CONCEPTUAL CHANGE: CREATIVITY, COGNITION, AND CULTURE

Nancy J. Nersessian
College of Computing
School of Interactive Computing
Georgia Institute of Technology
nancyn@cc.gatech.edu

1. Introduction

Concept formation and change—what I here call "conceptual innovation"—is one of the most creative dimensions of scientific practice. Throughout the history of the sciences changes in representational structure have provided "revolutionary" understandings of nature. As with other creative outcomes, conceptual revolutions are still widely perceived to be the outcomes of mysterious acts of individual genius, such as represented by an Isaac Newton, a Charles Darwin, or an Albert Einstein. The object of this paper is to dispel this notion by establishing how to incorporate both the undoubtedly unique contributions of individual scientists and the inherently socio-cultural nature of all scientific creations into the analysis of conceptual innovation. The route to meeting this objective lies in interpreting the conceptual practices scientists employ as deriving both from aspects of mundane human cognitive capabilities and from the social and cultural contexts, scientific and ordinary, in which they are embedded. What is required to construct such an interpretation is 1) knowledge of pertinent aspects of human cognition, 2) knowledge of specific practices implicated in cases of conceptual innovation, and 3) an understanding of how social and cultural contexts provide conceptual, analytical, and material resources that shape such practices.

2. Interpreting Conceptual Practices: Cognitive-Historical Analysis

In contemporary cognitive studies of science, the methodologies employed in investigating the practices scientists use in creating knowledge are ethnography, *in vivo* observation, laboratory experiments, and cognitive-historical analysis.

J. Meheus and T. Nickles (eds.), *Models of Discovery and Creativity*,
DOI 10.1007/978-90-481-3421-2_8, © Springer Science + Business Media B.V. 2009

Although it is possible to gain knowledge of conceptual practices by observing scientists in naturalistic settings, such as their own research laboratories, or by observing them in the setting of a problem solving experiment in the cognitive science laboratory, it is unlikely that conceptual innovation itself will be observed in these settings. It occurs infrequently and usually involves time spans longer than hours or days. Because of these facts, cognitive-historical analysis is the primary research method for investigating conceptual innovation (see, Nersessian, 1995).

Cognitive-historical analysis uses the customary range of historical records to recover how representational, methodological, and reasoning practices have been developed and used by scientists. These practices are studied over time spans of varying length, ranging from shorter spans defined by the activity itself to spans of decades or more. The records include notebooks and diaries, publications, correspondence, and material artifacts such as instruments. The historical practices are then examined in light of salient investigations of human representational and reasoning practices carried out by the fields within cognitive science. These comprise cognitive psychology, artificial intelligence, cognitive neuroscience, linguistics, cognitive sociology, and cognitive anthropology. One objective of cognitive-historical analysis is to explain the cognitive basis of the generativity of these practices. Some of the salient cognitive science research is directly on scientific cognition, but for the most part the studies are of cognition in mundane contexts. Saliency is determined by the nature of the practices under scrutiny. A "continuum hypothesis" underlies the cognitive-historical method: the cognitive practices of scientists are extensions of the practices humans employ in problem solving of a more ordinary kind within various physical and social environments. That is, human cognitive abilities give rise to and constrain scientific practices. Placing the historical practices within the broader framework of human cognitive activities aids in moving beyond the specific case study to more general conclusions about the nature and function of the scientific practices.

Margaret Boden (1990) makes a clarifying distinction between "P-creative" ideas that arise from episodes in which an individual creates something culturally available, but novel for the individual in question, and "H-creative" ideas, arising from episodes in which something fundamentally new in human history is created. Boden focused her attention on the nature of the mechanisms that lead to P-creative ideas. The primary foci of ethnographies, observations, and psychology experiments are the practices scientists use in coming to learn, appropriate, and employ existing concepts. The kind of conceptual change examined in these studies is primarily P-creative, that is, their novelty is for the individual. It is "H-creative" conceptual change I am concerned with here, that is, conceptual innovations with historical impact in that they have changed existing representations of nature. However, as Boden notes, some P-creative ideas are, of course, also H-creative. The hypothesis of continuity between mundane and scientific cognition that underlies the cognitive-historical method incorporates the possibility that mechanisms implicated in P-creative instances of conceptual innovation can also be employed in H-creative instances. For example, analogy could be a generative mechanism in both kinds of innovation (Gentner et al., 1997; Nersessian, 1984, 1992a). Thus, the findings about the conceptual practices of scientists derived from the other methodological approaches in science studies are relevant to developing a cognitive-historical analysis.

In addressing the problem of conceptual innovation, the historical practices can be investigated at the level of individuals and at the level of communities. The practices of designing and executing experiments, constructing models, using mathematical tools, devising means of communicating, and training practitioners, are all relevant to understanding the nature of conceptual innovation in science. A full analysis sets these in the social and cultural contexts of training, earlier research, knowledge base, community, collaborators, competitors, and material resources. Cognitive-historical investigations of conceptual change can focus ideographically, attempting to ferret out general cognitive factors underlying the uniquely individual dimensions of practice (see, e.g., Gooding, 1990; Nersessian, 1984, 1985, 1992a, 2002; Tweney, 1992). They can also focus on practices common to many instances with the intent to formulate a general account of how it is possible they produce the outcomes (see, e.g., Darden, 1991; Nersessian, 1992a; Thagard, 1992; Tweney, 1985). In both cases the source and generativity of such practices is located in what is, generally speaking, human. On the one hand, what is human includes those cognitive structures and capabilities humans have in common—that enable and constrain the unique application of an individual scientist. On the other hand, what is human is embeddedness in social and cultural systems. To date the focus of cognitive-historical analyses has been on the cognitive capabilities, structures and processes. Investigations of these have largely drawn from research in cognitive science within the traditional "GOFAI" ("Good Old Fashioned AI") framework.

On the traditional view, cognition comprises the representations internal to an individual mind and the processes that operate on these. Thinking is independent of the medium in which it is implemented, and the environment is represented in the content of thinking through being represented in memory. Recently, these founding assumptions of cognitive science were elaborated upon extensively by Alonso Vera and Herbert Simon (1993) in response to criticisms from within cognitive science. Following earlier work by Alan Newell and Simon (1972), the unit of analysis in studying cognition is called a "physical symbol system" (PSS). A PSS has a memory capable of storing and retaining symbols and symbol structures, and a set of information processes that form structures as a function of sensory stimuli. It makes no difference to understanding cognition whether the symbol processing is carried out by a human, a computer, or any other kind of PSS. In humans, and any natural or artificial PSS with sensory receptors and motor action, sensory stimuli produce symbol structures that cause motor actions and modify symbol structures in memory. Thus, a PSS can interact with its environment by 1) receiving sensory stimuli from it and converting these into symbol structures in memory and 2) acting upon it in ways determined by the symbol structures it produces, such as motor symbols. Perceptual and motor processes connect symbol systems with the environment and provide the semantics for the symbols. So, social and cultural environments are treated as abstract content on which cognitive processes operate. As with Simon's earlier "parable of the ant" (Simon, 1981, pp. 63–66), the complexity in human behavior is understood to arise from acting in the environment. So clearly, social and cultural factors are important to understanding cognition. However, the traditional contention is that what is important about the environment for thinking is abstracted through perception and represented in the symbols generated by the cognitive system. So, these dimensions need only be examined as residing internal to the mind of a human individual or other PSS as socio-cultural *knowledge*. One implication is that it makes little difference to understanding cognition whether the thinking is carried out in an authentic environment or in a psychological research laboratory.

The reductionism of the traditional account of cognition has led those on the social side of science studies to perceive social and cognitive accounts of science as fundamentally incompatible. Social accounts have tended to "black box" the individual entirely or to render cognitive explanatory factors inconsequential in comparison with socio-cultural factors. Indeed, the perceived *in-principle* incompatibility of cognitive and social accounts of science has led some in science studies to position themselves in opposition to cognitive analyses, as witnessed, e.g., by the now-expired "ten-year moratorium" on cognitive explanations called for by Bruno Latour and Steven Woolgar (Latour and Woolgar, 1979; Latour, 1987). As with the traditional view in cognitive science, this anticognitive stance has roots in the remnants of Cartesian dualism. The mind/body, individual/social, and internal/external dichotomies associated with Cartesianism are all in play in this stance. A "cognitive explanation" is seen as tantamount to maintaining the epistemological position that the source of knowledge is ideas internal to the mind (Latour, 1999).

What must be kept in mind in discussing scientific cognition, though, is that "thinking" is an inherently social and cultural activity. It rarely just goes on "in the head" in isolation from physical and social interactions. Even when a solitary thinker wrestles with a problem closed in her study, she is still engaged in a socio-cultural process. Educational training is present. Conversations with colleagues are recalled. Further, the process often involves external representation such as sketches and equations that are socio-cultural in origin. In science *what* one thinks about and *how* one thinks about it are highly dependent on one's socio-cultural environment. Take, for example, the quest for an electromagnetic theory by British and French scientists in the latter half of the 19th century. Their representations of the problem and their methods of analysis differed considerably. To understand how Maxwell derived the mathematical equations requires knowing that he was trained in Scotland in the methods of physical geometers and in Cambridge as a mathematical physicist; that he was located in a milieu that valued Faraday's theoretical speculations, which included teachers and colleagues such as Thomson and his penchant for analogical models; and that he was located in Victorian Britain with, among other factors, the cultural fascination with machines and mechanisms. These socio-cultural factors co-determined the nature of the theoretical, experimental, and mathematical knowledge and the methodological practices with which Maxwell formulated the problem and approached its solution. They are reflected in Maxwell's reasoning through mechanical models in deriving the equations, and one cannot understand his construction of these equations without taking these factors into account. Of course, now we re-derive them by different means. Continental physicists working on electromagnetism at the time, including the French physicist Ampère, employed quite different analytical practices and drew from fundamentally different theoretical assumptions and mathematical and physical representational structures. Differences in socio-cultural factors figure significantly into why members of these communities were not able to derive the field equations.

Clearly to produce scientific knowledge requires both sophisticated cognition and a rich socio-cultural environment. The objectives of cognitive-historical analysis are to determine what enables individual agency, while at the same time explaining how the products of individuals are communal products and how these products are transformed into communal resources, transported out of the specific localities of their construction into the accepted representational content of science. To carry this out, the difficult problem that needs to be addressed is how the cognitive, the social, and the material are fused in the processes of scientists' constructing knowledge of the world. One starting point that has significant potential for resolving the problem is reconceptualizing the notion of "cognition" along the lines of recent non-reductionist analyses that challenge traditional framing of the notion. I turn to these in the next section.

3. Cognition and Culture: Situated and Distributed Cognition

Those not engaged in or with cognitive science in the last several years continue to identify it exclusively with the "rules and representations" or "logicist" accounts of human cognition associated with "GOFAI" that initiated the "cognitive revolution". As James Greeno (1989a) points out, a framing assumption of that revolution was that "the locus of thinking" is assumed to be in an individual agent's mind, rather than in interaction between an agent and a physical or social situation" (p. 134). The founding "functionalist" assumption was that thinking or intelligence is an abstractable structure that can be implemented in many different media, independent of physical or social context. Although there are still many adherents to these assumptions, contemporary cognitive science possesses alternative accounts of reasoning, representation, and learning and richer, more contextualized studies of human cognition that have yet to be exploited by science studies. Where these accounts intersect, "cognition refers not only to universal patterns of information transformation that transpire inside individuals, but also to transformations, the forms and functions of which are shared among individuals, social institutions, and historically accumulated artifacts (tools and concepts)" (Resnick et al., 1991, p. 413).

