CURRENT DIRECTIONS IN PSYCHOLOGICAL SCIENCE

What Have Psychologists (And Others) Discovered About the Process of Scientific Discovery?

David Klahr[1] and Herbert A. Simon

Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania

Abstract

We describe four major approaches to the study of science—historical accounts of scientific discoveries, psychological experiments with nonscientists working on tasks related to scientific discoveries, direct observation of ongoing scientific laboratories, and computational modeling of scientific discovery processes—by viewing them through the lens of the theory of human problem solving. We compare and contrast the different approaches, indicate their complementarities, and provide examples from each approach that converge on a set of principles of scientific discovery.

Keywords

scientific discovery; problem solving

Early in the 20th century, Einstein, in reflecting on his own mental processes leading to the theory of relativity, said, "I am not sure whether there can be a way of really understanding the miracle of thinking" (Wertheimer, 1945, p. 227). However, in the past 25 years, several disciplines, including psychology, history, and artificial intelligence,[2] have produced a substantial body of knowledge about the process of scientific discovery that allows us to say a great deal about it.[3] Although the strengths of one approach are often the weaknesses of another, the work has collectively yielded consistent insights into the scientific discovery process.

ASSESSING THE FOUR MAJOR APPROACHES

Historical accounts of the great scientific discoveries—typically based on diaries, scientific publications, autobiographies, lab notebooks, correspondence, interviews, grant proposals, and memos of famous scientists—have high face validity. That is, it is clear that they are based on what they purport to study: real science. However, such studies have some weaknesses. For one thing, their sources are often subjective and unverifiable. Moreover, the temporal resolution of historical analysis is often coarse, but it can become much finer when laboratory notebooks and correspondence are available. Historical investigations often generate novel results about the discovery process, by focusing on a particular scientist and state of scientific knowledge, as well as by highlighting social and motivational factors not addressed by other approaches.

Although historical studies of discovery focus much more on successes than on failures, laboratory studies are designed to manipulate the discovery context in order to examine differences in processes associated with success and failure. Face validity of lab studies varies widely: from studies only distantly related to real scientific tasks to those that model essential aspects of specific scientific discoveries (e.g., Dunbar's, 1993, simulated molecular genetics laboratory; Schunn & Anderson's, 1999, comparison of experts' and novices' ability to design and interpret memory experiments; Qin & Simon's, 1990, study in which college sophomores rediscovered Kepler's third law of planetary motion). Laboratory studies tend to generate fine-grained data over relatively brief periods and typically ignore or minimize social and motivational factors.

The most direct way to study science is to study scientists in their day-to-day work, but this is extraordinarily difficult and time-consuming. A recent example is Dunbar's (1994) analysis of discovery processes in several world-class molecular genetics research labs. Such studies have high face validity and potential for detecting new phenomena. Moreover, they may achieve much finer-grained temporal resolution of ongoing processes than historical research, and they provide rigor, precision, and objectivity that is lacking in retrospective accounts.

A theory of discovery processes can sometimes be incorporated in a computational model that simulates and reenacts discoveries. Modeling draws upon the same kinds of information as do historical accounts, but goes beyond history to hypothesize cognitive mechanisms that can make the same discoveries, following the same path. Modeling generates theories and tests them against data obtained by the other methods. It tests the sufficiency of the proposed mechanisms to produce a given discovery and allows comparison between case studies, interpreting data in a common language to reveal both similarity and differences of processes. Modeling enables us to express a theory rigorously and to simulate phenomena at whatever temporal resolution and for whatever durations are relevant.

SCIENTIFIC DISCOVERY AS PROBLEM SOLVING

Crick argued that discoveries are major when they produce important knowledge, whether or not they employ unusual thought processes: "The path [to the double helix]. . . was fairly commonplace. What was important was *not the way it was discovered*, but the object discovered—the structure of DNA itself" (Crick, 1988, p. 67; italics added). Psychologists have been making the case for the "nothing special" view of scientific thinking for many years. This does not mean that anyone can walk into a scientist's lab and make discoveries. Practitioners must acquire an extensive portfolio of methods and techniques, and must apply their skills aided by an immense base of shared knowledge about the domain and the profession. These components of expertise constitute the *strong methods*. The equally important *weak methods* scientists use underlie all human problem-solving processes.

A problem consists of an initial state, a goal state, and a set of operators for transforming the initial state into the goal state by a series of intermediate steps. Operators have constraints that must be satisfied before they can be applied. The set of states, operators, and constraints is called a *problem space*, and the problem-solving process can be characterized as a search for a path that links initial state to goal state.

