Essays of an Information Scientist, Vol:4, p.488-493, 1979-80 Current Contents, #21, p.5-10, May 26, 1980

Current Comments

Premature Discovery or Delayed Recognition—Why?

Number 21 May 26, 1980

Most scientists can name an example of an important discovery that had little initial impact on contemporary research. Mendel's work is the classic example.[1] But even more interesting examples abound. For example, Karl Jansky discovered in 1933 that some of the radio static interfering with Bell Telephone's transatlantic circuits came from the Milky Way.[2] This discovery was ignored by Bell and the scientific community because they believed that the galaxy was too weak a radio source to detect.[3] However, an amateur astronomer, Grote Reber, published the first "radio maps" of the skies in the early 1940s.[4] When World War II ended, a new scientific field—radio astronomy—was born, more than ten years after Jansky made his discovery.

Apart from the fact that many scientists believe that *their* work is slow in being recognized by their peers, there seems to be some real basis for believing that delayed recognition or "premature discovery" is inevitable—especially in recent decades. There is a widespread belief that the information explosion, by clogging information channels or by sheer overload, prevents important ideas from penetrating the wall of established thinking or wisdom.

I've wondered about this problem for a long time. I've often thought we could identify some of these ignored ideas through citation analysis. While most highly-cited papers are recognized quite early, there are a small number of "slow starters." Why is it that these "premature" papers achieve recognition so many years after publication? Why do they go unnoticed for so long? Or did they? Certainly Gregor Mendel's work did not go unnoticed, but its significance was not appreciated for over 30 years.

I hope that citation analysis may draw attention to papers that were suddenly discovered or even "rediscovered" after several years of dormancy. This would be more systematic than simply asking everyone to cite interesting examples. By computer monitoring of citation patterns we can detect abrupt increases in citation.

By identifying enough of those instances, historians and sociologists of science may better understand the process of scientific innovation. For example, is it a problem of the *Zeitgeist* or is it simply that some scientists or inventors don't persevere in the "promotion" of their ideas? Hans Selye is a prime example. He made every possible effort to republish his ideas on stress[5] wherever an opportunity arose. How often does "rediscovery" simply mean that the author in question used the power of oral and written persuasion to convince enough people to try the ideas?

Derek Price, historian of science, Yale University, points out that an author's communication skills have a powerful effect on the recognition of one's papers. "Many scientists happen to be poor communicators, perhaps in their writing style or their expository style in mathematics. Or they have poor skill in writing in the right way for the right journal."[6] Price suggests his paper on cumulative advantage distribution[7] is suffering delayed recognition because it involves a heavy mathematical discussion. Clarity in writing is obviously important in communicating scientific ideas.[8]

Papers that initially are unappreciated but later go on to be recognized as significant are referred to as premature discoveries[9,10] or resisted discoveries.[11] Sociologist Stephen Cole, State University of New York at Stony Brook, suggests the more general term *delayed recognition.*[12] Resistance and prematurity are considered to be special subsets of the general phenomenon of delayed recognition. Whatever the term used to describe delayed recognition, the result is the same—significant discoveries are unused or unappreciated until they are "rediscovered" years later. In some instances this process may be gradual and in others it is sudden. Finding enough examples on which to draw some conclusions is not easy.

The phenomenon of delayed recognition has been investigated by several researchers.[9-12] Cole suggests two basic reasons for delayed recognition.[12] One reason is that the paper reports findings that don't agree with current, accepted theory. Another possible reason is related to the author's position in the hierarchy of science. The paper may initially be ignored because its author is a young researcher working in a relatively lesser-known institution.

Cole investigated the sources of delayed recognition to determine whether a paper's content or the author's visibility contributed more to the delay.[12] Using the *Science Citation Index® (SCI®),* he analyzed 74 papers published in all fields of science before 1961 which received three or fewer citations in 1961 and ten or more citations in 1966. These papers, considered as examples of delayed recognition, were compared with 513 papers published before 1961 which received more than three citations in 1961 and at least ten citations in 1966. Cole found that so-called "social stratification" indicators like age, institutional affiliation, and number of honorific awards "fail to distinguish the authors of delayed recognition papers from those not experiencing delayed recognition."[12] He concluded that delayed recognition was primarily the result of content rather than the author's prestige.