Investigations into "situated" and "distributed" cognition focus not only on the individual but also on the social group and on the various cultural artifacts and symbol systems involved in cognitive processes. It brings these directly into the purview of research on the customary cognitive science topics of representation, problem solving, and learning. The locus of analysis is always an "activity" and the unit of analysis of an activity is a "cognitive system". Analysis of a cognitive system can focus on an individual (reconceptualized as an embodied, social, tool-using agent), a group, or the material and conceptual artifacts of the context of an activity. The goal, however, is to understand cognition as an interaction among the participants in and context of an activity as it develops over time. Much of the research in this area is not conducted in standard laboratory settings, but focuses on cognitive activities in inherently social and collective contexts: in learning situations (see, e.g., Brown et al., 1989; Greeno, 1989a; Lave, 1988), in the workplace (see, e.g., Suchman, 1987; Woods, 1997), and "in the wild" (the world at large) (see, e.g., Hutchins, 1995; Norman, 1988; Shore, 1997). Further, much of it is concerned with how meaning and understanding is created collectively and addresses, directly, the problem of how cultural representations that are variable and context-relative could have universal properties of the human mind implicated in their development.

These challengers of GOFAI argue that the traditional view has mistaken the properties of a complex, *cognitive system*, comprising both the individual and the environment, for the properties of an individual mind. Thus, the critique is aimed at the traditional analytical framework in which cognitive processes are treated separately from the contexts and activities in which cognition occurs. For example, in arguing for a distributed notion of cognition, Edwin Hutchins (1995) contends that rather than construing culture as *content*, what is required is an integrated picture in which cognition and culture are interrelated notions construed in terms of *process*. Such construal leads to a shift in theoretical outlook from regarding cognitive and socio-cultural factors as independent variables to regarding cognitive processes as inherently socio-cultural. Thus the main point of contention is not whether the environment can be accommodated, but rather, *whether accounting for environmental factors requires altering fundamental notions of the structures and processes employed in cognition*. The argument is about the very nature of cognition itself.

Broadly characterized, the challenge posed to the traditional view centers on three interrelated issues: 1) the limitation of the cognitive system to the bounds of the individual mind, 2) the nature of the processing employed in cognition, and 3) the nature of, and the need for, mental representations in cognitive processing. The literature that addresses these issues is by now quite extensive and there are significant differences within and among the perspectives, so it will not be possible to lay out the positions in detail. Rather, I will highlight features of these views that seem pertinent to interpreting scientific cognition. I begin by discussing the "situative perspective" (Greeno, 1989a) and then link aspects of the other perspectives that are salient for our purposes.

Much of the impetus for developing theories of *situated cognition* has come from studies by cognitive anthropologists and sociologists concerned with learning and with work practices. Jean Lave, for instance, has attempted to explain ethnographical studies that establish striking disparities between mathematical problem solving, competency in real-world and in school learning environments. In real-world environments, such as supermarkets (Lave, 1988), adults and children exhibit high levels of competence in solving mathematics problems that are structurally of the same kind as those they fail at solving in standard school and test formulations. Lave argues that the disparities can be explained only by construing the relation between cognition and action as an interactive process that involves essentially the resources available in a specific environment. Cognition is a relation between the individual and the situation and does not just reside "in the head". Drawing on J. J. Gibson's theory of perception (Gibson, 1979), explanations of human cognition in the situative perspective employ the notion of *attunement to constraints and affordances*. On their adaptation of Gibson's notion, an affordance is a resource in the environment that supports an activity and a constraint is a regularity in a domain that is dependent upon specific conditions.

The structure of the environment provides the constraints and affordances needed in problem solving and these cannot be captured in abstract problem representations alone. Ethnographical studies of work environments by Lucy Suchman (1987), for example, have led her to argue that contrary to the traditional cognitive science view that problem solving involves formulating in the abstract the plans and goals that will be applied in solving a problem, plans and goals develop in the context of actions and are thus emergent in the problem situation. Problem solving involves improvisation and appropriation of affordances and constraints in the environment, rather than mentally represented goals and plans specified in advance of action.

Within the situative perspective, analysis of a cognitive system can focus at different levels: on the individual (conceptualized as an embodied, social, tool-using agent), a group of agents, or the material and conceptual artifacts of the context of an activity. The goal of an analysis at any level, though, is to understand cognition as an interaction among these participants in and, the context of, an activity. Cognition, thus, is understood to comprise the interactions between agents and environment, not simply the possible representations and processes in the head of an individual. Thus situated cognition is *distributed*.

As with the situative perspective, proponents of the notion of *distributed cognition* contend that the environment provides a rich structure that supports problem solving. The focus of distributed cognition research is on the claim that an environment does not just supply "scaffolding" for mental processes, as it is viewed in the traditional perspective, but that salient parts of the environment are an integral part of the cognitive system and, thus, enter essentially into the analysis of cognition. Thus they contend that a new account of cognitive processing is required—one that incorporates what is salient in the environment in a non-reductive fashion. Salient parts of an environment are, broadly characterized, those factors that can affect the outcome of an activity, such as problem solving. These cannot be determined *a priori* but need to be judged with respect to the instance. For ship navigators, for example, the nature of the function of a specific instrument can be salient, but not usually the material from which the instrument is made. For physicists, whether one is sketching on a blackboard or white board or piece of paper is likely irrelevant, but sketching on a computer screen might be salient because the computer adds resources that can affect the outcome.

Determining the *cognitive artifacts* within a specific system is a major part of the analytical task for the distributed perspective. Various kinds of external representations are candidates. Zhang & Norman (Zhang and Norman, 1995; see also Zhang, 1997), for example, have studied problem solving with isomorphic problems to ascertain potential cognitive functions of different kinds of external representations. They found that external representations differentially facilitate and constrain reasoning processes. Specifically, they argue that diagrams are cognitive artifacts in that they do not play just a supportive role in what is essentially an internal process, but that these external representations play a direct role in cognitive processing without the mediation of an internal representation of the information provided in them. On their account, affordances and constraints in the environment are construed, literally, as memory in cognitive processing. Thus, analysis of cognition in situations of problem solving with diagrams needs to be of the cognitive system that comprises both the mental and diagrammatic representations.

Research in the situative and distributed perspectives largely consists of observational case studies employing ethnographic methods. Although these studies focus on details of particular cases and often provide "thick descriptions" of these (Geertz, 1973), their objectives differ from historical, social, and cultural studies in STS that aim mainly to ferret out the specific details of a case. Rather, the aim of the cognitive research is to understand the nature of the regularities of cognition in human activity, i.e., those aspects that are common across cases. As Hutchins has framed the objective

> There are powerful regularities to be described at the level of analysis that transcend the details of the specific domain. It is not possible to discover these regularities without understanding the details of the domain, but the regularities are not about the domain specific details, they are about the nature of cognition in human activity. (Woods, 1997, p. 171)

Currently there are many research undertakings in cognitive science that share the objective of furthering an account of cognition that construes cognition and environment in relation to one another. These include research in a wide range of areas, including the embodied nature of mental representation and cognitive development in children and animals. At present there is little or no dialog among many of these. Research in each of these areas is very much research in progress, so it tends to focus internally to an area, with not much interaction across them. It is not possible to survey the various research areas that I see as comprising a body of interconnected research in the context of this paper. Instead I will focus on one issue: how culture might shape the very nature of cognitive capacities, structures, and processes.

Comparative studies in primatology and on cognitive development have led Michael Tomasello (Tomasello, 1999; Geertz, 1973), among others, to contend that cognition is inherently cultural. He argues that culture is central to the development of uniquely human cognitive abilities, both phylogenetically and ontogenetically. He begins by posing the problem of the origins of these abilities. In terms of biological evolution, the time span is just too short to account for the vast cognitive differences that separate humans from the primates closest to us genetically, e.g., chimpanzees. From comparative studies of ontogenesis in human children and other primates, he posits that the development of the uniquely human cognitive abilities began with a small phylogenetic change in the course of biological evolution: the ability to see conspecifics as like oneself, and thus to understand the intentionality of their actions. This change has had great consequences in that the processes of imitation and innovation enabled by it allowed for the accumulation of culture through transmission—or what he calls "cultural evolution".

On the account Tomasello develops, cultural evolution is the engine of cognitive evolution. That is, the expansion of cognitive capacities in the human primate has occurred as an adaptation to culture. Significantly then, culture is not something added to accounts of cognition—culture is what makes human cognition what it is. In ontogenesis, children absorb the culture and make use of its affordances and constraints in developing perspectivally-based cognitive representations. His analysis concentrates specifically on how language development creates cognitive capacities in the processes of ontogenesis, which supports the early speculation of Lev Vygotsky (1978) (whose work has influenced the development of the situative perspective discussed above) that cognitive development is socio-cultural in that it involves the internalization of external linguistic processes. However, this does not imply that cognitive processing need be all internal or linguistic. External representations seem indispensable in complex human thinking, and their development has been central to the process of cultural transmission. Merlin Donald's (1991) analysis of the evolutionary emergence of distinctively human representational systems underscores the importance of mimesis, or re-creation such as using the body to represent an idea such as the motion of an airplane, in the developments of such external representations as painting and drawing (40K years ago), writing (6K) and phonetic alphabets (4K). Donald argues for a distributed notion of memory as a symbiosis of internal and external representation on the basis of changes in the visuo-spatial architecture of human cognition with the development of external representation. So, affordances and constraints in the environment are *ab initio* part of cognitive processing.

This research into the relations between culture and cognitive development, along with developmental research in neuroscience can be construed as moving beyond the old "nature–nurture" debate through developing an *interactionist* approach. It attempts to provide an account of how evolutionary endowment and socio-cultural context act together to shape human cognitive development. Supporting this conception, neuroscience studies of the impact of socio-cultural depravation, enrichment, and trauma in humans and in non-human primates on brain structure and processes lead to a conception of the brain as possessing significant cortical plasticity (see, e.g, Elman et al., 1998; van der Kolk et al., 1996; Shore, 1997) and as a structure whose development takes place in response to the socio-cultural environment as well as genetic factors and biological evolution.

Finally, in so connecting cognition and culture, this body of research implies that human cognition should display both species-universal cognitive abilities and culturally specific cognitive processes. Tomasello discusses some of the universal learning abilities, such as those connected with language learning, among which he includes the ability to understand communicative intentions, to use role reversal to reproduce linguistic symbols and constructions, and to use linguistic symbols to contrast and share perspectives in discourse interactions (Tomasello, 1999, pp. 161–163). Although he does not discuss these, one implication is that the cognitive processes of learning, reasoning, problem solving, representation, decision making should display culturally specific features. Recent investigations into culturally specific features of cognition by Richard Nisbett and colleagues (Nisbett et al., 2001) has implications for the hypotheses linking cultural evolution and cognitive processes. This research was inspired by the substantial body of historical scholarship that maintains that there were systematic cultural differences between ancient Greek and Chinese societies, especially concerning what they call the "sense of personal *agency*" (p. 292, italics in original). Nisbett hypothesized that these differences between "eastern" and "western" cultures, broadly characterized as holistic vs. analytic thinking (p. 293), should still be detectable in cognitive processes such as categorization, memory, covariation detection, and problem solving in contemporary cultures whose development has been influenced by ancient China (China, Japan, Korea) or by ancient Greece (Western Europe, North America). In a series of experiments with subjects in East Asian and Western cultures, and subjects whose families have changed cultural location, they examined explanations, problem solving, and argument evaluation. Some significant systematic differences were found along the five dimensions they identified in the ancient cultures: 1) focusing on continuity vs. discreteness, 2) focusing on field vs. object, 3) using relations and similarities vs. categories and rules, 4) employing dialects vs. logic and first principles in reasoning, and 5) using experience-based knowledge vs. abstract analysis in explanations.

The implications of the research of the "environmental" perspectives reviewed above for the project of an integrative account of knowledge-producing practices in science are extensive. Working them out in detail is beyond the scope of this paper. One thing is clear though: situating the problem of interpreting these practices within the framework provided by environmental perspectives on cognition affords cognitive-historical analysis the possibility of analyzing from the outset the cognitive practices of scientists as bearing the imprint of human cognitive development, the imprint of the socio-cultural histories of the specific localities in which science is practiced, and the imprint of the wider societies in which science develops.