Initial state, goal state, operators, and constraints can each be more or less well-defined. For example, one could have a well-defined initial state and an ill-defined goal state and set of operators (e.g., make "something pretty" with these materials and tools), or an ill-defined initial state and a well-defined final state (e.g., find a vaccine against HIV). But well-definedness depends on the familiarity of the problem-space elements, and this, in turn, depends on an interaction between the problem and the problem solver.

Although scientific problems are much less well-defined than the puzzles commonly studied in the psychology laboratory, they can be characterized in these terms. In both cases, well-definedness and familiarity depend not only on the problem, but also on the knowledge that is available to the scientist. For that reason, much of the training of scientists is aimed at increasing the degree of well-definedness of problems in their domain. The size of a problem space grows exponentially with the number of alternatives at each new step in the problem (e.g., the number of possible paths one must consider at each possible move when planning ahead in chess). Effective problem solving must constrain search to only a few such paths. Strong methods, when available, find solutions with little or no search. For example, in chess, there are many standard openings that allow experts to make their initial moves with little search. Similarly, someone who knows algebra can use simple linear equations to choose between two sets of fixed and variable costs when deciding which car to rent instead of painstakingly considering the implications of driving each car a different distance. But the problem solver must first recognize the fit between the given problem (renting a car) and the strong method (high school algebra).

Weak methods, requiring little knowledge of a problem's structure, do not constrain search as much. One particularly important weak method is analogy, which attempts to map a new problem onto one previously encountered, so that the new problem can be solved by a known procedure. However, the mapping may be quite complex, and it may fail to produce a solution.

Analogy enables the problem solver to shift the search from the given problem space to one in which the search may be more efficient, sometimes making available strong methods that greatly abridge search. Prior knowledge can then be used to plan the next steps of problem solving, replace whole segments of step-by-step search, or even suggest an immediate solution. The recognition mechanism uses this store of knowledge to interpret new situations as instances of previously encountered situations. This is a key weapon in the arsenal of experts and a principal factor in distinguishing expert from novice performance.

In the past 25 years, analogy has assumed prominence in theories of problem solving and scientific discovery. Nersessian (1984) documented its role in several major 19th-century scientific discoveries. Recent studies of contemporary scientists working in their labs have revealed the central role played by analogy in scientific discovery (Dunbar, 1994; Thagard, 1998).

Although many strong methods are applied in scientific practice, weak methods are of special interest for scientific discovery because they are applicable in a wide variety of contexts, and strong methods become less available as the scientist approaches the boundaries of knowledge.

COMPLEMENTARITY OF APPROACHES

Viewing scientific discovery as problem solving provides a common language for describing it and facilitates studying the same discovery using more than one approach.

In the late 1950s, Monod and Jacob discovered how control genes regulate the synthesis of lactose (a sugar found in milk) in bacteria (Jacob & Monod, 1961). The literature explaining this discovery (e.g., Judson, 1996) tends to use terms such as "a gleam of perception," but to characterize a discovery as a gleam of perception is to not describe it at all. One must identify specific and well-understood cognitive processes and then determine their role in the discovery. Among the most important steps for Jacob and Monod in discovering the mechanisms of genetic control were representational changes that enabled them to replace their entrenched idea—that genetic control must involve some kind of activation—with the idea that it employed inhibition instead.

Dunbar (1993) created a laboratory task that captured important elements of Monod and Jacob's problem, while simplifying to eliminate many others. His participants—asked to design and run (simulated) experiments to discover the lactose control mechanism—faced a real scientific task with high face validity. Although the task was simplified, the problem, the "givens," the permissible research methods, and the structure of the solution were all preserved. Dunbar's study cast light on the problem spaces that Monod and Jacob searched, and on some of the conditions of search that were necessary or sufficient for success (e.g., knowing that there was such a thing as a control gene, but not exactly how it worked).

In this example, a historically important scientific discovery provided face validity for the laboratory study, and the laboratory provided information about the discovery processes with fine-grained temporal resolution.

CONVERGENT EVIDENCE OF DISCOVERY PRINCIPLES

In this section, we give a few additional examples of convergent evidence obtained by using two or more approaches to study the same discovery.

Surprise

Recently, reigning theories of the scientific method have generally taken hypotheses as unexplained causes that motivate experiments designed to test them. In this view, the hypotheses derive from scientists' "intuitions," which are beyond explanation. Historians of science have taken a less rigid position with respect to hypotheses, and include their origins within the scope of historical inquiry.

