real time. He then successfully replicated some of the results reported in the literature using this computational model.

2.3. Research phase 3

This is the interesting phase from the point of view of scientific cognition, and it started in late 2004. D11 started noticing interesting patterns in the way his computational model responded to different stimuli, patterns that were novel and distinct from what was known about the cultured dish. This novelty emerged because of three reasons, which I will label control, visualization, and cost.

1. *Control*: D11 had the ability to stop his network at any point he wanted, and start it again from there. He could also measure variables such as synaptic strength. The in-vitro system did not support these two features.
2. *Visualization*: D11 had a visualization of the network's activity that showed the *movement* of an activity pattern *across the network,* in real time (see Figs. 2 and 3). This feature was missing in the in-vitro system, as the neuronal activity is hidden, and had to be charted using display elements. The display in the in-vitro system tracked only activity at each electrode, using *graphs* (see Fig. 1). The graphs of activity in the electrodes captured which electrode was activated and by how much, but since it did not have a representation of the network itself, it did not capture how the *pattern moved across the network.*
3. *Cost*: D11 could run a large number of experiments at no cost, as the computational model could be changed easily and did not require the careful and laborious process involved in setting up a dish.

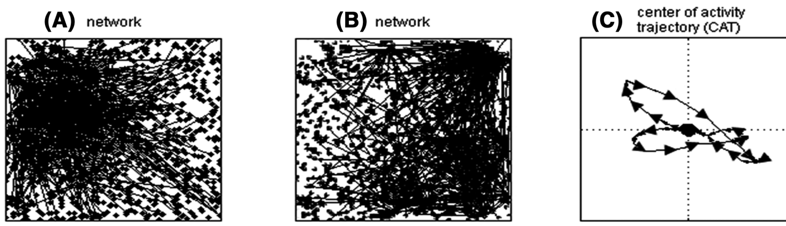

Fig. 2. Visual representations of changes in activity across the in silico network. (A) Time 1, (B) Time 2, (C) CAT T1 to T2.

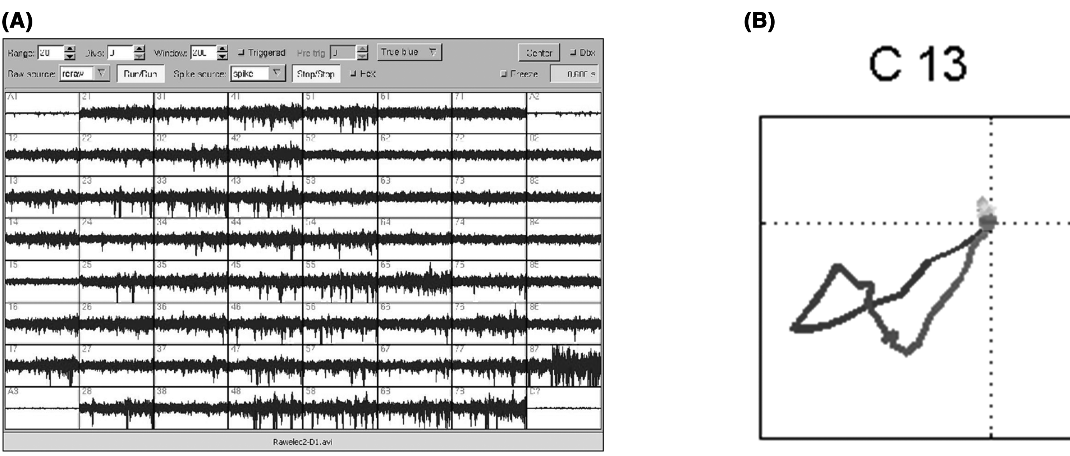

Fig. 3. A side-by-side comparison of a CAT visualization that corresponds to an MEAscope representation for an in vitro dish. The MEAscope representation (A) shows activity across each recording channel over time, but the CAT (B) is an integrated representation of activity across the entire network.

The combination of these three features (control, visualization, cost) proved very powerful: Control allowed D11 to stop and start the network as he pleased and to measure variables not accessible using the in-vitro system; the lack of cost allowed him to quickly set up the artificial dish in any configuration he found interesting; and the visualization allowed him to visually track the activity of the network as it was happening, that is, in real time. This gave him instant access to a range of stimulus–response configurations and data that the dish could not provide, and the data could be examined and reexamined, and comparison experiments run instantly, at almost no cost. From this, he noticed that there were *spatial* patterns in activity, as the activity propagated across the network. The spatial patterns were seen both when spontaneous bursts arose in the model network and when the model network responded to stimuli. Basically, he found that there were ‘‘similar looking bursts’’ that propagated across the network. He discussed this finding with D4, who was also having second thoughts about bursts, as she was not getting any results from her plasticity experiments on the burst-quieted network. The group then decided to investigate bursts further, as a possibly interesting pattern, that is, *a signal*.

Note the radical change in perspective here, where the burst moves from something akin to noise that needed to be quieted, to the status of a pattern, a signal that could possibly lead to a control structure for training the network.

2.4. Research phase 4

In this part of the work, the group came up with a range of ways to quantify the spatial properties of moving bursts, using clustering algorithms and statistical techniques. These measures were immune to the drift problem. Interestingly, these were first developed for the computational model, and equivalents were then developed for the real dish. The measures included the idea of burst types (small, medium, big), spatial extent (an estimate of the size and location of the burst in the dish), center of activity trajectory (CAT, a vector capturing the spatial location of the electrode along with the firing rate), and ''Occurance'' of bursts (what types of bursts occurred when). These spatial measures of bursts were then shown to be better indicators of plasticity than responses to probe stimuli. In the matter of a year, based on the patterns generated from the computational model, the group's theoretical position had shifted from bursts as noise to bursts as signal, and a possible structure to control the network of real neurons in the dish.

All these measures were driven by the spatial patterns of activity noticed by D11 in the computational model. But of these, CAT is probably the most noteworthy, both because it is a novel concept (it applies the idea of center of mass to an activity pattern) and because it emerged entirely from the computational model, and would be almost impossible to conceptualize and properly formalize without it. The center of mass is the unique point in an object or system that can be used to describe the system's response to external forces and torques. In analogy to this, CAT develops an averaging notion similar to the notion of population vectors, which captures how the firing rates of a group of neurons that are only broadly tuned to an action or stimulus (e.g., an arm movement), when taken together, provide an accurate representation of the action/stimulus. However, CAT is more complex than the population vector, because it tracks the spatial properties of activity as it moves through the network. If the network is firing homogenously, the CAT will be at the center of the dish, but if the network fires mainly at the left corner, then the CAT will *move* in that direction. CAT thus tracks the ''flow of activity'' (not just activity) at the population scale, and on a much quicker time scale than population vectors (see Fig. 3). It is thus a new concept in understanding neuronal activity.