Gunther Stent, professor of molecular biology, University of California, Berkeley, suggests that a paper will not be recognized immediately if its content cannot *conceptually* "be connected by a series of simple logical steps to canonical, or generally accepted, knowledge."[9] For example, Hannes Alfven proposed in 1937 that there was a galactic magnetic field created by electrical currents traveling through a small amount of ionized gas in space. However, this proposition could not conceptually be connected with the "well known fact" that space was a vacuum incapable of supporting electrical currents.[13]

In 1975, H.V. Wyatt, professor of biology, University of Bradford, UK, suggested that a paper will initially be ignored if its content cannot be extended *experimentally* to prevailing knowledge for technical reasons.[10] For example, in 1909 Simon Flexner identified poliovirus after injecting cerebrospinal fluid from patients into the spinal canal of monkeys. Flexner insisted that poliovirus could be cultivated only in nerve tissue of monkeys. This limited the number of scientists who could afford to work in polio research. However, in 1949 it became technically feasible to grow poliovirus in non-neural tissue of different animals. An orally administered vaccination against poliovirus was developed soon thereafter.[14]

Mendel's work on plant hybridization, which I've discussed before,[1] is a good example of how conceptual content and experimental method cause delayed recognition.[9] On the basis of his crossbreeding experiments with peas, Mendel discovered in 1866 the principle of segregation of hereditary units. However, Mendel's principle couldn't conceptually be connected with the prevailing scientific opinion that hereditary units were *blended* together when the blood of the parents intermingled in the offspring.[15]

Also, Mendel derived the principle of segregation from a statistical interpretation of his crossbreeding experiment results. Few of Mendel's contemporaries believed mathematics had any legitimate application to biological science. Thus, Mendel's results couldn't experimentally be confirmed and extended to other species of plants and animals because statistical methodology was not accepted in biological research.[9]

Mendel's 1866 paper remained unappreciated for 34 years until it was "rediscovered" in 1900. During that time, a number of advances provided the missing logical steps that finally connected Mendel's paper with canonical scientific knowledge. For example, staining techniques showed that chromosomes are a regular feature of the nucleus, which gave Mendel's abstract units a physical reality.[16] Also, it was observed under the microscope that chromosomes undergo *segregation* during the processes of mitosis and meiosis.[16] Lastly, biologists slowly accepted statistical methodology, and a journal devoted to mathematical biology, *Biometrika*, began publication in 1901.[11]

Sociologist Augustine Brannigan, University of Calgary, Alberta, Canada, argues that Mendel's work was rediscovered in 1900 mainly as the result of a priority dispute between Carl Correns, Hugo DeVries, and Erich Tschermak.[17] These researchers independently duplicated Mendel's segregation results, and DeVries was first to publish "his" laws. Correns quickly named the discovery Mendel's Law to both neutralize his loss of priority and undermine DeVries' claim to the discovery. Price suggests that delayed recognition discoveries are actually a special case of multiple discovery, in which one of the members is much earlier in time than another.[6] The priority dispute arising from multiple discovery can be settled when much older research is resurrected. Psychologist Dean Simonton, University of California, Davis, studied the statistical probability of multiple discoveries.[18]

There is a prevalent myth that a 1944 paper by Oswald Avery and colleagues[19] is a case of prematurity. Avery reported the results of studies that indicated DNA was basically the genetic material. Wyatt asserts that Avery's paper "was known as an isolated and interesting piece of information which could not immediately be assimilated as knowledge."[10] Stent agrees that "geneticists did not seem able to do much with [Avery's work] or build on it."[9] Rollin Hotchkiss, Rockefeller University, has recently published a detailed rebuttal of this view.[20] Joshua Lederberg, now president of Rockefeller University, says the significance of Avery's "formidable precedent" was recognized almost immediately, and some of the "most influential geneticists" of the time accepted the identification of "gene" with DNA fragment.[21]

Harriet Zuckerman, professor of sociology, Columbia University, points out that delayed recognition occurs in many different fields, and even involves work that would later win the Nobel prize.[22] For example, Hermann Staudinger won the Nobel prize in chemistry in 1953 for his work on high molecular weight polymers, like rubber, starch, and cellulose. When he first suggested in 1922 that these polymers consisted of long chains, his colleagues preferred to continue believing that they were actually aggregates of small molecules. In fact, "his colleagues ribbed him about it, asking him to show them how long his macromolecules were, as one would a fish that got away."[23]

