What is a Systematic Method of Scientific Discovery?

Herbert A. Simon

Department of Psychology

Carnegie Mellon University

Pittsburgh, PA 15213

Creativity in general, and scientific discovery as a particular form of creativity, has always been surrounded by an aura of mysticism. We continue to recount the mystical experiences of Kekule with his benzene ring, or Poincare stepping on the bus at Caen, or Archimedes running naked from his bath, shouting "Eureka!". Attempts to explain discovery in terms of normal thought processes are regretted or even resented as assaults on the mystery. Nevertheless, there is mounting evidence that scientific discovery is explicable – that there is method in the madness of the moment of inspiration or illumination.

The Evidence

The evidence of method takes several forms. The first, and perhaps most important to date, consists of careful and detailed historical analyses of specific discoveries. An excellent example is Larry Holmes' account of Hans Krebs' discovery of the ornithine cycle for the in vivo synthesis of urea (later simulated by the computer program KEKADA).

The second source of evidence comes from computer simulations that are able to re-create scientific discoveries of historical importance when they are given the same initial conditions as the original discoverer. The BACON program is an example of a system that has done this for a dozen or more discoveries of high significance, ranging from Kepler to Cannizzaro, and a number of other programs of this general kind have also been tested successfully.

A third kind of evidence is the demonstration that human subjects can re-create important discoveries from the same initial conditions as the original discoverer. One such demonstration was the rediscovery of Kepler's Third Law by four of fourteen college students who were given the (unidentified) data and asked to find a function to fit it. Another was the rediscovery (Klahr and Dunbar; Okada) of the Jacob-Monod-Lwoff mechanism of control genes by college students in the laboratory, working either singly or in pairs.

A fourth kind of evidence, all too rare, is automatic production of new discoveries by computer programs. There are some such examples of discoveries of original proofs for theorems in mathematics and logic, of reaction paths in chemistry, and perhaps others.

The Theory Supported by the Evidence

If one were to summarize what we have learned from this body of research it is that the mental processes leading to first-rank scientific discoveries (and presumably to those of lesser rank as well) are the same processes that have emerged from research on all other kinds of problem solving. The process of scientific discovery is a normal problem-solving process that can be summed up as a combination of heuristic search in one or more problem spaces and recognition of cues in a problem space.

Heuristic search in problem spaces is a familiar story in AI and Cognitive Science. The new chapter that has been contributed by research on discovery is that more than one problem space is usually involved: a space of observations, a space of hypotheses, a space of possible problem representations, and possibly spaces of possible instruments and research techniques, and others. Quite frequently, first-magnitude discoveries must await the prior discovery of an effective problem representation (e.g., differential equations), or an instrument (e.g., the thermometer, the ammeter, or a more powerful particle accelerator).

Recognition processes play an especially important role in the discoveries that are most closely associated with identifiable "aha's." Surprise, caused by an unexpected observation, evokes in good scientists a "surprise heuristic": first, delimit the surprising phenomenon, then search out its mechanism. Examples are legion: Faraday exploiting the unexpected production of an electrical current by the movement of a magnet; the Curies, surprised by an excess of radiation (more intense than pure uranium) while refining pitchblende to obtain uranium; Fleming, surprised by the lysing of bacteria in his Petri dish by a penicillium mold on its edge; Krebs, surprised by the volume of urea secreted upon the addition of a little ornithine to the ammonia. Each reacted to the surprise with an "aha"; each delimited the phenomena; each continued to search out the mechanism.

Surprise-induced recognition means the appearance of events that are contrary to expectations. No surprise without expectations, and no expectations without knowledge. Only a geologist can extract a fact from a rock, or, as Pasteur put it: "Accidents happen to the prepared mind."

This, then, is the outline of our contemporary theory of discovery: heuristic search in many spaces combined with knowledge-based recognition of cues, recognition that sometimes produces exploitable surprise.

Normative and Descriptive Theory

So far, we have been talking about how discoveries actually take place – a descriptive or explanatory theory of discovery. But the evidence for this theory has come from the examination of how scientists of first rank have actually made their discoveries: what strategies they have carried out, how they have responded to the unusual and unexpected. The theory, then, is a theory of how "good" science is done. Bunglers, or the unproductive, may proceed quite differently: they haven't been much studied.

What about a normative theory of scientific discovery? Can we construct from the descriptive theory a body of advice to aspiring scientists? Of course we can. The first principle is to do the kinds of things that successful scientists do. How can we turn this rather general kind of advice into something more specific? Let me propose some examples:

1. *Pick a good research problem.* A problem is good if (1) someone would care if you found an answer; and (2) you have some ideas about the first steps you could take in looking for it.

2. *Pick a problem you can solve first.* Often, if a problem is good (unless you discovered it through a surprising accident), many other people will also be seeking the answer. You will only have a chance of solving it first (second doesn't count) if you have a secret weapon.

   You have a secret weapon if you are much smarter than the other people who might work on the problem, or willing to work longer hours – but these are unlikely advantages to count on!

   You have a secret weapon if you have access to research tools others don't have (e.g., the first or biggest computer; a new technique for studying biological processes in vivo; a surprising phenomenon others haven't noticed; a different background from most people in the domain).

   Corollary: Avoid "hot" problems from the current literature. Everyone else also knows they are hot.

   Corollary: If you must choose a hot topic, put a cot in your office and have someone bring in your meals.

3. *Many (most?) good problems come from the "real world" – e.g., applications areas or puzzling phenomena.* There is no opposition between "basic" and "applied."

4. *When is a problem solved?* It is solved when you have a publishable research report. The solution of one problem is usually the beginning of another (on which you have a head start). All science is a series of progress reports.

In summary, a normative theory of discovery recommends picking a good question, whose answer people care about, and for which you have a secret weapon. It recommends acquiring lots of knowledge ("indexed" so it will be recalled when relevant – sometimes producing surprise). It recommends following procedures that will cause you to encounter phenomena. This can be done by designing a strategy, experimenting, exploring – and watching for the surprise. It recommends that you keep whittling at your problem representation until you find a clean way to think about it. It recommends that, as soon as you have found or demonstrated something, you publish your progress report.

It may be objected that this is a poor excuse for a theory: that its assertions do not follow deductively from its premises, hence it does not guarantee that following its prescriptions will produce outstanding (or even good) science. Quite so. It is an empirically based theory, an induction from observations about the world. It is, in a broad sense, a probabilistic theory: it tells the scientist how to improve his or her chances. Science is a gambling game, and those who are not gamblers should probably not remain in it.

But the theory is in the best tradition of normative theories. It gives advice about what to do in a world that is full of surprises, some of them welcome, and some fatal. The advice, like all advice, is valid to the extent that our theory of that world resembles reality.

Systematic Methods

We have pointed out reasons for believing that there is method in scientific discovery, and have sketched in broad brush what a normative theory of discovery looks like. This symposium examines systematic discovery methods in many fields of science; some methods are specific and implemented on computers, while others are of broader scope and are suitable, so far, only for human scientists to execute (our normative theory above is of the latter type). What, then, is a systematic discovery method in a specific field of science, organic chemistry say, or cognitive psychology?

The symposium, and the following pages, will propose answers to this question that are mutually supportive and based upon a substantial number of empirical studies of discovery in a wide range of domains.