The burst-quieting work was not wasted, though. In recent work, D2 and D11 have combined CAT and techniques developed for burst quieting to develop a set of stimulation patterns (a control structure) for the dish that led to supervised learning by the living neuronal network, in effect making the in vitro dish neuron network ''programmable.'' Using these stimulation patterns, the network in the dish was trained to control both the animat (virtual agent) and the drawing arm. Interestingly, this control structure consists of providing the network with a patterned stimulation, and then, instead of reinforcing this stimulus, the network is provided a random background stimulation to ''stabilize'' its response to the patterned stimulation. This training method is rather counterintuitive to existing ideas about learning and reinforcement, and it would not have emerged if not for the building of the two facsimile models.

I will end the case study here. It presents an instance of what we have termed elsewhere ''hybrid analogies''—analogies that make inferences from intermediary models, which are constructed (over time) by merging constraints from both source and target domains (Nersessian & Chandrasekharan, 2009). I outline below some of the specific questions raised by this case study.

1. Why did the researchers not think of the spatial nature of the activity before the computational model, even though the group used global and spatial concepts such as population vectors, and was precisely interested in network-level learning? Also, as D4 says, ''*we had the information always… the information was always there.*''
2. By what process did seeing the spatial pattern in the computational model lead to a change in the researchers' cognitive model of the network?
3. How did control contribute to the shift in perspective?
4. What type of cognitive structure underlying an internal model would allow such a broad change to occur quickly, across people, based just on observed movement patterns?
5. What desirable features ought such built external models possess, to support discovery?
6. And in general, how can building lead to innovation? How does building external facsimile models contribute to discovery?

In the following sections, I will develop a framework that extends the simulative model-based reasoning approach in a way that it accounts for how internal and external models interact, explain the role of construction in discovery, and how distributed cognitive systems such as research laboratories (Nersessian, Kurz-Milcke, Newstetter, & Davies, 2003) evolve over time. This account is based on the common coding framework, which I outline below.

3. Common coding

Recent work in cognitive science and neuroscience has illuminated a model of cognition where perception, execution, and imagination of movements share a common coding in the brain (Decety, 2002; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1992, 2005). The roots of common coding go back to the ideomotor principle, first outlined by William James:

> Every representation of a movement awakens in some degree the actual movement which is its object; and awakens it in a maximum degree whenever it is not kept from doing so by an antagonistic representation present simultaneously in the mind. (James, 1890: 526)

To illustrate, going round and round can make you dizzy, but equally, watching something go round and round can also make you dizzy. Note that watching a rotating disc could also generate such dizziness, so the effect is not limited to the observation of biological motion. This ideomotor effect is explained by a common coding in the brain that connects an organism's movement (motor activation), observation of movements (perceptual activation), and imagination of movements (simulation). First clearly articulated by Prinz (1992), this common coding allows any one of these movements to automatically generate the other two movements (Prinz, 2005; Sebanz, Knoblich, & Prinz, 2005; also see Decety, 2002; Hommel et al., 2001). One central outcome of common coding is a body-based ''resonance''—the body instantly replicates all movements it detects, generating an internal representation that is dynamic and based on body coordinates. This replication generates a dynamic trace, which can play a role in later cognition. All the replicated movements are not overtly executed or responded to. Most stay covert, as the overt movement is inhibited.

Common coding is closely related to embodied cognition (Barsalou, 1999; Gibbs, 2006), though the two theoretical streams developed almost in parallel till recently. One way to understand the relation between the two is to consider embodied cognition as a high-level description, and common coding as one of the neural mechanisms that support such cognition. The basic argument for common coding is an adaptive one, where organisms are considered to be fundamentally action systems. In this view, sensory and cognitive systems evolved to support action, and they are therefore dynamically coupled to action systems in ways that help organisms act quickly and appropriately. Common coding, and the resultant replication of external movements in body coordinates, provides one form of highly efficient coupling. Since both biological and nonbiological movements are equally important to the organism, and the two movements interact in unpredictable ways, it is beneficial to replicate both types of movements in body coordinates, so that efficient responses can be generated.

In implementation terms, common coding can be thought of as an artificial neural network encoding both action and perception elements, where the activation of one type of element automatically activates the other (associative priming), similar to connectionist implementations of semantic priming (Cree, McRae, & McNorgan, 1999). Imagination of movement, in this view, would be a form of implicit activation of the action network. It has been proposed that such common coding could arise from Hebbian learning (Heyes, 2005). Recent modeling work has shown that such common coding arises when organisms shift from a nonrepresentational mode to using representations, and this coding could emerge from both evolutionary and within-lifetime learning (Chandrasekharan & Stewart, 2007).

In operational terms, common coding implies that there are interactions between execution, perception, and imagination of movement. I review experimental evidence for different types of such interactions below. Most of the behavioral evidence for common coding is based on interference effects, where actions in one modality (e.g., imagination) leads to a difference in reaction time or accuracy in another modality (e.g., execution). This behavioral evidence is supported by neurophysiological experiments, including imaging, TMS, and patient studies.

3.1. Perception-action common coding

If common coding holds, perception of movement should interfere with execution of movement. Brass, Bekkering, and Prinz (2002) showed that when participants execute an action A (e.g., tapping fingers on a flat surface), while watching a noncongruent action on a screen (e.g., another person's finger moving in a direction perpendicular to the tapping), the speed of performed action A slows down, compared to the condition when the participant is watching a congruent action on screen. This is because the perceived opposite movement generates a motor response that interferes with the desired tapping pattern. A similar interference effect has been shown for competing movements within an individual—movement trajectories of participants veer away or toward the location of competing nontarget objects (Welsh & Elliott, 2004). Supporting this effect, perceiving an action has been shown to prime the neurons coding for the muscles that perform the same action (Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995). Establishing the common coding hypothesis further is the reverse of the above, where actions influence perception. Blindfolded subjects, after learning a new sequence of movements based just on verbal and haptic feedback (Casile & Giese, 2006), visually recognized the learned movements faster, compared to recognition of other movement sequences. Further, recognition performance correlated strongly with the accuracy of the execution during learning.

Supporting these behavioral data are a range of neuroimaging experiments that show that action areas are activated when participants passively watch actions on screen (Brass & Heyes, 2005 provides a good review). Expert performers of a dance form (such as ballet and capoeira) when watching video clips of the dances in which they are experts, show strong activation in premotor, parietal, and posterior STS regions, compared to when watching other dance forms. Nondancer control participants do not show this effect. Similar motor activation has been shown for expert piano players watching piano playing. When we observe goal-related behaviors executed by others (with effectors as different as the mouth, the hand, or the foot) the same cortical sectors are activated as when we perform the same actions (Gallese, Ferrari, Kohler, & Fogassi, 2002). The neuronal populations that support this blurring of first-person and third-person views have been termed ''mirror neurons'' (Fadiga, Fogassi, Gallese, & Rizzolatti, 2000). In contrast, motor areas are not activated when humans watch actions not part of human repertoire (such as barking). A similar effect has been replicated across a series of invasive studies in monkeys (see Hurley & Chater, 2005 for comprehensive reviews).