Peyton Rous was awarded the 1966 Nobel prize in medicine for his work on the cancer virus that bears his name—Rous sarcoma. But Rous discovered the virus in 1910! Only after a leukemia virus was isolated in 1951 could Rous' discovery be appreciated. The rediscovery of Rous sarcoma virus started the virus-theory trend in cancer research, and the Nobel committee noted that Rous' discovery increased in significance every year since 1951.[24]

Julius Comroe, Cardiovascular Research Institute, University of California, San Francisco, lists 132 discoveries in clinical medicine that were effectively applied after delays ranging from one month to more than 350 years.[25] His report is a valuable source for researchers interested in examining specific instances of delayed recognition of scientific discovery in medicine.

All the examples of delayed recognition I detailed—Mendel, Jansky, Staudinger, and Rous—involve discoveries made between 25 and 100 years ago. Since then, scientific institutions have become more established, more prominent, and more public. The checks and balances that are now built into science presumably make it unlikely that a significant breakthrough will go unnoticed for very long. Cole points out that it "would be unlikely to find a contemporary 'Mendel' working in an obscure monastery; modern day 'Mendels' would be in a university science department, or a government or industrial laboratory. Also we know...that the evaluation system of science operates so efficiently that most 'Mendels' would be in the top university departments."[12]

On the other hand, science literature has grown exponentially since the late 1950s. The volume of research generated in today's university, government, and industry laboratories is enormous. Does this increase the probability that *some* important work, even by well-established investigators, may go unnoticed? While researchers tend to cite more papers now than they did before, e.g., in biochemistry,[26] does the sheer bulk of literature increase uncitedness? Or is it all because there is so much trivia published that is not worth citing?

Stent cites a more recent example of research that cannot immediately be connected with today's scientific knowledge. He suggests that the macromolecular theory of memory[27] is a case of prematurity. This theory proposes that memories are stored in the brain in the form of protein or nucleic acid molecules, called "engrams." When memories are recalled, the engrams are somehow retrieved and the information encoded in them is "read" by the brain. Stent points out, "There is no chain of reasonable inferences by which our present, albeit highly imperfect, view of the functional organization of the brain can be reconciled with the possibility of its acquiring, storing, and retrieving nervous information by encoding such information in molecules of nucleic acid or protein."[9]

Stent notes that the macromolecular theory of memory still cannot be connected with current scientific knowledge,[28] 15 years after it was proposed by Babich and colleagues.[27] As a consequence, Stent observes, the theory is mentioned much less frequently today, if it is mentioned at all. His subjective impression is confirmed by citation data taken from the *SCI*: Babich's paper was cited 70 times from 1965-71, but it received only 11 citations from 1972-79. I've discussed Babich's paper in a recent essay on memory techniques.[29]

Is there any way to predict which papers will prove to be premature? Cole believes there is no way to tell in advance whether a paper will prove to be premature.[30] However, it is possible to identify delayed recognition papers from their citation histories. A number of basic problems must be solved before such a retrospective method can be devised. First, we would have to define typical citation patterns for the average research article. Each field will have its own characteristic citation pattern—citation patterns for biochemistry would not be the same for high energy physics or mathematics.

Second, we would have to decide what is a genuine "deviation" from the typical citation pattern. If we know when a typical paper is expected to attain a peak citation rate, we can isolate those papers that peaked many years before or after that point. These deviant papers may peak sooner or later than expected for a number of reasons, and not all deviant papers will be examples of delayed recognition. For example, a paper describing a method might have a different citation pattern from a theoretical paper. Or a paper published in a new subspecialty might have a different pattern from the field as a whole.

Lastly, we would have to set parameters for what truly qualifies as prematurity. How abrupt must a change in citation rate be in order to suggest delayed recognition? If a paper is uncited for ten years after its publication and then is cited ten times in one year—is that delayed recognition? If a paper is cited 20 times per year and then in one year increases to 50, does that signify some form of delayed recognition?

ISI®'s data base is a unique historical resource. We can use it to design a method for identifying premature research. Once we establish the specifications, we can call attention to those papers for which the citation patterns are typical of delayed recognition papers. Cole thinks delayed recognition is due to an inherent conservatism in science. He says scientists tend to stick with their existing paradigms and resist new ideas that don't fit the model.[30] Historians and sociologists can study scientific resistance by looking at delayed recognition papers identified through citation analysis. Whether this can contribute to more rapid advances in science remains to be seen.