4. Creativity in Conceptual Change: The Role of Model-Based Reasoning

As discussed in Section 2, the continuum hypothesis underlying cognitive-historical analysis holds that cognitive practices of scientists are extensions of the kinds of practices humans employ in coping with their environment and in problem solving of a more ordinary kind. The mental representations and processes used in human problem solving have developed out of an interaction between two inseparable processes: biological selection and adaptation and socio-cultural construction, selection, and adaptation. Thus, scientific cognition is shaped by the evolutionary history of the human species and by the developmental processes of the human child. Basic cognitive strategies are extended and refined in explicit and critically reflective attempts to devise methods for understanding nature. As with mundane modes of inquiry, the success of those created by science is rooted in human nature and the nature of the world.

What needs to be ascertained are the nature of the representations and of the processes employed in scientific cognition. Here I will focus on a specific kind of problem solving practice employed in conceptual innovation: "model-based" reasoning. The issue of the representational format of conceptual structures is especially significant for the problem of the nature of the reasoning through which inferences are made. Different representational formats enable different modes of reasoning. The predominant modes of analysis of conceptual change have viewed conceptual structures from the perspective of languages. Clearly concepts and conceptual structures can be represented linguistically. However, in earlier cognitive-historical analyses of conceptual change, I have proposed that from the perspective of understanding the reasoning practices leading to new concepts, conceptual structures are best viewed as models and conceptual change as a process of constructing and communicating new models. This proposal derives from extensive examination of scientific practices leading to conceptual innovation. This examination establishes, first, that conceptual innovation is a problem-solving process, and, second, that model-based reasoning practices, such as analogy, visual modeling, and thought experimenting (simulative modeling) (Nersessian, 1984, 1992a, 1992b, 1999, 1988), play a central role. My analyses draw from practices employed in physics, but investigations of other sciences establish that these practices are employed across the science*s* (see, e.g. Darden, 1991; Giere, 1988, 1992; Griesemer, 1991; Griesemer and Wimsatt, 1989; Latour and Woolgar, 1979; Latour, 1987; Lynch, 1985; Lynch and Woolgar, 1990; Shelley, 1996, 1999; Thagard, 1992). Although model-based reasoning practices are ubiquitous, I am, of course, not contending they are exhaustive of the practices that generate new representational structures. I have focused on these practices because they are ubiquitous and because within philosophy these practices have not traditionally been considered significant forms of scientific reasoning, even though there is abundant historical evidence in favor of their generativity. Philosophical accounts of scientific reasoning have restricted the notion of reasoning primarily to deductive and inductive arguments. Modeling practices, when considered at all, have mainly been held to perform an ancillary role as "mere aids" to reasoning. The approach taken here is to develop a cognitive basis for these practices as productive forms of reasoning more widely applicable in human reasoning than in the specific contexts in which they are employed in science. From this basis, one can mount a case for how they are productive forms of reasoning in science and how they function in conceptual innovation.

Most of the work on representation and reasoning in the cognitive sciences comes from considering individual cognition from the traditional perspective. Here I want to place scientific cognition within the framework of the environmental perspective discussed in Section 3 by starting from the assumption that scientific cognition is always situated and often distributed. However, since individual human agents are parts of cognitive systems an accounting of their role in the cognitive processing within the system is still required. Mainstream notions of mental representation, such as concepts and mental models, and of reasoning, such as analogy, can contribute to understanding the human component, with the caveat that modification will be necessary. Thinking about such notions from the perspective of cognition as situated and distributed can aid in creating alternative formulations of these. The most radical proponents of situated cognition discount the role of mental representations in cognitive processes. However, although one might not need to invoke the notion of mental representation in explaining how people drive cars around a familiar campus or measure a quantity of cottage cheese to be eaten on a diet program, it is difficult to see how one could begin to explain complex scientific problem solving without invoking it. Much of the research in distributed cognition seems consistent with the notion of mental representation. However, what kinds of mental representations and processes to accord the individuals that constitute significant components of cognitive systems remains an outstanding research problem. This section begins to address this problem in conjunction with the hypothesis that "model-based" reasoning is central in conceptual innovation. The cognitive hypothesis of reasoning through "mental modeling" is a significant component of the case for the cognitive basis of model-based reasoning. I will try to establish that a particular notion of mental modeling is more in accord with the situated and distributed nature of scientific cognition.

4.1 Mental Modeling

The notion of a "mental model" is an explanatory construct that plays a central role in much of cognitive science. It is employed widely in theories of comprehension and reasoning. In cognitive psychology there is an ongoing controversy about the nature of human reasoning that parallels the issues raised about reasoning in philosophy. On the traditional psychological view, reasoning consists of applying a mental logic to propositional representations. Critics of this view have contended that a purely syntactical account of reasoning cannot account for significant effects of semantic information exhibited in experimental studies of reasoning (see, e.g., Johnson-Laird, 1983; Mani and Johnson-Laird, 1982; McNamara and Sternberg, 1983; Oakhill and Garnham, 1996; Perrig and Kintsch, 1985; Wason, 1960, 1968). Instead, they propose adopting a hypothesis that in many instances people reason by manipulating internal models. Advocates of the mental modeling hypothesis argue that the original capacity developed as a means of simulating possible ways of maneuvering within the physical environment. It would be highly advantageous to possess the ability to anticipate the environment and possible outcomes of actions, so it is likely that many organisms have the capacity for some form of mental modeling. Given their linguistic abilities, humans should be able to create models from both perception and description, which is borne out by the research in narrative comprehension. The centrality of mental modeling to cognition is a hypothesis under investigation in numerous domains including: reasoning about causality in physical systems (see, e.g., de Kleer and Brown, 1983); the role of representations of domain knowledge in reasoning (see, e.g., Gentner and Gentner, 1983); logical reasoning (see, e.g., Johnson-Laird, 1983); narrative comprehension (see, e.g. Perrig and Kintsch, 1985); induction (see, e.g., Holland et al., 1986); and problem solving by contemporary scientists (see, e.g., Chi et al., 1981; Clement, 1989; Griffith et al., 1996). Further, a range of empirical investigations can be garnered in support of mental models as vehicles of cultural transmission, such as those into "prototypes" in concept representation (see, e.g., Rosch and Lloyd, 1978), "idealized cognitive models" in language understanding (see, e.g., Lakoff, 1987), and mental modeling in cultural transmission (see, e.g., Shore, 1997). Because the potential range of application is so extensive, some have argued that the notion can provide a unifying framework for the study of cognition (Gilhooly, 1986). For our problem, too, the hypothesis is attractive because it opens the possibility of furnishing a unified analysis of the widespread modeling practices implicated in conceptual change.

Philip Johnson-Laird (1983) credits the philosopher, psychologist, and physiologist Kenneth Craik (1943) with introducing the notion of mental modeling. Craik hypothesized that in many instances people reason by carrying out thought experiments on internal models, where a model is a structural or functional analog to a real-world phenomenon:

> By a model we thus mean any physical or chemical system which has a similar relation-structure to that of the process it imitates. By 'relation-structure' I do not mean some obscure non-physical entity which attends the model, but the fact that it is a physical working model which works in the same way as the process it parallels, in the aspects under consideration at any moment. Thus, the model need not resemble the real object pictorially; Kelvin's tide-predictor, which consists of a number of pulleys on lever, does not resemble a tide in appearance, but it works in the same way in certain essential respects..." (Craik, 1943, pp. 51–52)

Craik maintained that just as humans create physical models, such as, physical scale models of boats and bridges, to experiment with alternatives, so too the nervous system of humans and other organisms developed a way to create mental "'small scale model[s]' of external reality" (p. 61) for simulating potential outcomes of actions in a physical environment. Mental simulation occurs by the "excitation and volley of impulses which parallel the stimuli which occasioned them…" (p. 60). This internal process of reasoning results in conclusions similar to those that "might have been reached by causing the actual physical processes to occur" (p. 51). Craik based his hypothesis on the need for organisms to be able to predict the environment, thus simulation is central to mental modeling. In constructing the hypothesis he drew on existing research in neurophysiology and speculated that the ability "to parallel or model external events" (p. 51) is fundamental to the brain.

In the first place, a mental model is a form of knowledge organization. There are two main usages of the term 'mental model' that tend to get conflated in the literature: (1) a structure stored in long-term memory (LTM) and (2) a temporary structure created in working memory (WM) during a reasoning process. The first usage focuses on how the mental representation of knowledge in a domain is organized in LTM and the role it plays in supporting understanding and reasoning. The second usage focuses on the nature of the structure employed in WM in a specific comprehension and reasoning task. In considering model-based reasoning, our analysis can be restricted to WM representations and processes. This usage maintains that mental models are created and manipulated during narrative and discourse comprehension, deductive and inductive logical reasoning, and other inferential processes such as in learning and creative reasoning. In all cases, the inferencing takes place through specific operations on the model itself. Although Philip Johnson-Laird's own research focus has been on deductive and inductive reasoning tasks, and not mental modeling in other kinds of inferencing, his 1983 book provides a general theoretical treatment of mental models as WM structures that has had a wide influence. He holds that a mental model is a structural analog of a real-world or imaginary situation, process or event that the mind constructs in WM during reasoning. A mental model is a structural analog in that it embodies a representation of salient spatial, temporal, and causal structures relating the events or entities represented. The LTM knowledge drawn upon in the activity of mental modeling need not be represented in the form of a model. Johnson-Laird's account is uncommitted on the format of the LTM representation.

Although talk of mental modeling is ubiquitous in cognitive science today, unfortunately explicit accounts of just what a specific researcher means when invoking the notion are not. There is not a single fully-developed and agreed upon hypothesis about either the representational *format* of the model, where 'format' includes *structure* and *content*, or the nature of the *processing* involved in either generating a model or reasoning by means of it. So, the notion of understanding and reasoning via mental modeling is best considered as an explanatory framework under development for studying cognitive phenomena. What the various hypotheses within the framework share is that they postulate models as organized units of mental representation on which cognitive processing is carried out in diverse activities. The preponderance of research into mental modeling is concerned with either explaining logical inferencing or specifying the knowledge contained in the models in a specific domain with respect to a reasoning task or level of expertise, and not with either the format or processing issues. Here I will try to classify the major views on the format and processing issues that can be discerned from the literature.

Preliminary to discussing the issues of format and processing with respect to mental modeling, some sorting out of the terminology used in discussing mental representation, generally, will be useful. Since its inception, there has been a deep divide in the field of cognitive science between those who hold that all representation is language-like or 'propositional' (see, e.g., Fodor, 1975; Pylyshyn, 1981) and those who hold that at least some mental representation is perceptual or 'imagistic' in format (Kosslyn, 1980, 1994; Shepard and Cooper, 1932). Herbert Simon (1977) reported that this divide "nearly torpedoed the effort of the Sloan Foundation to launch a major program of support for cognitive science" (p. 385) at the inception of the field. Even though significant clarification of the issues has taken place and considerable experimental work conducted, the issue remains unresolved and most likely will continue to be until more is known about how the brain functions.

In much of the cognitive science literature 'propositional' is often treated as co-extensive with 'symbolic', comprising language-like and perceptual representations. Here I employ the term in its narrower philosophical usage of a language-like mental encoding that possesses a vocabulary, grammar, and semantics, such as Fodor's language of thought (Fodor, 1975). A propositional representation is interpreted as referring to physical objects, structures, processes, or events descriptively. The relationship between this kind of representation and what it refers to can be evaluated as being true or false. I will use the term 'iconic', rather than 'imagistic', for different kinds of analog representations, so as not to conflate these representations with mental pictures, which are only one kind of iconic representation. Iconic representation can be highly abstract and schematic, and not picture-like at all. What differentiates an iconic representation from a propositional one is that along some dimension(s) constraints are represented in a manner that is intended as isomorphic to its real-world analog. This is how I interpret Craik's notion of a 'relation-structure'. So, for example, a mental model of the tide might only capture functional constraints as does Kelvin's real-world analog predictor. Iconic representations represent spatial, temporal, causal, and functional information in analog format and procedures for constructing and manipulating the various kinds of representations. An iconic representation is interpreted as representing objects, structures, processes, or events demonstratively. The relationship between this kind of representation and what it represents is that of "similarity".