Two additional neural mechanisms supporting common coding have been reported. Canonical neurons fire both when a monkey grasps an object and also when it observes a ''graspable'' object (Oztop, Kawato, & Arbib, 2006), indicating a common coding between action and perception of action affordances (Gibson, 1979). Another supporting mechanism is the behavior of visio-tactile bimodal neurons during tool use. These neurons fire both when a monkey's hand is touched and also when light is shown near the hand. When the monkey uses a stick to get food, the visual fields of these neurons ''extend out,'' now firing when a light is shown near the end of the stick. This effect occurs only when the stick is actively used, not when it is held passively, indicating that the perceptual ''extending out'' is driven by common coding with action (Farne, Iriki, & Làdavas, 2005; Iriki, Tanaka, & Iwamura, 1996). This effect has been shown for people as well, including blindsight patients.

3.2. Imagination-action common coding

I will use mental rotation work to illustrate this case, though the interaction between imagination and action has been shown in many other areas (see Nersessian, 2002a, 2008; for a review of action-imagination links as they relate to scientific thinking). If imagination and execution of movement share a common code, imagining a movement should affect the execution of movement. Wohlschlager (2001) showed that while imagining a mental rotation, if participants plan another action, or move their hands or feet in a direction noncompatible to the mental rotation, their performance suffers. This effect is reversed for compatible movements. Unseen motor rotation leads to faster reaction times and fewer errors when the motor rotation is compatible with the mental rotation, and speeding/slowing the compatible motor rotation speeds/slows the mental rotation (Wexler, Kosslyn, & Berthoz, 1998).

Supporting the common code view further, it has been shown that the time to mentally execute actions closely corresponds to the time it takes to actually perform them (Decety, 2002; Jeannerod, 2006). Responses beyond voluntary control (such as heart and respiratory rates) are activated by imagining actions, to an extent proportional to that of actually performing the action. When sharpshooters imagine shooting a gun, their entire bodies behave as if they are actually shooting (Barsalou, 1999). Similarly, imagining performing a movement helps athletes perform the actual movement better (Jeannerod, 1997).

Links between imagination and action have also been found by experiments investigating mechanical reasoning, such as how people imagine the behavior of pulleys, gears, etc. (see Hegarty, 2004 for a review). Children who learn fractions by actually executing movements on blocks learn the fraction concepts better than others who do not perform such movements (Martin & Schwartz, 2005). Imaging experiments support these behavioral results and show that premotor areas are activated while participants do mental rotation (Vingerhoets, de Lange, Vandemaele, Deblaere, & Achten, 2002).

In the other direction, common coding would suggest that our action possibilities restrict imagination of novel actions and movements. Kosslyn (1994) reports that participants need more time to perform mental rotations that are physically awkward. People with writer's cramp (focal hand dystonia) take more time to do mental rotation of hand pictures, and people have difficulty mentally rotating manually difficult hand movements, such as right-sided stimuli at 120 degrees and left-sided stimuli at 240 degrees (Fiorio, Tinazzi, & Agiloti, 2006). According to common coding, we understand another person's actions by reenacting those actions using our own motor system. An example would be judging the weight of an object by watching how a person lifts a heavy object. Bosbach, Cole, Prinz, and Knoblich (2005) recently showed that people with compromised ability to activate their body, such as deafferented individuals, cannot make such predictions, suggesting that the action system is used in such judgments.

There is also evidence for a perception-action-imagination common coding. For instance, the perception of a drawing leads to participants simulating the actions that generated the drawing, and using this movement information while making judgments (see Viviani, 2002 for a review). Recent work has extended common coding to language processing, showing that there is motor activation while imagining words encoding movements, and processing sentences involving movements (Barsalou, 1999; Bergen, Chang, & Narayan, 2004; Holt & Beilock, 2006; Wilson & Gibbs, 2007). Finally, the ''resonance effect'' has been shown across people: when two participants perform reaction time tasks alongside each other, each actor's performance is influenced by the other's task movements (Knoblich and Sebanz, 2006; Sebanz et al., 2005; Welsh et al., 2007).

From this brief review, it is clear that there is significant evidence in favor of a common code linking execution, perception, and imagination of movement. The next section applies this theoretical position to the problem of built models.

4. Common coding and built models

An interesting theoretical property of common coding theory is that it combines representational and dynamic processing in an elegant fashion (Chandrasekharan & Osbeck, in press). By promoting common *coding* at the neural level *during learning,* the theory is unabashedly representationalist. However, one effect of this coding is that *during run-time*, the activation of movement in one modality (e.g., perception of movement) automatically, that is, *directly or instantaneously*, activates the others (imagination/execution of movement), thus translating an external movement into covert movement in body coordinates immediately. The traces generated by this direct/instantaneous translation also supports the covert activation of the motor system in a temporally extended fashion, such as in imagining a mental rotation.

Common coding is thus a representationalist account, but it proposes a representation that supports a motor simulation mechanism, which can be activated across different time-scales—instantaneous simulation of external movement, and also extended simulations of movement. The latter could be online, that is, linked to an external movement (as in mental rotations while playing Tetris, see Kirsh & Maglio, 1994), or can be offline (as in purely imagined mental rotation). As both ''direct'' and offline simulations can be activated by perception of external movement, the model offers a common body coordinate–based representation that connects external models with internal models (when both possess movement).

How does this notion of common coding of imagination, perception, and action help in understanding the way external models help generate novelty? I will answer this question in two parts, first laying out a macro level overview of the theoretical possibilities provided by common coding, and then going to the micro level, applying common coding to the case study.

At the macro level, common coding could provide an account of three features of building-based discovery—internal-external coupling, role of construction, and epistemic coordination (between team members).

4.1. Internal-external coupling

The common coding view provides a way of thinking about the interaction between external and internal models—as a trade in movement. In essence, the ideomotor effect supports an agent-world coupling based on movement, where every perceived movement generates a resonant movement in the organism. Some of this resonant movement is registered consciously, and some leads to overt execution of actions, but most are inhibited and remain unconscious. However, every perceived movement in the world registers a resonance and a ''trace,'' and these resonances and traces have the potential to make a change to internal models of the world, as these models (particularly in science and engineering) are also coded in movement terms.

