Barriers and Models: Comments on Margolis and Giere

Nancy J. Nersessian

Princeton University

Giere's paper calls for nothing less than a "paradigm change" in philosophy of science and provides a overview of what developments in cognitive science might contribute to a new philosophical theory of the growth, development, and change of scientific knowledge. Specifically, he challenges the position - still dominant in post-positivistic philosophy - that the basic representational device of the scientist is a linguistic structure. Rather, he proposes that the basic device is a "model" or "picture" by which I take him to mean some form of structural analog to the events, entities, or processes under investigation. I am in basic agreement with his analysis and with his assessment of current cognitive approaches to analyzing science; especially with his contention that, at least for the present, cognitive psychology has the most to offer a cognitive philosopher trying to understand how models are constructed and function within a scientific community.

As Giere points out, I have for some time been investigating in what ways joining philosophical analyses with those in the cognitive sciences might contribute to a more refined understanding of the nature and processes of scientific development and change. As part of this investigation, I have proposed a new interdisciplinary method - "cognitive-historical analysis" - as an approach to examining the actual practices of scientists who have created and changed representations of nature. The premises that form the basis of that proposal are: (1) the temporal perspective of history is needed to understand development and change in science and (2) the representational and problem-solving practices of scientists are outgrowths of ordinary cognitive resources. I see Margolis' analysis as fitting into the cognitive-historical genre, and thus want to focus my commentary on it. While his analysis is thought provoking, I believe it is fundamentally mistaken as an account of scientific change. And - more importantly for this session - it provides the wrong model for how a cognitive philosopher of science should investigate the nature and processes of scientific change.

Margolis' analysis is mistaken primarily because it gives no role to the scientist as the problem solver who brings about change. It provides the wrong model for a cognitive philosopher because it selectively and uncritically adopts data and speculative hypotheses from the cognitive sciences, gives more weight to these findings and interpretations than the state of these investigations warrant, and tries to fit cases of scien-

PSA 1990, Volume 2, pp. 441-444

tific change into the cognitive accounts. What is wrong with this approach is that it fails to take into account that the cognitive sciences are themselves in an embryonic state. There are no fully-developed theories one can simply take off the shelf and apply to science. In fact, at present, cognitive theories are largely uninformed by sophisticated scientific practices, making the fit between them and cognitive theories something that still needs to be determined. As Giere points out, a major task facing cognitive philosophers is to forge an interpretative framework adequate for understanding scientific change from cognitive analyses coupled with examinations of scientific practices. Margolis' project seems, rather, to be that of fitting the historical data to speculative hypotheses developed from work in the cognitive sciences.

Margolis' purpose is to provide a general model for how paradigm shifts occur in science. He takes as premise the claim of his recent work that all cognition reduces to pattern recognition. This claim emerges from: (1) an argument for how the mechanism of pattern recognition could have evolved in the brain and (2) evidence from psychological studies that humans are bad at drawing logical inferences from propositional representations and quite good at such things as perceptual inferences from diagrams and reasoning via mental models (non-propositional structural analogs of entities, events, and processes). Although I am not evaluating that work, the evidence he cites does not support the reductionist claim: all the empirical studies show is that pattern recognition is an important aspect of human cognition. I point this out here because Margolis' tendency is to generalize beyond what even cognitive scientists would see as the implications of their work. From the premise that all cognition reduces to pattern recognition he develops the thesis that scientific discovery can be explained in terms of the breaching of a cognitive "barrier" that allows recognition of a new pattern.

The "barrier" mechanism of scientific change works as follows. Margolis maintains that operating within a paradigm is exercising a "habit of the mind". Here he is drawing from work in cognitive science that treats the question of how knowledge is represented in the mind. Some theorists have argued that knowledge is represented in organized units variously called "schemata", "scripts", or "frames". The differences among these need not concern us here. The main claim is that specific organizational units are activated when we encounter or think about a situation or problem. Margolis' "habits" are reflex-like ways of understanding that arise from the automatic employment of the knowledge units available to a scientist working in specific paradigm. Additionally, he claims that among these "habits" there exists a unique barrier that prevents us from seeing another available and better pattern that is waiting to be discovered once the barrier is "breached". In his central example, he maintains that the Copernican hypothesis "lay on the table with no one able to see it" for 1400 years. "All the necessary information had been available" to anyone who would have been able to breach the nested spheres barrier of the Ptolemaic framework.

It is important to keep in mind that Margolis' intention is to use findings in the cognitive sciences to give more explicit formulation to some basic Kuhnian insights. Specifically, Kuhn's claim that doing normal science requires working under the exemplars it provides becomes, in Margolis' analysis, the claim that normal science involves exercising habits, i.e., automatically employing the specific schemata arising from the paradigm. These habits condition the expert's deepest intuitions employed in understanding and making judgments and present an obstacle to change. The incommensurability - or failure of communication - Kuhn maintains occurs in the initial stages of a scientific revolution - is, here, due to the psychological "fact" that those deeply entrenched in existing habits find it difficult to breach the barrier and are thus unable to recognize the new pattern even when it is pointed out to them. Kuhn himself offers the metaphor of the "gestalt switch" as a way of understanding these phenomena. It is difficult to see what Margolis offers in his analysis that goes beyond the "gestalt switch" metaphor in characterizing revolutionary change. Just as with that metaphor, Margolis' account makes recognizing a new pattern appear to be something that simply happens to scientists rather than the outcome of an extended period of problem solving and construction by scientists.