* * * * *

My thanks to Alfred Welljams-Dorof for his help in the preparation of this essay.

REFERENCES

1. **Garfield E.** Would Mendel's work have been ignored if the *Science Citation Index*® was available 100 years ago? *Current Contents* 12(47):5-6, 25 November 1969.*
2. **Jansky K G.** Electrical disturbances apparently of extraterrestrial origin. *Proc. IRE* 21:1387-98, 1933.
3. **Zuckerman H.** Theory choice and problem choice in science. *Sociol. Inq.* 48(3):65-96, 1978.
4. **Reber G.** Cosmic static. *Astrophys. J.* 91:621-4, 1940.
5. **Selye H.** General adaptation syndrome. *J. Clin. Endocrinol.* 6:117-230, 1946.
6. **Price D J D.** Telephone communication. 16 April 1980.
7. --------------. A general theory of bibliometrics and other cumulative advantage processes. *J. Amer. Soc. Inform. Sci.* 27:292-306, 1976.
8. **Garfield E.** On style in scientific writing. *Current Contents* (2):5, 10 January 1977.*
9. **Stent G S.** Prematurity and uniqueness in scientific discovery. *Sci. Amer.* 227(6):84-93, December 1972.

10. **Wyatt H V.** Knowledge and prematurity—journey from transformation to DNA. *Perspect. Biol. Med.* 18:596-602, 1961.
11. **Barber B.** Resistance by scientists to scientific discovery. *Science* 134:596-602, 1961.
12. **Cole S.** Professional standing and the reception of scientific discoveries. *Amer. J. Sociol.* 76:286-306, 1970.
13. **Dessler A J.** Nobel prizes: 1970 awards honor three in physics and chemistry. 1. Swedish iconoclast recognized after many years of rejection and obscurity. *Science* 170:604-6, 1970.
14. **Comroe J H.** How to succeed in failing without really trying. *Amer. Rev. Resp. Dis.* 114:629-34, 1976.
15. **Dobzhansky T.** Heredity. *Encyclopaedia Britannica.* Chicago: H.H. Benton, 1976. Vol. 8. p. 801-19.
16. **Glass B.** A century of biochemical genetics. *Proc. Amer. Phil. Soc.* 109:227-36, 1965.
17. **Brannigan A.** The reification of Mendel. *Soc. Stud. Sci.* 9:423-54, 1979.
18. **Simonton D K.** Multiple discovery and invention: zeitgeist, genius or chance? *J. Pers. Soc. Psychol.* 37:1603-16, 1979.
19. **Avery O T, MacLeod C M & McCarty M.** Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* 79:137-58, 1944.
20. **Hotchkiss R D.** The identification of nucleic acids as genetic determinants. (Srinivasan P R, Fruton J S & Edsall J T, eds.) *The origins of modern biochemistry: a retrospect on proteins.* New York: New York Academy of Sciences, 1979. p. 321-42.
21. **Lederberg J.** Reply to H.V. Wyatt. *Nature* 239:234, 1972.
22. **Zuckerman H.** The prize-winning research. *Scientific elite.* New York: Free Press, 1977. p. 163-207.
23. **Ashdown A A.** Nobel prizes to German, Hollander. *Chem. Eng. News* 31:4760, 1953.
24. Nobel prize to Rous, Huggins. *Chem. Eng. News* 44:23, 1966.
25. **Comroe J H.** Lags between initial discovery and clinical application to cardiovascular pulmonary medicine and surgery. Part 1 of Appendix B of *Report of the President's Biomedical Research Panel.* Washington, DC: US Department of Health, Education, and Welfare. DHEW Publ. No. (OS) 76-502.
26. **Garfield E.** Trends in biochemical literature. *Trends Biochem. Sci.* 4:N290-5, 1979.
27. **Babich F R, Jacobson A L, Bubash S & Jacobson A.** Transfer of a response to naive rats by injection of ribonucleic acid extracted from trained rats. *Science* 149:656-7, 1965.
28. **Stent G S.** Telephone communication. 26 February 1980.
29. **Garfield E.** Memory and super-memory—I'll never forget what's his name! *Current Contents* (9):5-9, 3 March 1980.
30. **Cole S.** Telephone communication. 23 April 1980.

*Reprinted in: **Garfield E.** *Essays of an information scientist.* Philadelphia: ISI Press, 1980. 3 vols.