The ideomotor effect thus provides a way in which external and internal models could work together seamlessly, as activity generated outside would be replicated in body coordinates by the activation of the motor system, and integrated easily into an internal model of movement that is also based on motor simulation. The results on imagination-action common coding point at internal models running partly on the motor system, and it has been argued that imagined models in science and engineering have movement components (see Nersessian, 2008; Hegarty, 2004; also see Barsalou, 1999). So if the external model generates resonant movements automatically, internal and external models share a motor ''lingua franca,'' which allows the externally generated movements to be integrated quickly into the internal model.

4.2. Role of construction

Common coding also shows why *building* an external model (construction) is important in generating novelty. I will lay out this feature as a set of points.

1. Building a system that exhibits the same behavior as a target system involves thinking of the system in component terms, and then imagining how these components would interact dynamically to produce the target behavior. This process automatically generates a flurry of internal movements in imagination. These are detailed movements that are not generated by just observing the target system. An example would be thinking of how to build a bicycle, as opposed to watching a passing bicycle or imagining a bicycle you owned. The latter two would not generate internal movements of, For example, the intricate and interconnected movements between the pedals, chain, gears, and the wheel, or ball bearings.
2. As the external system is built, it generates its own movements, which may be in conflict with the imagined internal model of movement patterns. This leads to revision—of either the internal or external model, or both.
3. The system is built component by component, and this provides more fine-grained *control* over the way the different components interact and the ability to manipulate each component. This generates a further level of fine-grained movement parameters, which are not provided either by observing the target system or by the imagined model of the target system.

4. Building an external model of behavior involves a *synthesis* of a large set of component movements. So the building process also provides integrated information—about how movements influence each other and global movement patterns. Also, this shows how the properties of different components relate to each other in dynamic terms, such as heat generated/attractive-repulsive forces/centrifugal-centripetal effects that arise during movement.
5. Building is based on primary actions, while imagination is based on secondary, stored, actions. This means building movements have more impact compared to imagined movements, and they can thus help override imagination's existing grooves of movement (functional fixedness, which arise from traces of previous actions and perceptions).

These features provide a way of thinking about how external facsimile models generate novelty: They generate external movements that ''perturb'' the internal models in a seamless and integrated fashion. Imagination by itself can perturb the internal models in many ways, but the changes it can generate are constrained by factors such as memory, cognitive load, repertoire of existing movements, use of standard manipulations (''functional fixedness,'' Adamson, 1952; German & Defeyter, 2000; German & Barrett, 2005), etc. At the other extreme, imagination running by itself can lead to unconstrained changes in parameters (''flights of fancy'').

The external model provides more freedom and flexibility, as it retains its states, thereby lowering memory and cognitive load. It also provides ways of escaping fixedness, as the building process generates more movements, or possibilities of movements. The model serves as a more flexible ''external imagination,'' as it can generate a wide variety of movements that are not supported or activated easily by imagination. This external imagination is also tightly constrained by executable possibilities. Common coding ensures that these external movements seamlessly perturb the internal model, almost as if external manipulation is the same as imaginative manipulation. Just as making a compatible physical action speeds up mental rotation (Wexler et al., 1998) and planning incompatible movements interfere with mental rotation (Wohlschlager, 2001), the act of construction could perturb the internal model.

There are a few strands of evidence that point to this role of external models. Martin and Schwartz (2005) showed that manipulating physical pieces facilitated children's ability to develop an interpretation of fractions. They also report that children who learned by adapting unstructured environments transferred to new materials better than children who learned with ''well-structured'' environments. Based on this, they argue that the opportunity to adapt an environment permits the development of new interpretations that can advance learning. In a related vein, German and Defeyter (2000) report that children 5 years old and younger are less prone to functional fixedness compared to children 7 years old and older. It is possible that this is because the movements of younger children are less ''set'' than the older ones, and hence more amenable to novel combinations. Craig, Nersessian, and Catrambone (2002) report a study using the radiation problem, where a doctor has to devise a way to kill a tumor using radiation, but any radiation strong enough to kill the tumor would also kill the surrounding tissue. The correct strategy is to use weak radiation from many directions. Before trying to solve the problem, students were first exposed to an analogical story (a general capturing a fortress using small groups). One group of students then recalled this analogy using sketches, another by manipulating wood blocks, and the third just verbally. While solving the radiation problem, the wood block group performed better than the sketch group, and both performed better than the verbal recall group, suggesting that the physical activity allowed them to better connect the analogy story and the radiation problem.

4.3. Epistemic coordination

The common coding approach helps in understanding how external models allow groups of people to reach a consensus on models, concepts, and control structures, and provide groups with a common way of thinking about problems and solutions. In existing work on epistemic coordination, distributed cognition theory considers as given the common conceptual structure of an external artifact across team members; it does not ask the question how this structure arises across people. Further, the theory does not specify whether the representations traded among team members are amodal or modal (Barsalou, Simmons, Barbey, & Wilson, 2003) in nature. If they are amodal, such a shared structure arising across people is a mystery, because amodal representations are considered to have arbitrary relations to entities in the world, and there is no reason why each person, particularly people new to the workspace, should have the same representation of the artifact.

On the other hand, in a movement-based model, the motor system ''resonates'' to an external movement automatically, which means every detectable movement is picked up and replicated in some fashion by the system, even though the movement is not always overtly performed. The level of arbitrariness in this replication is much more limited, as the replication of a movement by the motor system would be broadly similar in everyone. While there would be variation in the way the movement is integrated into the internal model, the common resonance of the movement generated by an artifact or model system provides a base structure that is shared by everyone in the workspace.

Further, such a replication of the artifact/system's movements, combined with replication of the actions generated by humans interacting with the artifact/system, could also support second-order understanding—based on repeated movement patterns, and their automatic replication across the group, everyone ''knows,'' or has a sense that, there is a common understanding of the artifact/system. These first- and second-order shared internal structures based on the artifact/system's movements support an implicitly shared, *and considered as shared*, internal model of the artifact/system. This leads to the coordination of workflow around the artifact/system and also supports the emergence of common approaches to solving problems. Note that this ''hub'' quality, combined with the ability to generate new movements, leads to the model system working as an ''imagination hub'' as well, as the shared internal model of the system allows people to generate similar or overlapping ideas and follow, support, or reject novel ideas generated by others easily. Work within common coding theory has recently started examining such sharing of movements across subjects during joint tasks (Sebanz et al., 2005; Welsh et al., 2007), and results show that such sharing indeed emerges.

Closely related to this ''hub'' notion is social learning, particularly imitation-based learning of techniques in science and engineering labs. Such learning is crucial to understanding distributed cognitive systems such as laboratories that evolve over time (Nersessian et al., 2003). One of the primary areas of application of common coding theory is in the understanding of imitation (the primary method of social learning), where it is argued that imitation is based on the automatic activation of others' actions by the observer's motor system (see Brass & Heyes, 2005; Hurley & Chater, 2005; for reviews). Our observation of the two biomedical engineering labs showed that lab members mentor juniors initially in techniques, based on an imitation-based apprentice style of learning (Alac & Hutchins, 2004; also see Becvar et al., 2007). Later on, the juniors pick up a style of learning we have termed ''agentive learning'' (Newstetter, Kurz-milcke, & Nersessian, 2004), which is an opportunistic and self-motivated learning that *weaves together people, systems, and environmental affordances* in an integrated fashion to solve problems. The common coding framework allows us to examine this style of learning in more detail, examining the combination opportunities provided by different system movements and internal movements, and how these could relate to such agentive learning.

