New Technology, New People, New Organizations: The Rise of the MOS Transistor, 1945-1975

Ross Bassett*

Department of History
North Carolina State University

The MOS (Metal-Oxide-Semiconductor) transistor, as the fundamental element in all digital electronics, is the base technology of late twentieth century American society. It has been the vehicle through which digital electronics has entered almost every area of American life, first through the electronic calculator, then through the digital watch, and finally through the microprocessor. And while the proliferation of the microprocessor has been most apparent through the personal computer, it has spread almost everywhere: to automobiles, sewing machines, cameras, and dishwashers, to name only a few of its endless applications. In 1994, over 2 billion 4- and 8-bit microprocessors and microcontrollers (long out of date for use as central processing units for computers) were produced [Slater, 1996, p. 41].

The MOS Transistor and Technological Revolutions

The story of the MOS transistor is the story of a technological revolution. The first technologically important transistor was the bipolar transistor, invented in 1948 by William Shockley at Bell Labs. It was extremely successful, and for most of the first twenty years of its life, the term transistor meant bipolar transistor. Bipolar transistors were in the first transistor radio, the first transistorized televisions, and the first integrated circuit. Today the bipolar transistor has been almost completely supplanted by the MOS transistor, a state of affairs that would have been unimaginable in the early 1960s.

This revolution has had very real winners and losers. Intel, the world's largest integrated circuit maker, made a decision early on in its history that it would abandon the bipolar transistor and cast its lot with the MOS transistor, a decision which was crucial to its later success. The MOS transistor played a

* The dissertation from which this piece was drawn was written at Princeton University in the Department of History (Program in History of Science) under the supervision of Michael S. Mahoney. Other members of the dissertation committee were M. Norton Wise and Stuart Leslie. This project was supported by a National Science Foundation Dissertation Improvement Grant, an SDF Project Fellowship from the Stanford University Libraries, and an IEEE Fellowship in Electrical History.

BUSINESS AND ECONOMIC HISTORY, Volume Twenty-seven, no. 1, Fall 1998.
 ISSN 0894-6825.

central role in IBM's ordeal of the early 1990s, by undercutting the mainframe computer to the point where it could no longer serve as the foundation of the company. Not coincidentally, as part of IBM's retrenchment, it abandoned its bipolar transistor work, deciding it could no longer afford to support both technologies, and thousands of people were laid off when its bipolar facility was shuttered.

This dissertation covers the years between 1945 and 1975, the critical period for the development of this new technology, and concentrates on the MOS efforts of four leaders in semiconductor research and development: Bell Labs, Fairchild Semiconductor, Intel, and IBM. The year 1945 marked the beginnings of Bell Labs' semiconductor research which led to the invention of the transistor, while by 1975 the MOS transistor had established itself as a major presence in the semiconductor industry, with Intel as the leading MOS firm.

The MOS transistor was by and large the work of a group of individuals who had little background in the older bipolar technology and the work on the MOS transistor was done primarily in organizations that had little or no commitment to the older bipolar technology. This would be expected in a case of technological paradigm change. But while the MOS transistor was revolutionary in terms of people and organizations, it was evolutionary in terms of knowledge and equipment. It built on the infrastructure that had grown up around the older bipolar transistor. Every scientist, engineer, technician, and operator working on the bipolar transistor (and in the mid-1960s there were thousands of them) was unwittingly enlisted in the cause of the MOS transistor because an improvement in bipolar technology often translated into an improved MOS transistor. The story of the MOS transistor is one of path dependency; the dominance of the bipolar transistor aided the MOS transistor's cause. However even though it was technically possible to use the exact same manufacturing line to make both MOS and bipolar transistors, the practice was to use separate lines, because almost every process used to make bipolar transistors had to undergo significant, but subtle modifications to make good MOS transistors. Gordon Moore, the head of research and development at Fairchild Semiconductor in the 1960s said "MOS was a religion as well as a technology," and to be a participant, "one was required to practice rites for reasons which bipolar people could not readily accept" [quoted in Bassett, 1998, p. 19].