Iconic representations are similar in aspects and degrees to what they represent, and thus can be evaluated as being accurate or inaccurate.

Because different kinds of representations enable different kinds of processing operations, propositional and iconic models support reasoning in different ways. Operations on propositional models include the customary logical and mathematical manipulations. The operations are rule-based and are truth-preserving if the symbols are interpreted in a consistent manner and the properties they refer to are stable in the environment. Additional operations can be defined in limited domains provided they are consistent with the constraints that hold in the domain. Manipulation of a model requires explicit representation of salient parameters, including structural constraints and transition states. Operations on iconic models involve transformations of the representations that change their properties and relations in ways consistent with the real-world constraints of the domain. Unlike propositional models, transformational constraints for iconic models can be implicit. For example, a person could perform simple simulative reasoning about what happens when a rod is bent without having an explicit rule, such as "given the same force a longer rod will bend farther", by employing constraints implicit in perceptual experiences.

The nature of the symbols that constitute the content of a mental model is important to processing issues. The distinction Lawrence Barsalou (1999) makes between 'amodal' and 'modal' symbols in discussing mental representation, generally, provides some clarification for thinking about mental models. Amodal symbols are arbitrary transductions from perceptual states, such as those associated with language. All propositional representations are composed of amodal symbols. Modal symbols are analog to the perceptual states from which they are extracted. Although perceptual in nature, modal symbols can be highly schematic. A cat-like image would be a modal symbol, ranging from an image of Fifi with her stripes to a more generic representation containing salient perceptual elements of 'catness' without definite feature such as stripes. The strings of letters 'cat' or 'chat' or 'Katze' are amodal symbols. Iconic representations can be composed of either. For example, a representation of the situation "the circle is to the left of the square, which is to the left of the triangle" could be composed of either modal tokens ● – ■ – ▲ or amodal tokens, standing for these entities, such as $C - S - T$. The latter is iconic in that it represents the spatial structure "to the left of" in an analog manner, but the tokens representing the entities are arbitrary.

The literature on mental models posits all possible representational flavors. Holland et. al.'s (1986) "induction" account considers mental models as propositional. On their view, mental models employ production-system type representations and are manipulated applying condition–action rules to propositional representations of a specific situation, such as making inferences about a feminist bank-teller using a model constructed of knowledge of feminists and banktellers. In the qualitative reasoning literature, the ontology of a mental model is represented propositionally and explicitly stated "qualitative equations" provide rules governing the possible state transitions of physical systems, such as "under condition X – move to next state" or "through behavior Y – move to next state" (see, e.g., Bobrow, 1985). The research by Johnson-Laird and colleagues (see, e.g., Johnson-Laird, 1983, 1989; Johnson-Laird and Byrne, 1993) on mental modeling in deductive and inductive reasoning tasks employs amodal iconic representations. These mental models are iconic in that they depict the salient structures among the entities in the problem, but the tokens representing entities are amodal, such as the $C - S - T$ in the example above. Making a logical inference such as *modus ponens* occurs by moving amodal tokens in a specific array that captures the salient structural dimensions of a problem, and then searching for models of counterexamples to the transformation. Modal iconic mental models—or 'perceptual models'—seem to be what Craik had in mind by an internal "'small-scale model' of external reality" (Craik, 1943). Simulation would involve mimicking physical transformations. "Depictive mental models" (Schwartz and Black, 1996b, 1996) provide a contemporary example of perceptual models. For example, in studies of gear rotation problems, Schwartz and Black argue that perceptual information is used to construct and manipulate a mental model of a set-up of machine gears. In a perceptual model, implicit knowledge embedded in physical constraints would be used to simulate possible behaviors in accord with real-world behaviors.

To aid in thinking about reasoning through simulation with perceptual models, there is an extensive literature that provides evidence that humans can perform simulative transformations in imagination which mimic physical transformations that can be recruited. The combinations and transformations using mental imagery are hypothesized to take place according to internalized constraints assimilated during perception. The literature on mental imagery indicates, for example, that people can mentally simulate combinations, such as in the classic example where subjects are asked to imagine a letter B rotated 90 degrees to the left, place an upside triangle below it and remove the connecting line. The simulation processes produce an image of a heart. Further, many experiments establish that in performing a mental simulation, such as rotating a figure, subjects exhibit latency times consistent with actually turning a mental figure around (see, e.g., Finke and Shepard, 1986; Kosslyn, 1980, 1994; Shepard and Cooper, 1932). There is also an extensive literature on spatial representation in mental models that indicates representation with respect to bodily orientation rather than a symmetrical Euclidian space (see, e.g., Franklin and Tversky, 1990; Glenberg, 1997). Additionally, research on mental modeling in discourse reasoning and comprehension tasks indicates that people can simulate various kinds of knowledge of physical situations in imaginary transformations. In these cases, too, such as when the imagined objects are separated by a wall, the spatial transformations exhibit latency times consistent with the reasoner having simulated moving an object around a wall rather than through it. Another significant line of research examines the role of causal knowledge in mental simulation. As mentioned earlier, Schwartz & Black have conducted studies focusing on gear rotations that provide evidence that people are able to perform simulative causal transformations on sets of gears, as does Mary Hegarty's research on problems with pulley systems (Hegarty, 1992; Hegarty and Just, 1989, 1994). In the gear problems, simulation ability was enhanced after the subject was told explicitly to imagine rotating the gears or given an visual display indicating simulation.

These interpretations are not without their critics from the camp which maintains that all mental representation is propositional. Zenon Pylyshyn (1981, 2001), for one, continues to argue that the data on visual mental imagery and transformation can be explained without having to invoke either the existence of imagery (as anything more than epiphenomena) or simulation. To explain the latency data, for example, he argues that the demand characteristics of the task could be such that they induce the subjects to perform calculations on how much time is required to traverse the distance and figure that into their responses. The arguments and counter-arguments on both sides of the "imagery debate" are too numerous to recount here. Again, I think the issue will continue to be unresolved for the foreseeable future, until the requisite neuroscience develops. In the meantime, there are good arguments and extensive research on mental imagery that can be recruited to develop theories of the nature of mental modeling, such as Stephen Kosslyn's (1980, 1994) theory of how transformation might take place in mental imagery.

Clearly much work remains to be done in developing a satisfactory understanding of mental modeling. What I am proposing here is that utilizing a minimalist notion provides a cognitive basis for interpreting the modeling practices exhibited in the historical records of conceptual change as indicative of mental modeling having played a central role in the past episodes. The minimalist notion is: in certain problem solving tasks humans reason by constructing an internal iconic model of the situations, events and processes that in dynamic cases can be manipulated through simulation. This will be considered more fully after we have discussed model-based reasoning. In the more mundane cases the reasoning performed in mental modeling is usually, though not necessarily, successful. For example, one usually is able to simulate successfully how to get the piece of furniture through the door, because the models and manipulative processes embody largely accurate assumptions about every-day real-world events. Admittedly it is some distance from the awkward furniture scenario and simulating causal transformations of rotating gears to employing the kinds of transformations requiring causal and other knowledge contained in a scientific theory. Further, it is likely the case that model-based reasoning does not take place all "in the head", as the furniture problem might. However, as with other kinds of representing and reasoning, it is consistent with the cognitive-historical method to consider the scientific practices as outgrowths of the mundane practice of mental modeling. In the case of science where the situations are more removed from human sensory experience and the assumptions more imbued with theory, there is less assurance that a reasoning process, even if carried out correctly, will yield "success". In the evaluation process, a major criterion for success remains the goodness of fit to the phenomena, but success in science can also include such factors as enabling the generation of a viable mathematical representation that allows for progress in spite of the lack of explicitly confirming data, such as that of Newton for gravitation and James Clerk Maxwell for the time delay in propagation for the electromagnetic field.

4.2 Model-Based Reasoning

The central problem of creativity in representational change is that of how is it possible to create something new given that the process must start with existing representations. The traditional account of reasoning as carrying out logical operations on propositional representations has been a major obstacle to understanding conceptual innovation as the outcome of reasoning processes. Because the kinds of modeling employed by scientists in discovery processes cannot be reduced to logic, they are discounted as generative reasoning. Conceptual innovation is viewed as occurring in sudden flashes of insight, with new concepts springing forth from the head of the scientist—like Athena—fully grown. This does accord with some scientists' retrospective accounts, but if one examines their deeds—their papers, diaries, letters, notebooks—these records support a quite different interpretation in most cases. As I have been arguing for some years, conceptual change results from extended problem-solving processes. The records of these processes display extensive use of practices that constitute forms of model-based reasoning: analogical, visual, and simulative modeling. Modeling practices are employed both in experimental and in theoretical settings. Embracing these modeling practices as "methods" of conceptual change in science requires expanding philosophical notions of scientific reasoning to encompass forms of creative reasoning, most of which cannot be reduced to an algorithm in application, are not always productive of solutions, and can lead to incorrect solutions.

Analyzing the conceptual innovation practices as various forms of model-based reasoning requires constructing a unified account of forms of modeling that are mostly treated separately in the literature in cognitive science, such as analogy and imagery. Although the practices of analogical and visual modeling and thought experimenting can occur separately, they most often are employed together in a problem-solving process. Examining Figure 1 exemplifies why

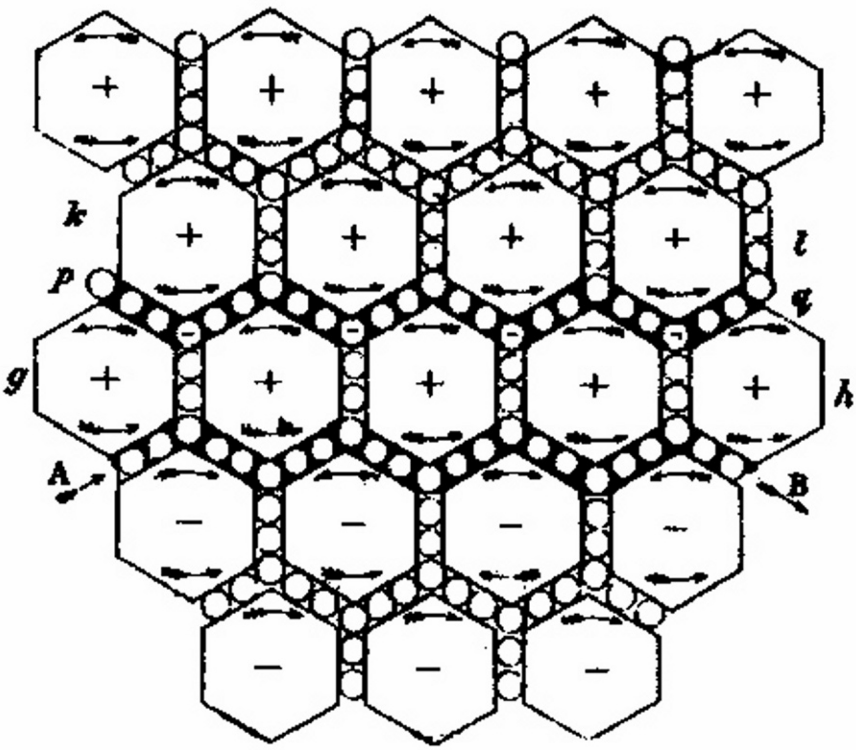

Figure 1. Maxwell's drawing of the vortex-idle wheel medium (Maxwell 1890, Vol. I, Plate VII)

a unified account is needed. The drawing is taken from a communication by Maxwell to his colleagues of his construction of a new, unified electromagnetic field concept, i.e., a paper published in *Philosophical Magazine*. The drawing is accompanied with instructions:

> Let the current from left to right commence in AB. The row of vortices *gh* above AB will be set in motion in the opposite direction to a watch... We shall suppose the row of vortices *kl* still at rest, then the layer of particles between these rows will be acted on by the row *gh* on their lower sides and will be at rest above. If they are free to move, they will rotate in the negative direction, and will at the same time move from right to left, or in the opposite direction from the current, and so form an *induced* electric current. (Maxwell, 1890, v. 1, p. 477, italics in original)

The figure is a visual representation of the analogical model Maxwell employed in constructing the electromagnetic field concept. The accompanying instructions assist the reader in animating it in thought. To understand and reason with the figure requires analogical, visual, and simulative modeling.