For example, the discovery of radium by the Curies started with their attempt to obtain pure radioactive uranium from pitchblende. As they proceeded, they were surprised to find in pitchblende levels of radioactivity higher than in pure uranium. As a surprise calls for an explanation, they conjectured that the pitchblende contained a second substance (which they named radium) more radioactive than uranium. They succeeded in extracting the radium and determined its key properties. In this case, a phenomenon led to a hypothesis, rather than a hypothesis leading to an experimental phenomenon. This occurs frequently in science. A surprise violates prior expectations. In the face of surprise, scientists frequently divert their path to ascertain the scope and import of the surprising phenomenon and its mechanism.

Response to surprise was investigated in a laboratory study (Klahr, Fay, & Dunbar, 1993) in which participants had to discover the function of an unknown key on a simulated rocket ship. They were given an initial hypothesis about how the key worked. Some participants were given a plausible hypothesis, and others were given an implausible hypothesis. In all cases, the suggested hypothesis was wrong, and the rocket ship produced some unexpected, and sometimes surprising, behavior. Adults reacted to an implausible hypothesis by proposing a competing hypothesis and then generating experiments that could distinguish between them. In contrast, children (third graders) tended to dismiss an implausible hypothesis and ignore evidence that supported it. Instead, they attempted to demonstrate that their favored hypothesis was correct. It seems that an important step in acquiring scientific habits of thinking is coming to accept, rather than deny, surprising results, and to explore further the phenomenon that gave rise to them.

Krebs's biochemical research leading to the discovery of the chain of reactions (the reaction path) by which urea (the end product of protein metabolism) is synthesized in the body has been the topic of convergent studies focusing on response to unexpected results. The discovery has been studied historically by Holmes (1991) and through the formulation of two computational models (Grasshoff & May, 1995; Kulkarni & Simon, 1988), both of which have modeled the discovery. After the models proposed an experiment and were given its outcome, they then proposed another experiment, using the previous outcomes to guide their decision about what sort of experiment would be useful. Using no more knowledge of biochemistry than Krebs possessed at the outset, both programs discovered the reaction path by following the same general lines of experimentation as Krebs followed. One of these models (Kulkarni & Simon, 1988) addressed the surprise issue directly (in this case, surprise in finding a special catalytic role for the amino acid ornithine). The simulated scientist formed expectations (as did Krebs) about experimental outcomes. When the expectations were violated, steps were taken to explain the surprise. Thus, historical studies, simulation models, and laboratory experiments all provide evidence that the scientist's reaction to phenomena—either observational or experimental—that are surprising can lead to generating and testing new theories.

Multiple Search Spaces

This reciprocal relation between hypotheses and phenomena arises in laboratory studies, historical studies, and computational models of discovery, enabling us to characterize scientists' thinking processes as problem-solving search in multiple spaces.

Dual Search

The discovery process can be characterized as a search in two spaces: a hypothesis space and an experiment space. When attempting to discover how a particular control button worked on a programmable device, participants in the "rocket ship" study described earlier (Klahr & Dunbar, 1988) had to negotiate this dual search by (a) designing experiments to disclose the button's functions (searching the experiment space) and (b) proposing rules that explained the device's behavior (searching the hypothesis space). Thus, participants were required to coordinate two kinds of problems, and they approached this dual search with different emphases. Some ("experimenters") focused on the space of possible manipulations, whereas others ("theorists") focused on the space of possible explanations of the responses.

Historical studies usually reveal both hypothesis-space search and experiment-space search. For example, most histories of Faraday's discovery of induction of electricity by magnets place much emphasis on the influence of Ampère's theory of magnetism on Faraday's thought. However, a strong case can be made that Faraday's primary search strategy was in the space of experiments, his discovery path being shaped by phenomena observed through experimentation more than by theory.

The number of search spaces depends on the nature of the scientific problem. For example, in describing the discovery of the bacterial origins of stomach ulcers, Thagard (1998) demonstrated search in three major spaces: hypothesis space, experiment space, and a space of instrumentation.

Analogy in Search for Representations

Bohr used the solar system analogy to arrive at his quantum model of the hydrogen atom. He viewed the electrons in the hydrogen atom as planets orbiting the nucleus, although, according to classical understanding of the solar system, this would mean that the charged electrons would dissipate energy until they fell into the nucleus. Instead of abandoning the analogy, Bohr borrowed Planck's theory that energy could be dissipated only in quantum leaps, then showed that these leaps would produce precisely the spectrum of light frequencies that scientists 30 years previously had demonstrated hydrogen produces when its electrons move from a higher-energy stationary state to a lower-energy one.