The technical advantages of the MOS transistor were that it was simpler to build and improve. Although the MOS transistor was initially much slower than the bipolar transistor, it needed only a fraction of the processing steps that the bipolar transistor required, and furthermore the MOS transistor could be improved by scaling – by shrinking all its dimensions by a constant factor. The bipolar transistor did not scale – it could be improved but each generation required increasing ingenuity. The MOS transistor's advantages were particularly compelling for integrated circuits (a contemporaneous invention), for one could put many more MOS transistors than bipolar transistors on an integrated circuit.

Science and Technology, Research and Development

David Hounshell has described the post-World War II ascendancy of Vannevar Bush's linear model that world class science leads to technological breakthroughs [Hounshell, 1996, pp. 41-46]. Hounshell has chronicled how this played out at DuPont, with a greatly increased basic research budget in the hopes of finding "new nylons." IBM, which previously had had almost no scientific research lab, in 1956 established a large central research organization aimed at competing with the best research labs in the world, although it had given very little thought to what it wanted this lab to do or how it would fit in with the rest of the company. In 1958 a senior manager of IBM Research stated "we, in general, undertake no projects just because they are of interest to IBM, but rather select our work so that it will, at once, be good science and of value to the Company" [quoted in Bassett, 1998, p. 114].

The MOS transistor did not follow this model of technology springing from science. And in fact that was one of the things that made it unattractive to researchers, there was no exciting new science associated with it (or so they thought). Instead it was a case of one technology leading to another technology. The MOS transistor represents in a certain sense the triumph of technology over science. Science did not say how to make good MOS transistors. Development was instead pedestrian engineering work, consisting of dozens of small things which made it possible to manufacture the MOS transistor reliably.

Furthermore the MOS transistor story is in a certain sense about the end of research. Throughout the 1950s and 1960s when research managers looked for the next generation of electronic device, their typical practice, in line with Bush's dictum, was to find an interesting scientific effect or a material with interesting scientific properties and then attempt to build a technology around it. Two examples of this are Josephson Junctions, which take advantage of electron tunneling, and gallium arsenide, a material in which electrons move at very high speeds. In spite of the great hopes these programs aroused, they invariably ran into problems which delayed them. In the meantime the relentless advance of the MOS and bipolar technologies which used the existing infrastructure passed by these would-be new technologies and left them uncompetitive. Today the MOS transistor stands unchallenged, with no industrial research lab willing to spend the money to attempt to overthrow it.

Technological Communities

The work of Thomas Kuhn and Edward Constant suggests that the community is a promising focus for the study of scientific or technological change [Kuhn, 1970; Constant, 1980, 1984, 1987]. (The definition of community used here is broader than Constant's and includes organizations within corporations.) This a community study of innovation using the methods of the new social history, analogous to Boyer and Nissenbaum's *Salem Possessed.* It examines the social conditions behind the rise of the MOS transistor, but instead of asking who saw witches and who did not, it asks who saw the MOS transistor and who did not. What people and organizations embraced the MOS transistor and what was their position within the larger corporation?

This focus on technological communities has been productive in several ways. One of the most intriguing things about the MOS transistor is that there is no publication marking its birth, even though it was invented at Bell Laboratories, at a time when Bell Labs published and distributed its most important work throughout the industry. Furthermore the person most clearly associated with its invention, M.M. Atalla, is virtually unknown even in the industry. An examination of the early history of Bell Labs' transistor work makes sense of these anomalies.

By 1955 Bell Labs had made substantial progress in its bipolar transistor development effort and had made the decision that a particular version of that transistor was going to be the workhorse technology in the Bell System. The main task at hand for Bell's transistor development group was to implement versions of this transistor. The next year a new group was formed within the semiconductor area. Its manager was an Egyptian mechanical engineer, M.M. Atalla, who had just transferred over from the electromechanical relay area and had no experience with semiconductors. The department consisted entirely of newly hired engineers and scientists. It was from this department that the MOS transistor emerged

There were two problems with the MOS transistor: it met no business need of the Bell System, and to be at all useful it required a solution to a problem that had vexed semiconductor researchers from before the time of the initial invention of the transistor, the problem of semiconductor surfaces. The group from which the MOS transistor emerged was incompletely socialized into Bell Lab's transistor development effort – anyone with experience working with semiconductors would have known that the MOS transistor was a bad idea. One senior person in the main transistor effort claimed to have been entirely unaware of Atalla's work and found out about it only after he had left Bell Labs to teach at Stanford. Atalla gave a presentation on his work at one of the industry's leading conferences, but it attracted almost no attention. The senior manager of Bell's transistor development effort thought the work was unpromising, and it was halted before a paper could even be published on the subject. It was almost three years later that the MOS transistor came onto the agenda of the semiconductor industry, and this was due to RCA and Fairchild Semiconductor. Bell Labs, which had been the leader in work on the bipolar transistor, would be a laggard in MOS work.