4.4. Common coding and the dish system

How does the common coding framework help us answer the questions raised about the case study? In the following section, I will take up the questions raised in section 1 (with reference to how D11's computational model changed D4's notion of bursts-as-noise) and lay out how the common coding framework could address these questions.

> Why was it not possible to think of the spatial nature of the activity before the computational model, even though the group used global and spatial concepts, such as population vectors, and was precisely interested in network-level learning?

Obviously, the answer is visualization. But visualization viewed traditionally is a representation, where the ''standing in'' relation it has to the underlying data is what is crucial. In the common coding framework, movements in the visualization are *also* a way of generating equivalent movements in body coordinates. By this resonance system, the movements in the visualization influence the internal model directly. This means visualization is not just a user-friendly mapping of the data generated by the system, *it is a way of transferring the system's behavior to the internal model of the system.*

However, both the in-vitro and the computational model used visualizations, so why did not the former lead to a breakthrough? The in-vitro dish's output (MEAScope graphs, see Fig. 1) did not have global movement-generating properties, so it was not able to generate global spatial concepts such as CAT and burst types. For one, the MEAScope visualizations were detached from the dish itself and showed activity at each node using graphs of spikes. So while there were movements on screen, they were spiking movements detached from the dish, and the spatial activity pattern of the spikes was hidden among the different graphs. Further, the spiking movements required a mapping function to link it to the internal (network) model of the dish. The spikes thus acted as both a limitation (submerging the global movement patterns) and a form of fixation (activating spike-based internal movements more than other possible ones, such as network-based internal movements).

The computational model, on the other hand, used a visualization where the activity moved across the dish network itself, thus allowing an easy integration of the network activity seen on screen with the internal model of activity in the dish, *in movement terms*. The internal model (movement of electrical activity through neuronal connections) was linked to a specific pattern of such movement, namely a movement with a spatial focus, but ''jumped'' around (CAT). The combination of the internal model of the dish and the spiking activity in the computational model generated a model of dish activity that had spatial parameters and movement behavior across the dish, leading to the new spatial concepts.

How did seeing the spatial pattern lead to a change in the internal model?

Seeing the spatial pattern generated novel motor simulations, which pushed the internal model out of the spike-based movement groove (local minima) it was in, and raised the types of manipulative movements *available* for the internal model (see availability heuristic, Kahneman & Tversky, 1982). Also, the spatial movement patterns and the stop-go control were related, which led to more correlated movements. This helped in getting out of the spike-based minima as well. Since the structure of the dish and the network were integrated in the visualization, the spatial movement pattern could influence the internal model's movements more directly. This integrated visualization also helped lower memory and cognitive load, as the external model retained its states, could be stopped at any time, and there was no mapping required between the movements in the visualization and the movements in the internal model of the dish.

How did control contribute to the shift in perspective?

The primary contribution of stop-go control is its ability to generate more fine-grained movements. This is particularly useful in cases where imagination is constrained by more ''available'' parameters. In such cases, the external model makes different parameters available for movement. D11 could stop and start the network activity as he pleased, thereby raising the number and types of manipulations his, *and others'*, imagination could execute on their internal model of the network. In our particular case, this was very important, because this large number of varied movements generated the case for types of bursts, and movements of such bursts across the system, and these two ideas ultimately led to discarding the notion of burst as noise. If the number of movements generated were limited and less varied (in terms of both output and input), the classification based on types and location would not arise, as the results from the limited trials would be unrecognizable from noise, and thus would not generate these patterns.

The second contribution of control, in common coding terms, is the activation of an ''interventional stance,'' where the agent actively tracks the external movements and generates intervention plans. An example would be a video game, where the player constantly generates and revises action plans, and tracks the environment actively in relation to them. In contrast, watching an action movie only generates replication of the external movements. Control leads to constant tracking of the simulation, both in surface activity as well as internal activity. This helped in qualitatively judging the rate of activation (of the CSIM neurons), and then judging the rate of propagation of activity through the network, both in terms of movement. This, in turn, supported judging the spatial extent, and comparing the extent of different patterns that are generated, leading to the idea of burst types. Stop-go control allowed building stronger correlations between stimulation and activation of the model. The fine-grained movements provided by such control, together with the tracking activity involved, helped in detecting and standardizing common movements and their effects (thus forming causal relations).

What type of structure underlying an internal model would allow such a broad change to occur so quickly across people, based just on observed patterns?

The fact that the visualization of movement changed the internal model so quickly is an indirect indication that the underlying structure of the internal model has movement properties. If the internal model has an amodal structure, for example, with quantifiers and algebraic notations, it would not change as quickly, across people, based on visualization of movement. The case of D11 is not an isolated instance of conceptual change based on such visualization. Entire methodologies, disciplines, and phenomena challenging existing models have been built just from observed movements on computer screens. These include Complexity Theory (Langton, 1984, 1990), Artificial Life (Reynolds, 1987; Sims, 1994), models of plant growth (Prusinkiewicz, Lindenmayer, & Hanan, 1988; Runions et al., 2005), computational bio-chemistry (Banzhaf, 1994; Edwards, Peng, & Reggia, 1998), and computational nanotechnology (reported in Lenhard, 2004; Winsberg, 2006a). All these novel areas of exploration are based on movement systems and could not exist without visualizing movements. From a common coding perspective, the ability of such visualizations to challenge existing models is not a mystery, because most internal models in science involve movement, and visualized movements can interact with these internal models in a seamless way, across the population who possess such internal models. Before the advent of the computer, such external movements were approximated by diagrams and giving directions to the audience to move the elements in the diagram in imagination, in specific ways (Nersessian, 2002a, 2008), sometimes using arrows. This process is very close to the simulation of actions from end-point movements such as writing and drawing (Viviani, 2002).

How can building lead to innovation? How does building external facsimile models contribute to discovery?

The process of building a model requires both executing novel movements and generation of novel movements by the model. These movements alter internal movement-based models of a target system. Since the objective is to build a working model, it sets constraints on which internal movements are activated. Building external models is thus a way of *perturbing* the imagination in a focused and constrained fashion, and this perturbation is one way in which facsimile models can generate novelty. A related way in which such ''perturbation'' could contribute to discovery is by generating random combinations of internal and external movements, and thereby ''connecting'' brain regions that are not activated together ordinarily, while just imagining movement (see Schubotz & von Cramon, 2004, for imaging evidence for such a ''thread'' across different perceived movements).