Search in the Strategy Space

Finally, changes in strategy, even while the representation of a problem is fixed, may enable discovery. Often the change in strategy results from, or leads to, the invention of new scientific instruments or procedures. Breeding experiments go back to Mendel (and experiments for stock breeding go much further back), but the productivity of such experiments depended on mutation rates. Müller, with the "simple" idea that x-rays could induce higher rates of mutation, substantially raised that productivity.

CREATIVITY AND PROBLEM SOLVING IN SCIENCE AND BEYOND

Scientific discovery is a type of problem solving using both weak methods that are applicable in all disciplines and strong methods that are mainly domain-specific. Scientific discovery is based on heuristic search in problem spaces: spaces of instances, of hypotheses, of representations, of strategies, of instruments, and perhaps others. This heuristic search is controlled by general mechanisms such as trial-and-error, hill-climbing, means-ends analysis, analogy, and response to surprise. Recognition processes, evoked by familiar patterns in phenomena, access knowledge and strong methods in memory, linking the weak methods to the domain-specific mechanisms.

All of these constructs and processes are encountered in problem solving wherever it has been studied. A painter is not a scientist; nor is a scientist a lawyer or a cook. But they all use the same weak methods to help solve their respective problems. When their activity is described as search in a problem space, each can understand the rationale of the other's activity, however abstruse and arcane the content of any special expertise may appear.

At the outer boundaries of creativity, problems become less well structured, recognition becomes less able to evoke prelearned solutions or domain-specific search heuristics, and more reliance has to be placed on weak methods. The more creative the problem solving, the more primitive the tools. Perhaps this is why "childlike" characteristics, such as the propensity to wonder, are so often attributed to creative scientists and artists.

Recommended Reading

Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge, MA: MIT Press.

Klahr, D., & Simon, H.A. (1999). Studies of scientific discovery: Complementary approaches and convergent findings. *Psychological Bulletin*, *125*, 524–543.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, *20*, 99–149.

Acknowledgments—Preparation of this article and some of the work described herein were supported in part by a grant from the National Institute of Child Health and Human Development (HD 25211) to the first author. We thank Jennifer Schnakenberg for a careful reading of the penultimate draft.

Notes

1. Address correspondence to David Klahr, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213.

2. In addition, the sociology of science explains scientific discovery in terms of political, anthropological, or social forces. The mechanisms linking such forces to scientific practice are usually motivational, social-psychological, or psychodynamic, rather than cognitive. Although this literature has provided important insights on how social and professional constraints influence scientific practices, we do not have much to say about it in this brief article.

3. This article summarizes an extensive review listed as the second recommended reading. Full references to historical sources alluded to in the present article can be found there, as well as in the first recommended reading. The third recommended reading focuses on developmental aspects of the discovery process.

References

Crick, F. (1988). *What mad pursuit: A personal view of scientific discovery*. New York: Basic Books.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science*, *17*, 397–434.

Dunbar, K. (1994). How scientists really reason: Scientific reasoning in real-world laboratories. In R.J. Sternberg & J. Davidson (Eds.), *The nature of insight* (pp. 365–395). Cambridge, MA: MIT Press.

Grasshoff, G., & May, M. (1995). From historical case studies to systematic methods of discovery: Working notes. In *American Association for Artificial Intelligence Spring Symposium on Systematic Methods of Scientific Discovery* (pp. 45–56). Stanford, CA: AAAI.

Holmes, F.L. (1991). *Hans Krebs: The formation of a scientific life, 1900-1933, Volume 1*. New York: Oxford University Press.

Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, *3*, 318–356.

Judson, H.F. (1996). *The eighth day of creation: Makers of the revolution in biology* (expanded ed.). Plainview, NY: Cold Spring Harbour Laboratory Press.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, *12*, 1–55.

Klahr, D., Fay, A.L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology*, *24*, 111–146.

Kulkarni, D., & Simon, H.A. (1988). The process of scientific discovery: The strategy of experimentation. *Cognitive Science*, *12*, 139–176.

Nersessian, N.J. (1984). *Faraday to Einstein: Constructing meaning in scientific theories*. Dordrecht, The Netherlands: Martinus Nijhoff.

Qin, Y., & Simon, H.A. (1990). Laboratory replication of scientific discovery processes. *Cognitive Science*, *14*, 281–312.

Schunn, C.D., & Anderson, J.R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science*, *23*, 337–370.

Thagard, P. (1998). Ulcers and bacteria: I. Discovery and acceptance. *Studies in the History and Philosophy of Biology and Biomedical Science*, *9*, 107–136.

Wertheimer, M. (1945). *Productive thinking*. New York: Harper & Row.