A major theme in the early history of the MOS transistor is the importance of people and organizations at the margins of power within the corporation in advancing the MOS technology. The bipolar transistor had given every evidence of continuing to be successful and people and organizations working on it had no incentive to give it up. In addition to Atalla, the outsiders included a maverick physicist at Fairchild Semiconductor who saw himself as an inventor rather than a scientist and a young engineer who used his positions at the margins of several large companies, Douglas Aircraft, IBM, and Fairchild, to pursue his own vision of what MOS technology should be. The marginal organizations included a Research group at IBM that had failed to make substantial contributions to the company, an IBM development laboratory in Boeblingen West Germany whose possibilities were circumscribed by its position outside the United States, and several small start-ups in the Santa Clara Valley of California.

The early years of the MOS transistor were a period of great uncertainty. Even while people from a number of companies worked on the technology, no one knew how to make MOS transistors reproducibly and no one knew what the MOS transistor would be suitable for. One way this uncertainty was managed was through the exchange of information between competing firms. The elite scientists (nearly always Ph.D.'s in chemistry, physics, or engineering) presented their work at professional conferences that were modeled on academic science. Here researchers were simultaneously members of two communities – their corporation and the community of MOS researchers. Through these conferences a number of major research findings were quickly spread throughout the industry.

There were other areas of interfirm communication as well. Researchers described scientific effects that were important in making MOS transistors, but they did not tell how to make MOS transistors. (That was not their job.) Much was craft. One of the most important means of spreading information between corporations was through the wanderings of one individual, Frank Wanlass. Wanlass had a Ph.D. in physics and was the first to make MOS transistors at Fairchild. But he considered himself an inventor rather than a scientist and did not have the patience for the detailed analytical studies presented at professional conferences. Wanlass rarely published and if one were to judge the significant contributors to MOS technology based on their publications, Wanlass would not be among them.

What Wanlass had was know-how. He had more experience making MOS transistors than anyone else in the industry. And he had a number of skills in making MOS integrated circuits that were not dependent on science. Wanlass was extremely bright, but very independent. He was more interested in advancing the technology than in the fortunes of the company he worked for and was willing to be much more candid in talking to engineers from competitive firms than most people were. I have traced interactions between Wanlass and almost every other major MOS program in the industry.

A focus on communities and the movement of personnel within the industry has been particularly helpful in pointing out the unique status of IBM. By the mid-1960s IBM's semiconductor operations were essentially a closed system with nobody coming in from other firms to take a position of significant responsibility at IBM and nobody leaving IBM. In some ways the only similarities between IBM's semiconductors and those made by other firms were that they all operated according to the same laws of physics. IBM's position as a computer manufacturer allowed its semiconductors to have dramatically different characteristics from merchant semiconductor producers. IBM was a vertically integrated company marketing computing systems – not semiconductors.

Alfred Chandler has described how firms substituted managerial control for market control in internalizing functions [Chandler, 1977, pp. 6-7]. This is what IBM did in its semiconductor operations and as they developed, they were immune to market forces. As long as IBM's semiconductor technology was grossly competitive with what its competitors were using, IBM could do whatever it pleased. A firm's decision as to whether or not it would use IBM data processing equipment might be affected by a number of factors, such as IBM's marketing strength or the quality of its service organization, but the specific semiconductor technology IBM used was not one of them.

IBM's development and manufacturing groups did not closely follow what was happening in the rest of the industry and were very slow to pick up innovations that had originated elsewhere. This is not to say that IBM was not innovative, but they were primarily interested in pursuing their own ideas. While by the early 1970s Intel had developed a wide variety of MOS products, ranging from RAMs, EPROMs, and digital watches to microprocessors, IBM had essentially one MOS product, RAM. Ignoring work done elsewhere was to prove very costly to IBM.