A second way in which building leads to innovation is via the role it plays in judgment, where other group members decide whether the novel concept is worth pursuing, and whether it would address the problems they are facing. This judgment is easier to do with a built and ''manifest model'' (Nersessian & Chandrasekharan, 2009) than an internal model, because the manifest model allows group members to perform manipulations and thus form common movement representations of the proposed concept. The manifest model also improves group dynamics. One need not say ''you're wrong'' only that ''the model doesn't support that claim.''

What desirable cognitive features should such external models have to support discovery?

Based on the common coding framework, one desirable feature would be lots of movement—building and manipulation activity. Another feature would be a range of ways to generate movements in the built system. The movements need not always be visualizations. For instance, protein structure has been generated as music (Dunn & Clark, 1999), and scanning microscope output has been used to generate haptic feedback (Sincell, 2000).

A third desirable feature would be more control, to generate more fine-grained movements. Four, a closer integration between the movement generation mechanism and the internal model of the target system, similar to D11's use of the dish to display network activity.

4.5. *Objections to a movement-based approach to discovery*

In this section, I will consider a set of objections that have been raised to a movement-based explanation of the coupling between internal and external models, and the role of movements in discovery.

> In the common coding model, all transactions between internal and external models happen through motor activation. Are all the internal models of the world we possess coded in movement terms? What about objects and colors and labels? Static images and symbols (such as graphs and histology stains) do exist in science, and they provide information. How are they incorporated into internal models?

In the common coding view, the brain is a control mechanism that evolved to coordinate actions and movements in the world. There are two ways in which static images and symbols could be integrated with such a movement-based account. One is a model of representation where the ''standing-in'' relation between a symbol and the world arises out of actions and movements (see Chandrasekharan & Stewart, 2007). In this view, the static nature of a symbol is an illusion—the symbol is part of a dynamic sense-action network, and what appears static is the part that remains constant across actions and movements. A close metaphor would be the persistent core of a bee swarm or a tornado. A related possibility is that static images and symbols are starting points for internal simulations, as in the use of the Two-thirds power law to generate movement patterns from drawings (Viviani, 2002). This second option comes for free if the first view is accepted (for details see Chandrasekharan & Stewart, 2007). Generally speaking, any system designed to process dynamic structures can process static structures, for example, as equilibrium states or time slices. At the level of neural systems, there are recent efforts to explore how object movements could be coded by body movement areas (see Schubotz, 2007; Schubotz & von Cramon, 2004).

For our purposes here, it is highly likely that for a large majority of models in science and engineering, movement is a central component, given the focus on causation and co-variation, and the high use of dynamics as an explanatory device in science. This means external movements could lead to changes in such models via the ideomotor effect. As for objects and colors, Noe (2004) has made a persuasive case that vision, particularly object perception and color, requires movement components, either self-generated (such as eye movements or internal simulations) or object-generated. On labels, research into processing of concepts and sentences shows that processing of labels involves movement components. However, there could still be elements processed in a static, purely pattern-matching, fashion. The common coding view does not deny this possibility.

If any movement can activate the motor system, wouldn't all movement influence internal models?

In a general sense, all movements in the world do influence the brain, and the representations it contains. However, changes to specific internal movement modules (models) can occur only when these modules are active. For instance, the motor areas involved in dancing and piano playing are not usually active when dancers and piano players are, for example, driving. So the movements generated in driving are not influencing the dancing/piano-playing module. However, when they are watching a dance/piano concert, or planning one while driving, the module is activated. And any movements on stage, or movements encountered while doing the planning, have the ability to influence the internal model. Similarly, the external movements can contribute to the internal models of scientists and engineers only when they are imagining or interacting with the model.

The common coding experiments show only that movements outside and movements inside influence each other. The influence could be positive (as in speeding up mental rotation) or negative (as in slowing down mental rotation). Given this, why should construction, and external movements in general, always play a positive role?

In the view I have outlined, the process of construction serves to generate more movements than is possible in imagination. That by itself is always positive, as more movements help in overcoming functional fixedness. The negative role arises only in the interaction of the externally generated movement with the internal model. It is highly probable that such negative integration or interference does happen (as in the noise interpretation of bursts by D4), but as the construction process moves further, it leads to other movements, which can dislodge such interpretations. One way to think of the construction process is to think of it as working similar to a genetic algorithm, where the system can settle into local minima, but random mutations always dislodge the system from such states. The movements generated by the construction process are similar to these random mutations. Also, ultimately, all integrations of movements into the internal model are ''grounded'' by the building of the facsimile system, which need to exhibit the behavior of the target system. So misinterpretations are dislodged by this requirement as well. Note also that the building process involved in developing facsimile models just perturbs internal models; it does not guarantee new discoveries and insights. Perturbation is just a strategy—it is not a method—so it does not always have positive effects.

Does all building lead to new understanding?

No. Kirlik (1998) describes how humans built projectiles thousands of years before we understood dynamics and mechanics. If building always leads directly to understanding, ancient societies would have developed abstractions and theories that accelerated their progress, building more sophisticated projectiles. Since this did not happen, it is clear that sophisticated internal theoretical models need to exist for any building to contribute to understanding. It is the interplay between the building process and the internal model that leads to novel understanding and discovery.

5. Limitations of the common coding account

The model developed above is preliminary, and it does not account for a range of features of the building-to-discover strategy. I outline some of these limitations below:

1. The model is limited to domains involving movement, and it does not account for the generation of novelty in domains where built models or internal models are not based on movements, such as in decision making, logic, and probability.
2. The model does not account for the role of mathematical, logical, and causal knowledge in the development of new concepts. It also does not account for how new mathematical, logical, and causal relationships could be generated from the building process.
3. The model is highly limited by its use of perception of movement as the primary link between external and internal models. Because of this, it cannot account for the generation of novel concepts using simulations without visualizations (or other similar perceivable effects, such as sounds or processing speed).
4. The model does not account for the role of analogical reasoning in generating novel concepts. Although it is worth mentioning here that, at a very high level, common coding itself could be considered to have an analogical structure. It is a mapping mechanism focused on movements, translating movements in one domain into another in a streamlined fashion. In this view, common coding offers a more detailed account of how analogical transfer works at the cognitive and neural levels. This is similar to the earlier proposed relation between embodied cognition and common coding.

These limitations are significant, and accounting for them would require a detailed understanding of how language and other symbolic frameworks interact with the motor system. The representational aspect of the common coding model offers the possibility that connections could be built between motor simulations and symbolic frameworks. There is growing evidence of a link between language and motor simulations, particularly in relation to metaphors and concepts, though no links have yet been shown between motor simulations and mathematical concepts or logic. There is some emerging work on the connection between causality and the perception of force (Wolff, 2007), and representation of causality in language and the sense of force (Talmy, 1988). It is unclear how these connections relate to the motor system and the simulation process, but these present promising avenues for further exploration. Together with the common coding model and simulative model-based reasoning, further developments in these areas could provide the possibility of a unifying cognitive account of how novel concepts arise from building models.