To explain why modeling practices figure centrally in conceptual innovation in science requires a fundamental revision of the understandings of concepts, conceptual structures, conceptual change, and reasoning customarily employed explicitly in philosophy and at least tacitly in the science studies fields more generally. The basic ingredients of that revision are to view the representation of a concept (whatever its format) as providing a set of constraints for generating members of a class of models, and a conceptual structure, as an agglomeration of these constraints. Concept formation and change is, then, a process of generating new and changing existing constraints. Model-based reasoning promotes conceptual change because these forms of reasoning are effective for abstracting, generating, integrating, and changing constraints. Genuine novelty is produced through combinations made possible through the generic abstraction process discussed below.

To engage in analogical modeling in science one calls on knowledge of the generative principles and constraints for a known source domain. These constraints and principles may be represented mentally in different informational formats and long-term knowledge structures that act as tacit assumptions employed in constructing and transforming models during problem solving. Inter- or intra-domain analogical models can be retrieved and applied with suitable adaptation, but often, and especially in cases of conceptual innovation, no direct analogy exists and construction of an initial source model is required. In these cases the analogical domain serves as the source for constraints to be used in interaction with those provided by the target problem to create an initial novel analog model, as well as subsequent models. Evaluation of the analogical modeling process is largely in terms of how well the salient constraints of a model fit the salient constraints of a target problem, with key differences playing a significant role in further model generation (Griffith et al., 1996).

As with other instances of analogical modeling, when employed in conceptual innovation the process often requires recognition of potential similarities across disparate domains, and a means of integrating information from them. "Generic abstraction" is a key reasoning process that enables recognition, adaptation, and integration. Constraints in both the target and the source domains are domain-specific. For retrieval, transfer and integration to occur in the reasoning process, they need to be understood at a sufficient level of abstraction. The various representations employed have to function with some of their features considered as unspecified, that is, as generic. In model-based reasoning processes, a central objective is to create a model that is of the *same kind* with respect to salient dimensions of the target phenomena one is trying to represent. Thus, although an instance of a model is specific, inferences made with it in a reasoning process are generic. In viewing a model generically, one takes it as representing features common to members of a class of phenomena. The relation between the generic model and a specific instantiation is similar to the type–token distinction in logic. Generality in representation is achieved by interpreting the components of the representation as referring to object, property, relation, or behavior types rather than tokens of these. In reasoning about a triangle, for instance, one cannot draw or imagine a generic triangle, but only some specific instance of a triangle. However, in considering what it has in common with all triangles, humans have the ability to imagine it as lacking specificity in the angles and the sides. That is, the reasoning context demands that the interpretation of the concrete polygon be as generic. The same is the case in considering the behavior of a physical system. To consider what a specific representation of a spring has in common with all springs, one needs to reason as though it lacked specificity in length and width and number of coils; to consider what it has in common with all simple harmonic oscillators, one needs to reason as though it lacked specificity in structure and some aspects of behavior. The analogical model, understood generically, represents what is common among the members of specific classes of physical systems, viewed with respect to a problem context.

The kind of creative reasoning employed in conceptual innovation involves not only applying generic abstractions, but creating and transforming them during the reasoning process. Maxwell's vortex-idle wheel analogy represented visually in Figure 1 is one example. In constructing the model Maxwell was considering what certain continuum mechanical and electromagnetic systems have in common. The construction and subsequent reasoning required that the dynamical relations among the idle wheels and vortices be treated as generic. That is, they must be understood to represent the class of such dynamical systems, and that class includes electric and magnetic interactions on the assumptions of Maxwell's treatment (Nersessian, 1992a, 2002). There are many significant examples of generic abstraction in conceptual innovation. In the domain of classical mechanics, for example, Newton can be interpreted as employing generic abstraction in reasoning about the commonalities among the motions of planets and projectiles in formulating a unified mathematical representation of their motions. Newton's inverse-square law of gravitation abstracts what a projectile and a planet have in common in the context of determining motion, such as that both can be represented as point masses. After Newton, the inverse-square-law model itself served as a generic model of action-at-a-distance forces for those who tried to bring all forces into the scope of Newtonian mechanics.

A variety of perceptual resources are used by scientists in modeling. Visual modeling figures prominently in conceptual change across the sciences. This may be because employing the visual modality enables the reasoner to bypass constraints inherent in current linguistic or formulaic representations of conceptual structures. As discussed in the previous section, there is a vast literature in cognitive science on mental imagery that provides evidence that humans can perform simulative transformations in imagining that mimic physical spatial and causal transformations. External visual representations provide support for the processes of constructing and reasoning with a mental model. They aid significantly in organizing cognitive activity during reasoning, such as fixing attention on the salient aspects of a model during reasoning, enabling retrieval and storage of salient information, and exhibiting salient interconnections, such as structural and causal, in appropriate co-location. Further they facilitate the construction of shared mental models in a community and the transportation of scientific models out of the local milieu of their construction.

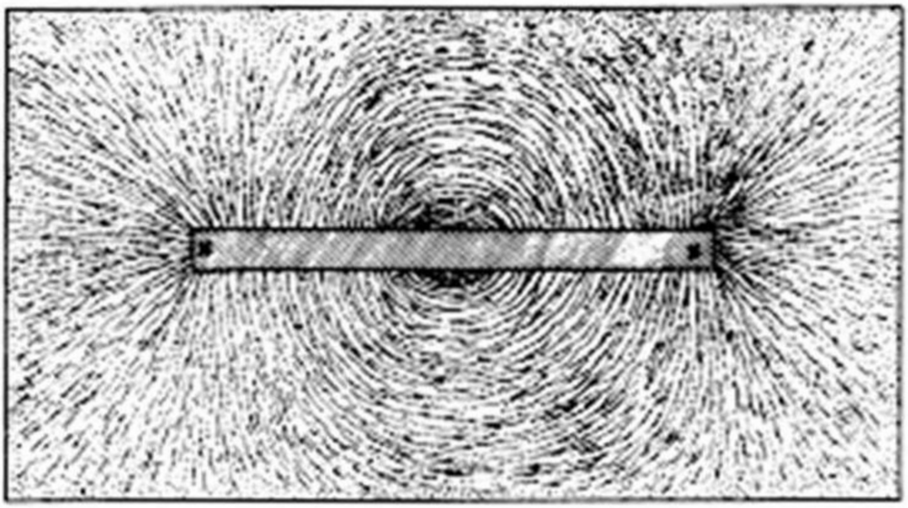

Figure 2. Lines of force

As used in model-based reasoning in physics, external visual representations tend to be schematic. These representations can aid modeling phenomena in several ways, including providing abstracted and idealized representation of aspects of phenomena and embodying aspects of theoretical models. For example, early in Michael Faraday's construction of a field concept the visual model represented in Figure 2 provided an idealized representation of the lines of force surrounding a magnet. Later in his development of the field concept, the visual model of lines of force functioned as the embodiment of a dynamical theoretical model of the transmission and interconversion of forces, generally, through stresses and strains in, and various motions of, the lines. The visual analogical model represented by Maxwell in Figure 1, however, was intended as an embodiment of an imaginary system, displaying a generic dynamical relational structure, and not as a representation of the theoretical model of electromagnetic field actions in the aether.

As a form of model-based reasoning, thought experimenting is a specific kind of simulative reasoning, which can occur in other forms of model-based reasoning. In the case of scientific thought experiments implicated in conceptual change, the main historical traces are in the form of narrative reports, constructed after the problem solving has taken place. These have often provided a significant means of effecting conceptual change within a scientific community. Accounting for the generative role of thought experimenting, thus begins with examining how these narratives support modeling processes and, by means of cognitive-historical analysis, infers that the original experiment involves a similar form of model-based reasoning (Nersessian, 1992b). What needs to be determined are: (1) how a narrative facilitates the construction of a model of an experimental situation in thought and (2) how one can reach conceptual and empirical conclusions by mentally simulating experimental processes.

From a mental modeling perspective, the function of the narrative form of presentation of a thought experiment would be to guide the reader in constructing a mental model of the situation described by it and to make inferences through simulating the events and processes depicted in it. A thought-experimental model can be construed as a form of "discourse" model (Perrig and Kintsch, 1985; Johnson-Laird, 1982), with the operations and inferences performed not on propositions but on the constructed model. Unlike a fictional narrative, however, the context of the scientific thought experiment makes the intention clear to the reader that the inferences made pertain to potential real-world situations. The narrative has already made significant abstractions, which aid in focusing attention on the salient dimensions of the model and in recognizing the situation as prototypical (generic). Thus, the experimental consequences are seen to go beyond the specific situation of the thought experiment. The thought-experimental narrative is presented in a polished form that "works", which should make it an effective means of getting comparable mental models among the members of a community of scientists. Undoubtedly experimental revision and tweaking goes on in the original reasoning and in the narrative construction, although accounts of this process are rarely presented.

Although some kinds of mental modeling may employ static representations, those derived from thought-experimental narratives are usually dynamic. The narrative delimits the specific transitions that govern what takes place. In constructing and conducting the experiment a scientist makes use of inferencing mechanisms, existing representations, and scientific and general world knowledge to make constrained transformations from one possible physical state to the next. Much of the information employed in these transformations is tacit. Thus, expertise and learning play a crucial role in the practice. So does the know-how derived from perceptual experience which David Gooding (1992) calls "embodiment". The thought-experimental process links the conceptual and the experiential dimensions of human cognitive processing. Thus, the constructed situation inherits empirical force by being abstracted both from our experiences and activities in the world and from our knowledge, conceptualizations, and assumptions of it. In this way, the data that derive from thought experimenting have empirical consequences and at the same time pinpoint the locus of the needed conceptual reform. The derived understanding forms the basis of further problem-solving efforts to construct an empirically adequate conceptualization.

All three forms of model-based reasoning are complex forms of reasoning that integrate information represented in multiple formats—propositions, models, and equations—into mental models. There are several key ingredients common to the various forms of model-based reasoning. They are semantic reasoning processes in that the models are intended as interpretations of target physical systems, processes, phenomena, or situations. The models are retrieved or constructed on the basis of potentially satisfying salient constraints of the target domain. In the modeling process, various forms of abstraction, such as limiting case, idealization, generalization, and generic modeling, are utilized. Evaluation and adaptation take place in light of structural, causal, and/or functional constraint satisfaction and enhanced understanding of the target problem obtained through the modeling process. Simulation can be used to produce new states and enable evaluation of behaviors, constraint satisfaction, and other factors.