This brings us to the heart of the problem. Margolis promises new insights into "the evolution of radically new ideas"; i.e., into the mechanisms of representational change in science. But his mechanism - breaching a cognitive barrier - gives a relatively minor role to the actual activities of the cognitive agents: the scientists. While I agree with him that a radical social constructivist approach cannot explain how new representations arise and become viable competitors to accepted views, the proper response is not to take all construction out of scientific change. There is much historical research to support the view that new scientific representations are constructed in a problem-solving process and that the scientists involved in this process employ a wide range of cognitive strategies for solving problems. Some things Margolis says indicate that he might agree with this, but his central thesis that there is a unique cognitive barrier in every paradigm and that the new patterns are out there waiting in the wings to be seen once this is breached - available, e.g., in the Copernican case for 1400 years - gives no role to the specific problem situation of the scientist in creating the new paradigm.

That seeing a new pattern lies at the heart of major scientific change is obvious. To say scientific change involves seeing new patterns neither explains anything nor gives any insights into the evolution of new ideas. We still want to know why, e.g., Copernicus saw a pattern that Ptolemy failed to recognize. Historically it is simply wrong that nothing had changed in 1400 years and that Copernicus had nothing available to him that was not available to Ptolemy. Among other factors, numerous technical problems had been encountered in those 1400 years and, significantly, the equant that Ptolemy had introduced as the center of motion of the system was seen by Copernicus as conflicting with the neo-platonic desideratum of uniform motion about the geometrical center of a planet's orbit. As a result, Copernicus' framing of the problem situation differed significantly from that of Ptolemy. Contrary to Margolis' analysis, a new pattern only becomes available when there is a set of problems to which it seems to offer a solution and whatever might constitute a cognitive barrier in a situation is relative to a specific framing of the problem situation. The new pattern is constructed as a solution to the new framing of the problem and it is the structure of this framing that enables specific patterns to emerge as possible solutions.

Since Margolis aims to provide a general account of scientific change I will underscore this point by way of different historical case. What can Margolis' analysis provide in the way of illumination about why Einstein saw the pattern of the special theory of relativity that many of his contemporaries came so close to but failed to recognize? Margolis' response should be that he was able to breach the aether barrier. But this response does not explain why Einstein was able to do this and how the new pattern emerged. My answer would be: Einstein framed the problem situation in a fundamentally different way. The new framing came from a process of reflection on and reassessment of the foundations upon which classical mechanics and electromagnetism rested. His framing of the problem situation included: (1) the failure of attempts to reduce electromagnetism to mechanics and vice versa; (2) Planck's analysis of blackbody radiation, which indicated to him that both electromagnetism and mechanics are inadequate in regions small enough for fluctuation phenomena to count; and (3) specific problems with Lorentz' electron theory. Relative to this framing of the problem situation, Einstein claimed the solution was to put electromagnetism and mechanics on an equal footing. In this case the aether becomes "superfluous", i.e., the barrier, and the special theory of relativity, the new pattern.

There are several historical puzzles Margolis' framework would need to explain relating to this case. One puzzle is that Faraday - with a quite different framing of the problem situation - recognized and breached the aether barrier. Yet, this did not lead him to the pattern of the special theory of relativity. Faraday's field theory is not that of Einstein. Another is that Lorentz recognized the aether barrier (and understood the special theory) but refused to breach it. Among other reasons, for him, a satisfactory explanation of the contraction of moving rods had to include a causal explanation for the phenomenon and this involved understanding the interaction of the aether with the molecular forces in the rod. Historical data such as these cannot be accommodated by Margolis' model of scientific change.

In sum, I find the model of scientific change presented by Margolis offers little insight into the processes of representational change. This being said, I do think the cognitive sciences provide a valuable resource for philosophy and history of science. In particular, the cognitive sciences can contribute to inquires into the dynamic forces through which a vague speculation gets articulated into a new scientific theory, gets communicated to other scientists, and comes to replace existing representations of a domain. As Giere points out, ongoing investigations in cognitive psychology are contributing new insights into the agency of individual scientists in constructing new representations of nature. These will aid philosophers in their attempts to construct new theories of development, evaluation, and choice of scientific theories. They will also enrich historical analyses of specific instances of scientific creativity, whether the innovation led down a dead end, to "winning" science, to the same point via different routes, etc.

Putting emphasis on the individual scientist, though, by no means requires a "new-fashioned" internalist analysis, as critics such as Fuller have claimed. Nature may endow the individual with specific cognitive mechanisms, but nurture influences acquisition, development, and use of cognitive strategies and conceptual resources. Investigations are underway as to how "situated" cognition is and also to what role interests and motivations play in cognition. Historical and sociological research over the last twenty years has shown us that there is no denying the importance of the facts that scientists are situated within scientific communities and wider social contexts and that they acquire cognitive strategies and conceptual resources from both of these contexts. Indeed, what makes cognitive-historical analysis so attractive is its potential to provide a much-needed synthesis of the cognitive and the social dimensions of knowledge construction.