6. Conclusion

Starting from simulative model-based reasoning, which proposes that internal models in science and engineering have movement components instantiated by action simulations, I have outlined a model of how building of facsimile models of phenomena leads to the development of new concepts. The model is based on the common coding framework, which proposes a common neural representation connecting execution, perception, and imagination of movements. This common representation leads to the action system covertly replicating external movements, such as generated by an external facsimile model. This replication generates a movement ''lingua franca'' that seamlessly connects internal simulations with external movements generated by built models, allowing the states of external models to change internal models directly. The generation of new concepts from the building process is explained in two steps. One, building a model involves generating new and fine-grained movements in imagination, but in ways limited by the building process. Two, these movements, together with the movements generated by the external model, ''perturb'' existing internal models in a constrained fashion, allowing the internal models to move away from standard movement grooves and generate new patterns. These new patterns lead to new concepts.

The strategy of building to discover control mechanisms will grow in importance in the future, especially since the problems tackled in interdisciplinary fields are getting more and more complex. We thus need an account of how the building strategy works, its effectiveness, limits, and possibilities. The common coding framework provides a possible way of approaching this thorny problem, though clearly a lot remains to be done.

Acknowledgments

This research was conducted under a grant from the U.S. National Science Foundation to Nancy Nersessian and Wendy Newstetter (DRL 0411825). I thank Dr. Nancy Nersessian for many insightful discussions on the role of models in scientific cognition, and for providing me the wonderful opportunity to be part of the CLIC (Cognition and Learning in Interdisciplinary Cultures) project. Her extensive and insightful comments and clear pointers have helped focus the ideas presented in this paper. I also thank members of the CLIC group, particularly Lisa Osbeck, Chris Patton, Ellie Harmon, and Todd Fennimore for wide-ranging discussions and feedback on the ideas presented here. I thank the anonymous reviewers for their detailed comments, which have contributed significantly to the paper. Reviewer 2 pointed out how the manifest model improves group dynamics. Any errors and opaqueness that remain are entirely my own.

References

Adamson, R. (1952). Functional fixedness as related to problem solving: A repetition of three experiments. *Journal of Experimental Psychology*, *44*, 288–291.

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*(3), 629–661.

Banzhaf, W. (1994). Self-organization in a system of binary strings. In R. Brooks & P. Maes (Eds.), *Artificial life IV* (pp. 109–119). Cambridge, MA: MIT Press.

Barsalou, L. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

Barsalou, L., Simmons, W., Barbey, A., & Wilson, C. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*, 84–91.

Becvar, L., Hollan, J., & Hutchins, E. (2007). Representational gestures as cognitive artifacts for developing theory in a scientific laboratory. In C. Halverson & M. Ackerson (Eds.), *Artifacts in workplace practice* (pp. 117–143). New York: Kluwer Academic Publishers.

Bergen, B., Chang, N., & Narayan, S. (2004). Simulated action in an embodied construction grammar. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual meeting of the cognitive science society* (pp. 108–113). Chicago. Hillsdale, NJ: Erlbaum.

Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Understanding another's expectation from action: The role of peripheral sensation. *Nature Neuroscience*, *8*, 1295–1297.

Brass, M., Bekkering, H., & Prinz, W. (2002). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*, 3–22.

Brass, M., & Heyes, C. (2005). Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*, 489–495.

Casile, A., & Giese, M. (2006). Non-visual motor learning influences the recognition of biological motion. *Current Biology*, *16*(1), 69–74.

Chandrasekharan, S., & Osbeck, L. (in press). Rethinking situatedness: Environment structure in the time of the common code. *Theory & Psychology*.

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

Craig, D. L., Nersessian, N. J., & Catrambone, R. (2002). Perceptual simulation in analogical problem solving. In L. Magnani & N. Nersessian (Eds.), *Model-based reasoning: Science, technology, and values* (pp. 167–191). New York: Kluwer Academic/Plenum Publishers.

Cree, G., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, *23*(3), 371–414.

Decety, J. (2002). Is there such a thing as a functional equivalence between imagined, observed and executed actions. In A. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution and brain bases* (pp. 291–310). Cambridge, England: Cambridge University Press.

Dunn, J., & Clark, M. (1999). Life music: The sonification of proteins. *Leonardo*, *32*(1), 25–32.
Edwards, L., Peng, Y., & Reggia, J. (1998). Computational models for the formation of protocell structure. *Artificial Life*, *4*(1), 61–77.
Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402.
Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (2000). Visuomotor neurons: Ambiguity of the discharge or ''motor'' perception? *International Journal of Psychophysiology*, *35*, 165–177.
Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, *73*, 2608–2611.
Farne, A., Iriki, A., & Làdavas, E. (2005). Shaping multisensory action-space with tools: Evidence from patients with cross-modal extinction. *Neuropsychologia*, *43*, 238–248.
Fiorio, M., Tinazzi, M., & Agiloti, S. M. (2006). Selective impairment of hand mental rotation in patients with focal hand dystonia. *Brain*, *129*, 47–54.
Gallese, V., Ferrari, P., Kohler, E., & Fogassi, L. (2002). The eyes, the hand and the mind: Behavioral and neurophysiological aspects of social cognition. In M. Bekoff, C. Allen, & M. Burghardt (Eds.), *The cognitive animal* (pp. 451–462). Cambridge, MA: MIT Press.
German, T., & Barrett, H. (2005). Functional fixedness in a technologically sparse culture. *Psychological Science*, *16*, 1–5.
German, T., & Defeyter, M. (2000). Immunity to functional fixedness in young children. *Psychonomic Bulletin and Review*, *7*(4), 707–712.
Gibbs, R. W. (2006). *Embodiment and cognitive science*. New York: Cambridge University Press.
Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge, England: Cambridge University Press.
Hegarty, M. (2004). Mechanical reasoning as mental simulation. *Trends in Cognitive Sciences*, *8*, 280–285.
Heyes, C. (2005). Imitation by association. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From cognitive neuroscience to social science* (pp. 157–174). Cambridge, MA: MIT Press.
Holt, L., & Beilock, S. (2006). Expertise and its embodiment: Examining the impact of sensorimotor skill expertise on the representation of action-related text. *Psychonomic Bulletin and Review*, *13*, 694–701.
Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–878.
Hurley, S., & Chater, N. (2005). *Perspectives on imitation: From neuroscience to social science. Imitation, human development and culture*. Cambridge, MA: MIT Press.
Hutchins, E. (1995a). How a cockpit remembers its speeds. *Cognitive Science*, *19*, 265–288.
Hutchins, E. (1995b). *Cognition in the wild*. Cambridge, MA: MIT Press.
Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurons. *Neuroreport*, *7*, 2325–2330.
James, W. (1890). *Principles of psychology*. New York: Holt.
Jeannerod, M. (1997). *The cognitive neuroscience of action*. Oxford, England: Blackwell.
Jeannerod, M. (2006). From volition to agency: The mechanism of action recognition and its failures. In N. Sebanz & W. Prinz (Eds.), *Disorders of volition* (pp. 175–192). Cambridge, MA: MIT Press.
Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.
Kirlik, A. (1998). The design of everyday life environments. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 702–712). Oxford, England: Blackwell.
Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.
Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Psychological Science*, *15*(3), 99–104.