From the perspective of conceptual innovation and change as involving processes of generating and transforming constraints, model-based reasoning is particularly effective. Analogy is a means through which constraints are abstracted from existing representations, including quite disparate domains, and integrated into models providing candidate constraints for new concepts. Thus, although analogical modeling enables arguments, the power of analogy lies in employing generic abstraction in the service of model construction, manipulation, and evaluation. Understood in this way, analogical modeling is a powerful form of reasoning, as opposed to the standard philosophical evaluation of "argument by analogy" as a weak form of reasoning. Visual modeling appears to be a highly developed and effective means of displaying constraints in a form in which humans can grasp them and follow through their consequences immediately and efficiently. As philosophers have worried for centuries, visual representations do indeed have the potential to lead a reasoner astray, but visual modeling is an effective tool for science when sufficient constraints are guiding the reasoning process. Finally, although many thought experiments can be reconstructed as arguments (Norton, 1991), their modeling function cannot be supplanted by an argument. The argument is not evident until after the thought experiment has been constructed and executed. Thought experimenting plays a crucial role in conceptual change by showing that existing systems of constraints cannot be integrated into consistent models of the physical world. Thought experimenting may facilitate recognizing the undesirable consequences of a conceptualization in much the way that experimenting by computer simulation exposes undesirable consequences of the constraints of a scientific representation. By creating a simulative model that attempts to integrate specific systems of constraints, thought experimenting enables a scientist to grasp essential points of conflict and infer their consequences more readily than they would by reasoning through the logical consequences of a representation. Once the initial experimenter understands the implications of a thought experiment, she can guide others in the community to see them as well by crafting a description of the experiment into a narrative.

This account of the generative nature of model-based reasoning in conceptual innovation lends support to the position of other contemporary philosophers (see, e.g., Cartwright, 1989; Giere, 1988) that in reasoning with or about a theory, the basic units scientists employ are not axiom systems, not propositional networks, but models. The term 'model' is used in these accounts not in the logical sense of an abstract mapping of things to terms, but in the analog sense of a structure intended as isomorphic to some aspect of a physical system. Together these hypotheses about model-based reasoning in creating and using theories make a claim that no matter how theories and concepts may *in principle* be represented, models are the mental representations with which a scientist carries out much reasoning and by means of which she thinks and understands through the lens of a conceptual structure.

5. Model-based Reasoning as Situated and Distributed Reasoning

In an obvious but non-trivial sense, model-based reasoning is situated: the scientist constructs a model and reasons in the situation it represents. Of course, here the reasoning needs to apply to the type of phenomena and not just the specific instance. On my analysis of model-based reasoning, the generic abstraction process enables reasoning in the situation to lead to inferences applying to the appropriate class of phenomena represented in the situation. The process is similar to reasoning about "triangularity" from a representation of a specific triangle. Taking the Maxwell case of conceptual innovation, in constructing the mathematical representation of the electromagnetic field concept field, Maxwell created several models of an imaginary fluid medium drawing from the source domains of continuum mechanics and of machine mechanics. On my analysis, these analogical domains served as sources for constraints used together with those provided by the target problem to create the imaginary analog models that served as the basis of his reasoning. Maxwell also employed several imagistic representations, such as that in Figure 1. In constructing the various continuum mechanical models Maxwell was explicit about creating physical situations in which to carry out the abstract reasoning involved in determining the structural relations governing electromagnetic interactions and how to represent these mathematically.

Although ignored by many philosophers and historians, Maxwell's own comments on his method of analysis are most insightful. In investigating a new area in science, Maxwell asserted that one begins with a process of "simplification and reduction of the results of previous investigation to a form in which the mind can grasp them" (Maxwell, 1855, p. 155). That process requires a "method of investigation, which allows the mind at every step to lay hold of a clear physical conception, without being committed to any theory founded on the physical science from which that conception is borrowed so that it is neither drawn aside from the subject in pursuit of analytical subtleties, nor carries beyond truth by favorite hypotheses" (ibid., p. 156) A "physical analogy" is "that partial similarity between the laws of one science and those of another which makes each of them illustrate the other" (ibid.). In Craik's terminology, Maxwell's physical analogies are "relation structures".

It does not matter whether the mechanical systems employed in the models do or do not exist in nature; all that matters is that they are "mechanically conceivable". That is, that they supply mechanisms belonging to the classes of phenomena with dynamical relational structure common to mechanics and electromagnetism. The models provide an environment in which to carry on reasoning. Throughout his reasoning processes Maxwell abstracted from the specific mechanism to find the mathematical form of that class of mechanism, i.e., of the generic dynamical structure. In this manner, Maxwell was able to formulate the laws of the electromagnetic field by abstracting from specific mechanical models the dynamical properties and relations continuum-mechanical systems, certain machine mechanisms, and electromagnetic systems have in common. In their mathematical treatment these common dynamical properties and relations were separated from the specific instantiations provided in the models through which they had been rendered concrete. The generic mechanical relationships represented by the imaginary systems of the models served as the basis from which he abstracted a mathematical structure of sufficient generality that it represented *causal processes* in the electromagnetic medium without requiring knowledge of specific *causal mechanisms*—similar to the achievement of Newton and the universal law of gravitation.

Model-based reasoning is often a distributed process, too, where the reasoning employs not only representations and processes in the head but also in the environment. Putting it paradoxically, the mental modeling process can be construed as not only taking place in the mind. When considering external representations part of the cognitive system, it is possible the process can make direct use of information in the environment. Many instances of model-based reasoning in science and engineering employ 'external' representations that are constructed for and during the reasoning process, such as diagrams, sketches, and physical models, and these can be viewed as providing constraints and affordances essential to problem solving that augment whatever the mental representations used during the process provide. One finds evidence of their use in the historical records and in current-day scientific practices. While it might be difficult to say with surety that Maxwell sketched or had a visual representation in front of him as he reasoned, there is sufficient evidence of contemporary use both in practice and in problem-solving protocols with scientists (including gestural representations). Within cognitive systems, external representations can instantiate part of the current model of the phenomena, allow manipulation, and facilitate simulative processes. The external representation can change the nature of the processing task, such as when the TIC-TAC-TOE grid is placed over the mathematical problem "15" (Zhang and Norman, 1995). Even in the simple case of simulating how to get a piece of furniture through a doorway, it is much easier to do so when the furniture and doorway are in front of you. One line of criticism against mental modeling simulation and in favor of logical reasoning over propositions has been that it is just too complex to be psychologically realistic (Rips, 1986). There are two lines along which to answer this criticism. First, not everything needs to be represented in the head to carry out a simulation. From a distributed cognition perspective, one can expand the notion of mental modeling to comprise both what are customarily held to be the internal thought of the human agent and the external representations. Simulative model-based reasoning would, under this construal, involve a process of co-constructing 'internal' models of the phenomena and 'external' models, each of which are incomplete. Understood in this way, simulating the mental model would consist of processing information both in memory and in the environment—see Greeno, 1989b for a similar view.

Second, much of the speculation about the nature of simulation comes from considering constraints of computational modeling. Psychological theories that claim simulation utilizes perceptual and motor mechanisms have the potential to provide better solutions. As discussed earlier, my analysis of model-based reasoning in conceptual change requires adopting only a *minimalist* hypothesis: that in certain problem solving tasks humans reason by constructing an internal iconic model of the situations, events and processes that in dynamic cases can be manipulated through simulation. In constructing such a model, it does however, need to be possible to utilize information in various formats, including linguistic, formulaic, and deriving from various perceptual modalities. However, the issue of whether the content of the representation is modal or amodal and what the generative processes are for creating and operating on mental models do not have to be resolved before we can make progress on an account of model-based reasoning in science. The minimalist hypothesis locates the cognitive basis for the hypothesis that the modeling practices of scientists constitute a form of reasoning through which new conceptual structures are constructed within a major thrust in the mental modeling framework. Still, I think there are some considerations weighing in favor of perceptual mental models. Developing these fully is beyond the scope of this paper, so I make only brief allusion to them in concluding this section.

As we discussed in Section 4.1 it is on the basis of extensive cognitive science research in numerous domains that mental modeling has been proposed as a fundamental form of human reasoning. It is hypothesized to have evolved as an efficient means of navigating the environment and solving problems in matters of significance to existence in the world. Following on the evolutionary hypothesis, the perceptual mental model position appears more in accord with the ability to simulate the environment. The ability should not be unique to humans, since, for example, other animals need to anticipate and predict their environment for survival purposes. In these non-human cases, perceptual and motor mechanisms would need to be employed. What makes humans unique is that we can construct models from both perceptual and descriptive information. Possibly for the human ability of logical inferencing that Johnson-Laird investigates, amodal tokens could suffice. But, what I am calling simulative model-based reasoning is closer to imaginative thinking than logical reasoning, and there is mounting evidence from neuropsychology that the perceptual system plays a significant role in imaginative reasoning (see, e.g., Kosslyn, 1994). Again, this makes sense from an evolutionary perspective. The visual cortex is one of the oldest and most highly developed regions of the brain. As Roger Shepard has put it, perceptual mechanisms "have, through evolutionary eons, deeply internalized an intuitive wisdom about the way things transform in the world. Because this wisdom is embodied in a perceptual system that antedates, by far, the emergence of language and mathematics, imagination is more akin to visualizing than to talking or to calculating to oneself" (Shepard, 1988, p. 180). He argues that although the original ability to envision by mental modeling would have developed as a way of simulating possible courses of action in the world, it is highly plausible that, as human brains have developed, this ability has been "bent to the service of creative thought" (ibid.). Once extended to scientific reasoning, for instance, the nature and richness of models one can construct and one's ability to reason would develop as one learns domain-specific content and techniques. Comparative studies of expert and novice reasoning do indicate that skill in mental modeling develops in the course of learning (Chi et al., 1981). Thus, facility with mental modeling is a combination of an individual's biology and learning. The ability of the scientist or engineer to reason about technical material through mental modeling should differ significantly from that of ordinary folk. The salient point is that it originates in the cognitive endowment of ordinary folk.

For performing simulation tasks, a mental model would need to capture the causal coherence of a system and other relevant behavioral constraints of the kinds of physical entities represented in the model and possible relations among them. These can in principle be represented in either propositional or iconic structures. However, being able to make direct use of perceptual affordances and perceptual and motor processing would increase the ease of reasoning with a mental model, and is consistent with other creatures having the ability. Unlike an amodal representation, in a modal representation perceptually-relevant information about objects, processes, situations is available directly for use. Running through a series of logical inferences such as "if I move the right corner up and to the right then the other corner will swivel to the left" or performing a set of trigonometric calculations on a model with amodal tokens for the door and the chair in proper spatial configurations, seems a more cumbersome process than simulating possible movements in a spatial configuration using a token with perceptual features approximating the chair and the door. Perceptual mental models need only be schematic in that they contain selective representations of aspects of objects, situations, and processes, making for flexibility in reasoning and comprehension tasks. Performing a simulation with a perceptual mental model is mentally re-enacting perceived information and, thus, would facilitate inferences about the real-world phenomena. The simulation should comply with the same constraints as the system it represents, such as a catcher simulating the path of a baseball and anticipating where it will land. Modal representations would have an advantage in simulation, if, as Barsalou (1999) argues, constructing a modal representation is likely to involve reactivation of patterns of neural activity in the perceptual and motor regions of the brain that were activated in the initial experience of something. This simulation ability would require that there is at least a component of long-term memory representations that is perceptually-based. On Barsalou's theory, the long-term representation of a perceptual experience—the "perceptual symbols"—can be stored separately and reactivated in thinking to create novel combinations. The major advantage with a perceptual model is that simulation would employ perceptual (and possibly motor) information and processing directly in the inferencing process. Cast in situated and distributed terms, perceptual mental models would enable a reasoner to take direct advantage of affordances and constraints inherent in the situation being modeled. The interaction with the environment could be enacted in imagination, or in a combination of imagination and external representation. The internal processing would be making direct use of the situational information in the format in which it was encoded (Yeh and Barsalou, 1996). How simulation would take place in the brain is an open question, though Kosslyn's theory of visual mental imagery, which postulates perceptual and motor processes in image transformation, might be extended to mental modeling. Shepard (1984) and others have attempted to develop a mathematical representation of psychokinetic processing in the nervous system.