The capstone of the MOS story is the formation and early history of Intel. Intel created the most important early market for MOS technology, semiconductor memories, and by 1975 it was the largest vendor of MOS products. While it is well known that Gordon Moore and Robert Noyce came to Intel from Fairchild, Intel had a complex relation to its Fairchild past. Although Fairchild provided a large proportion of the early Intel engineers and managers (including an entire manufacturing group) Intel made a conscious effort to transcend the organizational problems that had plagued Fairchild. Specifically Intel had no separate R&D group and did not do the kind of research often leading to published papers that Fairchild had. In contrast to the previously quoted IBM manager, Intel did not do "good science" by academic standards; it did enough to solve its problems and moved on [Moore, 1996, pp. 168-69].

Perhaps the best known event in this study is the invention of the microprocessor. It is the best known because Intel has a large stake in promoting that history. But things are more complicated than Intel lets on. As one looks at the organizations involved in MOS work in the late 1960s, one finds that the idea that it would be possible to put a computer on a chip was widespread at the time. Furthermore in the particular case of Fairchild, there were two distinct groups working on MOS. One group had a background in physics and chemistry and was skilled in the processes of making integrated circuits. The other was made up of electrical engineers who understood how to design things with MOS integrated circuits. The core of the first group left Fairchild to form Intel, while the second left to form a computer company, Four-Phase Systems. Four-Phase actually built a chip that could have been called a microprocessor prior to Intel's work. It did not see what it had done as a microprocessor, a computer on a chip or even a discrete invention. To the engineers at Four-Phase their chip represented merely an evolutionary extension of previous ideas. Because they were in a computer company, they did not want to publicize their chip for fear others would copy it. Although the chip was technically similar to Intel's early microprocessors (and in many ways superior), its location in a computer company led to a vastly different trajectory for it.

This is a dissertation which covers a lot of technical material. But very little is merely technical. Different social groups shape technology in different ways. Historians have shown this in looking at labor and management and also at producers and consumers [Noble, 1984; Kline and Pinch, 1996]. Engineers in different organizations (even within a single corporation) can come to different conceptions of a technology and these different conceptions have real implications.

References

Bassett, Ross K., *New Technology, New People, New Organizations: The Rise of the MOS Transistor, 1945-1975* (Ph.D. dissertation, Princeton University, 1998).

Boyer, Paul and Stephen Nissenbaum, *Salem Possessed: The Social Origins of Witchcraft* (Cambridge, MA, 1974).

Chandler, Alfred D. Jr., *The Visible Hand: The Managerial Revolution in American Business* (Cambridge, MA, 1977).

Constant, Edward W., II, "Communities and Hierarchies: Structure in the Practice of Science and Technology," Rachel Laudan ed., *The Nature of Scientific Knowledge: Are Models of Scientific Change Relevant?* (Dordrecht, 1984), 27-46.

———, *The Origins of the Turbojet Revolution* (Baltimore, 1980)

———, "The Social Locus of Technological Practice: Community, System, or Organization?" Wiebe E. Bijker, Thomas P. Hughes and Trevor J. Pinch, eds., *The Social Construction of Technological Systems* (Cambridge, MA, 1987), 223-42.

Hounshell, David A., "The Evolution of Industrial Research in the United States," Richard S. Rosenbloom and William J. Spencer, eds., *Engines of Innovation: U.S. Industrial Research at the End of an Era* (Boston, 1996), 13-85.

Kline, Ronald, and Trevor Pinch, "Users as Agents of Technological Change: The Social Construction of the Automobile in the Rural United States," *Technology and Culture,* 37 (1996), 763-95.

Kuhn, Thomas S. *The Structure of Scientific Revolutions,* 2nd edn. (Chicago, 1970).

Moore, Gordon E. "Some Personal Perspectives on Research in the Semiconductor Industry," Richard S. Rosenbloom and William J. Spencer, eds., *Engines of Innovation: U.S. Industrial Research at the End of an Era* (Boston, 1996), 165-74.

Noble, David F. *Forces of Production: A Social History of Industrial Automation* (New York, 1984).

Slater, Michael, "The Microprocessor Today," *IEEE Micro* (December 1996), 32-44.