Kosslyn, S. (1994). *Image and brain*. Cambridge, MA: MIT Press.

Langton, C. G. (1984). Self-reproduction in cellular automata. *Physica D*, *10*, 135–144.

Langton, C. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D*, *42*, 12–37.

Lenhard, J. (2004). Surprised by a nanowire: Simulation, control, and understanding. *Philosophy of Science*, *73*, 605–616.

Lenhard, J. (2006). Simulation and a revolution in modelling style: From hierarchical to network-like integration. In J. Lenhard, G. Küppers, & T. Shinn (Eds.), *Simulation: Pragmatic construction of reality* (pp. 89–106). Sociology of the Sciences Yearbook 25. Dordrecht, The Netherlands: Springer.

Martin, T., & Schwartz, D. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science*, *29*(4), 587–625.

Millet, G., Lecuyer, A., Burkhardt, J., Haliyo, D., & Regnier, S. (2008). Improving perception and understanding of nanoscale phenomena using haptics and visual analogy. *Lecture Notes in Computer Science*, *5024*, 847–856.

Nersessian, N. (2002a). The cognitive basis of model-based reasoning in science. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 133–153). Cambridge, England: Cambridge University Press.

Nersessian, N. (2002b). Maxwell and ''the Method of Physical Analogy'': Model-based reasoning, generic abstraction, and conceptual change. In D. Malament (Ed.), *Essays in the history and philosophy of science and mathematics* (pp. 129–166). Lasalle, IL: Open Court.

Nersessian, N. (2008). *Creating scientific concepts*. Cambridge, MA: MIT Press.

Nersessian, N., & Chandrasekharan, S. (2009). Hybrid analogies in conceptual innovation in science. *Cognitive Systems Research*, *10* (3) 178–188. Special Issue on Analogy.

Nersessian, N., Kurz-Milcke, E., Newstetter, W., & Davies, J. (2003). Research laboratories as evolving distributed cognitive systems. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th annual conference of the cognitive science society* (pp. 857–862). Boston: Cognitive Science Society.

Nersessian, N. J., & Patton, C. (2009). Model-based reasoning in interdisciplinary engineering. In A. Meijers (Ed.), *Handbook of the philosophy of technology and engineering sciences* (pp. 687–718). Amsterdam: Elsevier.

Newstetter, W., Kurz-milcke, E., & Nersessian, N. (2004). Cognitive partnerships on the bench tops. In Y. B. Kafai, W. A. Sandoval, N. Enyedy, A. S. Nixon, & F. Herrera (Eds.), *Proceedings of the 6th international conference on learning sciences* (pp. 372–379). Hillsdale, NJ: Erlbaum.

Noe, A. (2004). *Action in perception*. Cambridge, MA: MIT Press.

Oztop, E., Kawato, M., & Arbib, M. (2006). Mirror neurons and imitation: A computationally guided review. *Neural Networks*, *19*, 254–271.

Prinz, W. (1992). Why don't we perceive our brain states? *European Journal of Cognitive Psychology*, *4*, 1–20.

Prinz, W. (2005) An ideomotor approach to imitation. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science*, *vol. 1* (pp. 141–156). Cambridge, MA: MIT Press.

Prusinkiewicz, P., Lindenmayer, A., & Hanan, J. (1988). Developmental models of Herbaceous plants for computer imagery purposes. *Computer Graphics*, *22*(4), 141–150.

Reynolds, C. (1987). Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics*, *21*(4), 25–34.

Rheinberger, H.-J. (1997). *Towards history of epistemic things: Synthesizing proteins in the test tube*. Stanford, CA: Stanford University Press.

Runions, A., Fuhrer, M., Lane, B., Federl, P., Rollang-Lagan, A., & Prusinkiewicz, P. (2005). Modeling and visualization of leaf venation patterns. *ACM Transactions on Graphics*, *24*(3), 702–711.

Schubotz, R. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218.

Schubotz, R., & von Cramon, D. (2004). Sequences of abstract non-biological stimuli share ventral premotor cortex with action observation and imagery. *Journal of Neuroscience*, *24*(24), 5467–5474.

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1234–1246.

Sims, K. (1994). Evolving virtual creatures. *Computer Graphics*, *8*, 15–22.

Sincell, M. (2000). NanoManipulator lets chemists go mano a mano with molecules. *Science*, *290*, 1530.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*, 49–100.

Vingerhoets, G., de Lange, F., Vandemaele, P., Deblaere, K., & Achten, E. (2002). Motor imagery in mental rotation: An FMRI study. *Neuroimage*, *17*, 1623–1633.

Viviani, P. (2002). Motor competence in the perception of dynamic events: a tutorial. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance XIX* (pp. 406–442). New York: Oxford University Press.

Welsh, T., & Elliott, D. (2004). Movement trajectories in the presence of a distracting stimulus: Evidence for a response activation model of selective reaching. *Quarterly Journal of Experimental Psychology – Section A*, *57*, 1031–1057.

Welsh, T., Lyons, J., Weeks, D., Anson, G., Chua, R., Mendoza, J., & Elliott, D. (2007). Within-and between-nervous system inhibition of return: Observation is as good as performance. *Psychonomic Bullettin and Review*, *14*(5), 950–956.

Wexler, M., Kosslyn, S., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, *68*, 77–94.

Wilson, N., & Gibbs, R. (2007). Real and imagined body movement primes metaphor comprehension. *Cognitive Science*, *31*(4), 721–731.

Winsberg, E. (2003). Simulated experiments: Methodology for a virtual world. *Philosophy of Science*, *70*, 105–125.

Winsberg, E. (2006a). Handshaking your way to the top: Simulation at the nanoscale. *Philosophy of Science*, *73*, 582–594.

Winsberg, E. (2006b). Models of success vs. the success of the models: reliability without truth. *Synthese*, *152*(1), 1–19.

Wohlschlager, A. (2001). Mental object rotation and the planning of hand movements. *Perception and Psychophysics*, *63*(4), 709–718.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General*, *136*, 82–111.