6. Culture and Cognition: Implications for Creativity

We are now in a position to return briefly to the role of model-based reasoning in conceptual innovation with an eye to indicating how the analysis in the previous sections provides potential resources for explaining how cognitive, social, and material elements are fused in the representations of science. Thus far I have argued that the account needs to be rooted in the interplay between individual mental activity and the environmental context in which reasoning takes place. First, mental modeling is not just something scientists do, but is fundamental to various aspects of human existence. The claim is that in the process of developing scientific approaches to understanding nature, the cognitive tool was extended. Skill in employing it in scientific reasoning now develops through acquiring expertise. Second, scientific modeling always takes place in a material environment that includes the natural world and socio-cultural artifacts (stemming from both within science and outside of it), including instruments devised by scientific communities to probe and represent that world. Modeling employs a range of representational resources. Third, scientists employ modeling practices not only in creating representations, but in communicative attempts at creating shared understanding. That is, modeling plays a central role in creating, comprehending, transmitting, and adopting scientific representations. In short, the modeling practices and models of scientists are cognitive and socio-cultural achievements and artifacts.

In dispelling the genius mythology discussed at the outset, two points made throughout the paper need to be re-emphasized here. First of all, social and cultural context is crucial to understanding *any* creative process in science, and conceptual innovation is no exception. As discussed in the previous sections, model-based reasoning has scientists reasoning in situations. The representations and processes employed in constructing and manipulating these situations are culturally laden. Returning to the figure Maxwell drew in constructing the mathematical representation of the electromagnetic field concept (Figure 1), what it represents is the situation he created in which to carry out abstract reasoning about certain relations between electric and magnetic phenomena. It also represents his attempt to communicate his reasoning and is a representational device with the potential to create mental models similar to his own within his community. But from where did the representation in the figure derive and why did he use the "method of physical analogy" approach rather than, say, pure mathematical analysis? As was noted previously, Maxwell's educational in Scotland and Cambridge led to his training as mathematical physicist of a certain kind. This training was significantly determinative of the nature of the theoretical, experimental, and mathematical knowledge and the methodological practices with which he formulated the problem and approached its solution. The mathematical and physical representations and methods of continuum mechanics were in his tool kit, more than action-at-a-distance representations, which of course he was aware of and could use. Continental physicists working on electromagnetism at the same time employed quite different methodological practices and drew from the fundamentally different action-at-a-distance mathematical and physical representational structures. Further, the theoretical speculations of Faraday as to the active nature of space surrounding bodies and charges made continuum mechanics more salient to Maxwell in approaching the problem. Finally, William Thomson's (later, Lord Kelvin) practice of constructing mathematical representations on the basis of analogies, though different from *how* Maxwell used analogy were especially important to *that* he started from analogical models. In sum, the culture of the specific scientific environment is evident both in the representational content of the models Maxwell's constructed and in his using analogical, visual, and simulative modeling as reasoning processes at all.

Secondly, in the process of creating new concepts, concepts from all aspects of a scientist's experience are candidates for redeployment as analogical sources which, with suitable abstraction, can be applied in specific problem-solving processes. This fact helps with the problem of how wider socio-cultural context could be implicated in the context of scientific practices. Here generic abstraction, as discussed in Section 4.2 can provide a mechanism for importing representations drawn from wider culture into the representational content of science. One can, for example, interpret the historical claim that Faraday's religious views about the "unity of nature" had a significant impact on the specific form of field concept he developed (Cantor, 1985; Nersessian, 1984, 1985; Williams, 1964) in the following way. A generic concept of the unity of nature can be abstracted from the specific religious context. Redeployed with respect to the problem of the nature of physical forces, it could provide a constraint of "the unity of all forces in nature", that then facilitated Faraday's developing a dynamical model of the interaction and interconversion of all the forces of nature—chemical, electric, magnetic, gravitational—which he did by using the forms of model-based reasoning we have discussed.

To take a more challenging example, part of Maxwell's modeling process can similarly be interpreted. Maxwell's modeling processes involved adjusting multiple constraints drawn from electromagnetism, continuum mechanics, and machine mechanics. Consider Maxwell's introduction of "idle wheel particles" into his model of the electromagnetic medium in developing the field concept (Figure 1). Maxwell's first model had vortices packed in the aether without separation (illustrated by me in Figure 3). Simulating a preliminary version

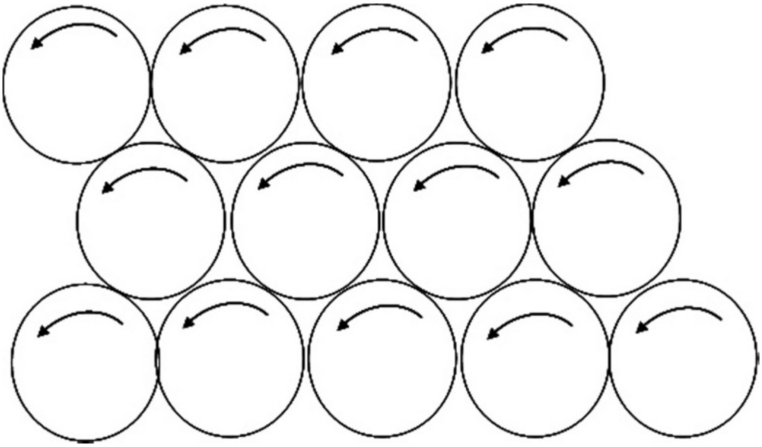

Figure 3. Cross section of model of vortex fluid medium

of the model described by Maxwell, provides a constraint that friction would bring the spinning vortices to a halt. Maxwell next utilized an intuitive model of machine gears and fly wheels. This resource is not readily connected with the hybrid of the continuum mechanical and electromagnetic domains from which he was reasoning, but it could have quickly come to mind because of a widely accessible cultural resource: the Victorian fascination with machines—especially the steam engine. Through a generic abstraction process, such as that illustrated in Figure 4, the cultural model could provide constraints to be redeployed in the vortex-idle wheel model drawn by Maxwell in Figure 1. Further model-based reasoning led to the construction of a unified mathematical representation of the electromagnetic field concept, which was the object of the problem-solving process (for a detailed technical discussion, see Nersessian, 2002).

Although these examples provide only sketches, they are based on the more detailed research cited earlier. To fully interpret any instance of conceptual innovation requires that level of analysis. A deep understanding of scientific cognition, from the mundane to the creative, and how it leads to knowledge requires no less than ascertaining what it means to be a human thinker acting in specific complex physical and socio-cultural worlds. This is a complex and multi-faceted problem. The analysis here has aimed to build a framework that will enable progress; specifically by providing a cognitive basis in support of the claim that the practices exhibited in the historical records of major conceptual innovations constitute reasoning generative of concept formation and change. Clearly more work remains to be done in filling out this account of conceptual innovation. However, the present analysis demonstrates that the perceived division between the individual and the socio-cultural, between cognition and

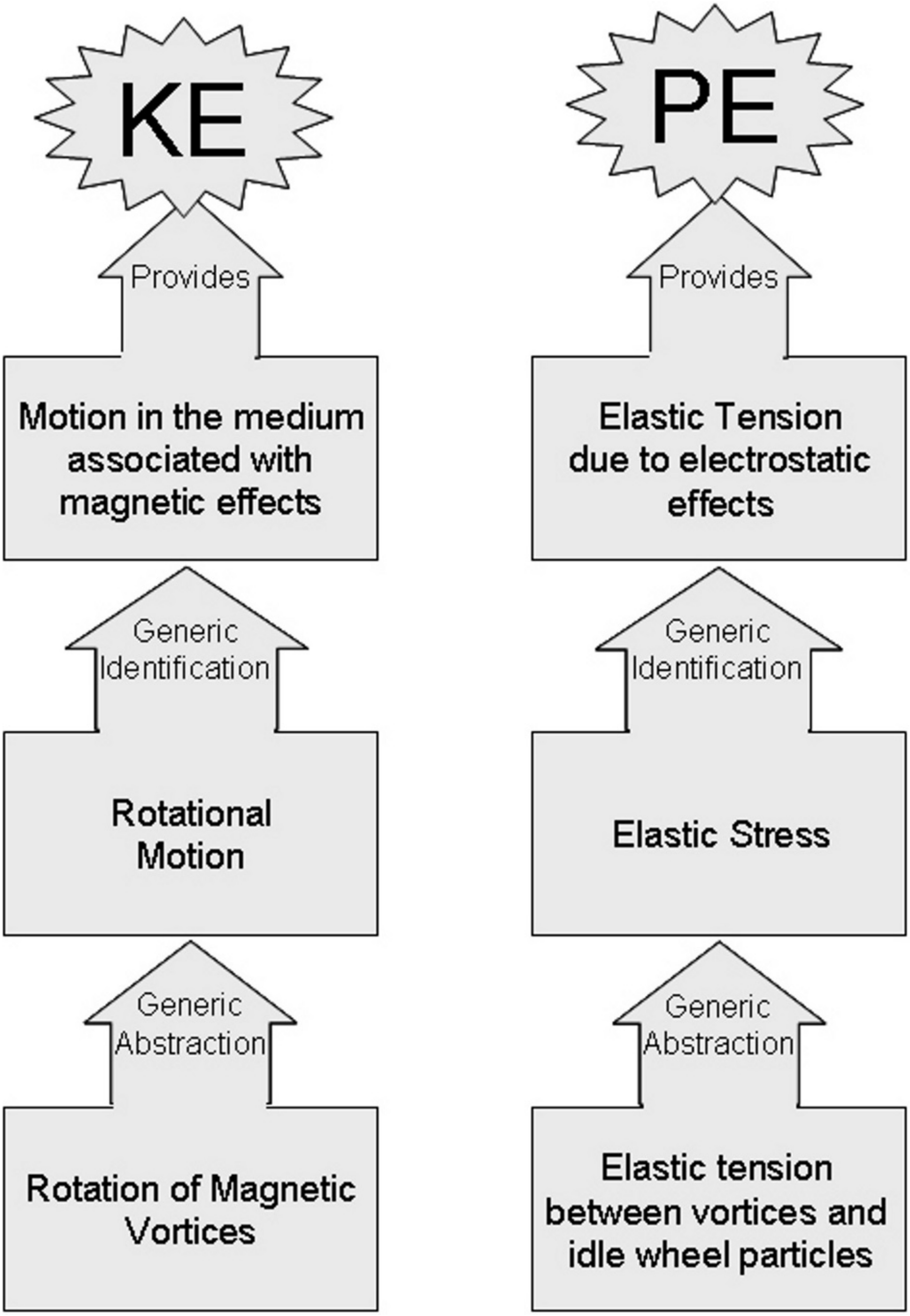

Figure 4. Identifying emergency components via generic modeling

culture, in constructing scientific knowledge is artificial. The scientific "genius" who creates in isolation from social and cultural contexts is, indeed, a myth. The cognitive-historical method of analysis provides the resources for studying the social–cognitive–cultural nexus from a unified perspective.

Acknowledgments

I gratefully acknowledge the support of grants from the National Science Foundation in carrying out this research: STS Scholar's Award SBE 9810913 and ROLE research grants REC106773 and REC0411825.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22:577–609.

Bobrow, D. G. (1985). *Qualitative Reasoning About Physical Systems*. MIT Press, Cambridge, MA.

Boden, M. A. (1990). *The Creative Mind: Myths and Mechanisms*. Basic Books, New York.

Brown, J. S., Collins, A., and Duguid, S. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18:32–42.

Cantor, G. N. (1985). Reading the book of nature: The relation between Faraday's religion and his science. In Gooding and James, 1985, pages 69–82.

Cartwright, N. (1989). *Nature's Capacities and Their Measurement*. Clarendon, Oxford.

Chi, M. T. H., Feltovich, P. J., and Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5:121–152.

Clement, J. (1989). Learning via model construction and criticism. In Glover, G., Ronning, R., and Reynolds, C., editors, *Handbook of Creativity: Assessment, Theory, and Research*, pages 341–381. Plenum, New York.

Craik, K. (1943). *The Nature of Explanation*. Cambridge University Press, Cambridge.

Darden, L. (1991). *Theory Change in Science: Strategies from Mendelian Genetics*. Oxford University Press, New York.

de Kleer, J. and Brown, J. S. (1983). Assumptions and ambiguities in mechanistic mental models. In Gentner and Stevens, 1983, pages 155–190.

Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Harvard University Press, Cambridge, MA.

Elman, J. L., Bates, E. A., Johnson, M., Karmiloff-Smith, A., Parisi, D., and Plunkett, K. (1998). *Rethinking Innateness: A Connectionist Perspective on Development*. MIT Press, Cambridge, MA.

Finke, R. A. and Shepard, R. N. (1986). Visual functions of mental imagery. In Boff, K. R., Kaufman, L., and Thomas, J. P., editors, *Handbook of Perception and Human Performance*, pages 37.1–37.55. John Wiley & Sons, New York.

Fodor, J. A. (1975). *The Language of Thought*. Thomas Y. Crowell Company, New York.

Franklin, N. and Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology*, 119:63–76.

Geertz, C. (1973). *The Interpretation of Cultures*. Basic Books, New York.

Gentner, D., Brem, S., Ferguson, R., and Wolff, P. (1997). Analogical reasoning and conceptual change: A case study of Johannes Kepler. *The Journal of the Learning Sciences*, 6:3–40.

Gentner, D. and Gentner, D. R. (1983). Flowing waters and teeming crowds: Mental models of electricity. In Gentner and Stevens, 1983, pages 99–130.

Gentner, D. and Stevens, A. L., editors (1983). *Mental Models*. Lawrence Erlbaum, Hillsdale, NJ.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston.

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

Giere, R. N., editor (1992). *Cognitive Models of Science*. University of Minnesota Press, Minneapolis.

Gilhooly, K. J. (1986). Mental modeling: A framework for the study of thinking. In Bishop, J., Lochhead, J., and Perkins, D., editors, *Thinking: Progress in Research and Teaching*, pages 19–32. Lawrence Erlbaum, Hillsdale, NJ.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, 20:1–19.

Gooding, D. (1990). *Experiment and the Making of Meaning: Human Agency in Scientific Observation and Experiment*. Kluwer, Dordrecht.

Gooding, D. (1992). The procedural turn: Or why did Faraday's thought experiments work? In Giere, 1992, pages 45–76.

Gooding, D. and James, F. A. J. L., editors (1985). *Faraday Rediscovered*. Stockton Press, New York.

Greeno, J. G. (1989a). A perspective on thinking. *American Psychologist*, 44:134–141.

Greeno, J. G. (1989b). Situations, mental models, and generative knowledge. In Klahr, D. and Kotovsky, K., editors, *Complex information processing*, pages 285–318. Lawrence Erlbaum, Hillsdale, NJ.

Griesemer, J. R. (1991). Must scientific diagrams be eliminable? The case of path analysis. *Biology and Philosophy*, 6:177–202.

Griesemer, J. R. and Wimsatt, W. (1989). Picturing Weismannism: A case study of conceptual evolution. In Ruse, M., editor, *What the Philosophy of Biology is: Essays for David Hull*, pages 75–137. Klwuer, Dordrecht.

Griffith, T. W., Nersessian, N., and Goel, A. (1996). The role of generic models in conceptual change. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ. Lawrence Erlbaum Associates.

Hegarty, M. (1992). Mental animation: Inferring motion from static diagrams of mechanical systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18:1084–1102.

Hegarty, M. and Just, M. A. (1989). Understanding machines from text and diagrams. In Mandl, H. and Levin, J., editors, *Knowledge Acquisition from Text and Picture*. North Holland, Amsterdam.

Hegarty, M. and Just, M. A. (1994). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32:717–742.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., and Thagard, P. A. (1986). *Induction: Processes of Inference, Learning, and Discovery*. MIT Press, Cambridge, MA.

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

Johnson-Laird, P. N. (1982). The mental representation of the meaning of words. *Cognition*, 25:189–211.

Johnson-Laird, P. N. (1983). *Mental Models*. MIT, Cambridge, MA.

Johnson-Laird, P. N. (1989). Mental models. In Posner, M., editor, *Foundations of Cognitive Science*, pages 469–500. MIT Press, Cambridge, MA.

Johnson-Laird, P. N. and Byrne, R. (1993). Precis of the book, deduction with peer review commentaries and responses. *Brain and Behavioral Sciences*, 16:323–380.

Kosslyn, S. M. (1980). *Image and Mind*. Harvard University Press, Cambridge, MA.

Kosslyn, S. M. (1994). *Image and Brain*. MIT Press, Cambridge, MA.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press, Chicago.

Latour, B. (1987). *Science in Action*. Harvard University Press, Cambridge, MA.

Latour, B. (1999). *Pandora's Hope: Essays on the Reality of Science Studies*. Harvard University Press, Cambridge, MA.

Latour, B. and Woolgar, S. (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press, Princeton.

Lave, J. (1988). *Cognition in Practice: Mind, Mathematics, and Culture in Everyday Life*. Cambridge University Press, New York.

Lynch, M. (1985). *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. Routledge and Kegan Paul, London.

Lynch, M. and Woolgar, S., editors (1990). *Representation in Scientific Practice*. MIT Press, Cambridge, MA.

Mani, K. and Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory and Cognition*, 10:181–187.

Maxwell, J. C. (1855). On Faraday's lines of force. *Scientific Papers*, 1:155–229.

Maxwell, J. C. (1890). *The Scientific Papers of James Clerk Maxwell*. Cambridge University, Cambridge, MA.

McNamara, T. P. and Sternberg, R. J. (1983). Mental models of word meaning. *Journal of Verbal Learning and Verbal Behavior*, 22:449–474.

Nersessian, N. J. (1984). *Faraday to Einstein: Constructing Meaning in Scientific Theories*. Martinus Nijhoff/Kluwer Academic Publishers, Dordrecht.

Nersessian, N. J. (1985). Faraday's field concept. In Gooding and James, 1985, pages 377–406.

Nersessian, N. J. (1988). Reasoning from imagery and analogy in scientific concept formation. In Fine, A. and Leplin, J., editors, *PSA* 1988, pages 41–47. Philosophy of Science Association, East Lansing, MI.

Nersessian, N. J. (1992a). How do scientists think? Capturing the dynamics of conceptual change in science. In Giere, 1992, pages 3–45.

Nersessian, N. J. (1992b). In the theoretician's laboratory: Thought experimenting as mental modeling. In Hull, D., Forbes, M., and Okruhlik, K., editors, *PSA* 1992, pages 291–301. Philosophy of Science Association, East Lansing, MI.

Nersessian, N. J. (1995). Opening the black box: Cognitive science and the history of science. In Thackray, A., editor, *Constructing Knowledge in the History of Science*, volume 10 of *Osiris*, pages 194–211.

Nersessian, N. J. (1999). Model-based reasoning in conceptual change. In Magnani, L., Nersessian, N. J., and Thagard, P., editors, *Model-Based Reasoning in Scientific Discovery*, pages 5–22. Kluwer/Plenum, Dordrecht.

Nersessian, N. J. (2002). Maxwell and the "method of physical analogy": Model-based reasoning, generic abstraction, and conceptual change. In Malamet, D., editor, *Reading Natural Philosophy: Essays in the History and Philosophy of Science and Mathematics*, pages 129–165. Open Court, Lacily, IL.

Newell, A. and Simon, H. A. (1972). *Human Problem Solving*. Prentice Hall, Englewood Cliffs, NJ.

Nisbett, R., Peng, K., Choi, I., and Norenzayan, A. (2001). Culture and systems of thought: Holistic v. analytic cognition. *Psychological Review*, 108:291–310.

Norman, D. A. (1988). *The Psychology of Everyday Things*. Basic Books, New York.

Norton, J. (1991). Thought experiments in Einstein's work. In Horowitz, T. and Massey, G., editors, *Thought Experiments in Science and Philosophy*, pages 129–148. Rowman and Littlefield, Savage, MD.

Oakhill, J. and Garnham, A., editors (1996). *Mental Models in Cognitive Science: Essays in honor of Philip Johnson-Laird*. Psychology Press, Philadelphia, Brighton.

Perrig, W. and Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language*, 24:503–518.

Pylyshyn, Z. (1981). The imagery debate: Analog media *vs*. tacit knowledge. *Psychological Review*, 88:16–45.

Pylyshyn, Z. (2001). Visual indexes, preconceptual objects and situated vision. *Cognition*, 80:127–158.

Resnick, L. B., Levine, J. M., and Teasley, S., editors (1991). *Perspectives on Socially Shared Cognition*. APA Press, Washington DC.

Rips, L. (1986). Mental muddles. In Brand, H. and Hernish, R., editors, *The Representation of Knowledge and Belief*, pages 258–286. University of Arizona Press, Tuscon, AZ.

Rosch, E. and Lloyd, B. (1978). *Cognition and Categorization*. Lawrence Erlbaum, Hillsdale, NJ.

Schwartz, D. L. and Black, J. B. (1996a). Analog imagery in mental model reasoning: Depictive models. *Cognitive Psychology*, 30:154–219.

Schwartz, D. L. and Black, J. B. (1996b). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, 20:457–497.

Shelley, C. (1996). Visual abductive reasoning in archeology. *Philosophy of Science*, 63:278–301.

Shelley, C. (1999). Multiple analogies in evolutionary biology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 30:143–180.

Shepard, R. (1988). Imagination of the scientist. In Egan, K. and Nadaner, D., editors, *Imagination and the Scientist*, pages 153–185. Teachers College Press, New York.

Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, 91:417–447.

Shepard, R. N. and Cooper, L. A. (1932). *Mental Images and their Transformations*. MIT Press, Cambridge, MA.

Shore, B. (1997). *Culture in Mind: Cognition, Culture and the Problem of Meaning*. Oxford University Press, New York.

Simon, H. A. (1977). *Models of Thought*. Reidel, Dordrecht.

Simon, H. A. (1981). *The Sciences of the Artificial*. MIT Press, Cambridge, MA.

Suchman, L. A. (1987). *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge and New York, Cambridge University Press.

Thagard, P. (1991). *Conceptual Revolutions*. Princeton University Press, Princeton, NJ.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Harvard University Press, Cambridge, MA.

Tweney, R. D. (1985). Faraday's discovery of induction: A cognitive approach. In Gooding and James, 1985, pages 189–210.

Tweney, R. D. (1992). Stopping time: Faraday and the scientific creation of perceptual order. *Physis*, 29:149–164.

van der Kolk, B., McFarlane, A. C., and Weisaeth, L., editors (1996). *Traumatic Stress: The Effects of Overwhelming Experience on Mind, Body, and Society*. Guilford Press, New York.

Vera, A. and Simon, H. (1993). Situated cognition: A symbolic interpretation. *Cognitive Science*, 17:4–48.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press, Cambridge, MA.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 32:109–123.

Wason, P. C. (1968). On the failure . . . — A second look. In Watson, P. C. and Johnson-Laird, P. N., editors, *Thinking and Reasoning*, pages 307–314. Cambridge University Press, Cambridge.

Williams, L. P. (1964). *Michael Faraday: A Biography*. Basic Books, New York.

Woods, D. D. (1997). Towards a theoretical base for representation design in the computer medium: Ecological perception and aiding human cognition. In Flach, J., Hancock, P., Caird, J., and Vincente, K., editors, *The Ecology of Human: Machine Systems*, pages 157–188. Lawrence Erlbaum, Hillsdale, NJ.

Yeh, W. and Barsalou, L. W. (1996). The role of situations in concept learning. In *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, Lawrence Erlbaum Associates, Hillsdale, NJ.

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, 21:179–217.

Zhang, J. and Norman, D. A. (1995). A representational analysis of numeration systems. *Cognition*, 57:217–295.