Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms

Rebecca M. Henderson; Kim B. Clark


Stable URL:
http://links.jstor.org/sici=sici=0001-8392%28199003%2935%3A1%3C9%3AAITROE%3E2.0.CO%3B2-U

Administrative Science Quarterly is currently published by Johnson Graduate School of Management, Cornell University.
Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms

Rebecca M. Henderson
Massachusetts Institute of Technology
Kim B. Clark
Harvard University

This paper demonstrates that the traditional categorization of innovation as either incremental or radical is incomplete and potentially misleading and does not account for the sometimes disastrous effects on industry incumbents of seemingly minor improvements in technological products. We examine such innovations more closely and, distinguishing between the components of a product and the ways they are integrated into the system that is the product “architecture,” define them as innovations that change the architecture of a product without changing its components. We show that architectural innovations destroy the usefulness of the architectural knowledge of established firms, and that since architectural knowledge tends to become embedded in the structure and information-processing procedures of established organizations, this destruction is difficult for firms to recognize and hard to correct. Architectural innovation therefore presents established organizations with subtle challenges that may have significant competitive implications. We illustrate the concept’s explanatory force through an empirical study of the semiconductor photolithographic alignment equipment industry, which has experienced a number of architectural innovations.

The distinction between refining and improving an existing design and introducing a new concept that departs in a significant way from past practice is one of the central notions in the existing literature on technical innovation (Mansfield, 1968; Moch and Morse, 1977; Freeman, 1982). Incremental innovation introduces relatively minor changes to the existing product, exploits the potential of the established design, and often reinforces the dominance of established firms (Nelson and Winter, 1982; Ettlie, Bridges, and O’Keefe, 1984; Dewar and Dutton, 1986; Tushman and Anderson, 1986). Although it draws from no dramatically new science, it often calls for considerable skill and ingenuity and, over time, has very significant economic consequences (Hollander, 1965). Radical innovation, in contrast, is based on a different set of engineering and scientific principles and often opens up whole new markets and potential applications (Dess and Beard, 1984; Ettlie, Bridges, and O’Keefe, 1984; Dewar and Dutton, 1986). Radical innovation often creates great difficulties for established firms (Cooper and Schendel, 1976; Daft, 1982; Rothwell, 1986; Tushman and Anderson, 1986) and can be the basis for the successful entry of new firms or even the redefinition of an industry.

Radical and incremental innovations have such different competitive consequences because they require quite different organizational capabilities. Organizational capabilities are difficult to create and costly to adjust (Nelson and Winter, 1982; Hannan and Freeman, 1984). Incremental innovation reinforces the capabilities of established organizations, while radical innovation forces them to ask a new set of questions, to draw on new technical and commercial skills, and to employ new problem-solving approaches (Burns and Stalker, 1966; Hage, 1980; Ettlie, Bridges, and O’Keefe, 1984; Tushman and Anderson, 1986).
The distinction between radical and incremental innovation has produced important insights, but it is fundamentally incomplete. There is growing evidence that there are numerous technical innovations that involve apparently modest changes to the existing technology but that have quite dramatic competitive consequences (Clark, 1987). The case of Xerox and small copiers and the case of RCA and the American radio receiver market are two examples.

Xerox, the pioneer of plain-paper copiers, was confronted in the mid-1970s with competitors offering copiers that were much smaller and more reliable than the traditional product. The new products required little new scientific or engineering knowledge, but despite the fact that Xerox had invented the core technologies and had enormous experience in the industry, it took the company almost eight years of missteps and false starts to introduce a competitive product into the market. In that time Xerox lost half of its market share and suffered serious financial problems (Clark, 1987).

In the mid-1950s engineers at RCA’s corporate research and development center developed a prototype of a portable, transistorized radio receiver. The new product used technology in which RCA was accomplished (transistors, radio circuits, speakers, tuning devices), but RCA saw little reason to pursue such an apparently inferior technology. In contrast, Sony, a small, relatively new company, used the small transistorized radio to gain entry into the U.S. market. Even after Sony’s success was apparent, RCA remained a follower in the market as Sony introduced successive models with improved sound quality and FM capability. The irony of the situation was not lost on the R&D engineers: for many years Sony’s radios were produced with technology licensed from RCA, yet RCA had great difficulty matching Sony’s product in the marketplace (Clark, 1987).

Existing models that rely on the simple distinction between radical and incremental innovation provide little insight into the reasons why such apparently minor or straightforward innovations should have such consequences. In this paper, we develop and apply a model that grew out of research in the automotive, machine tool, and ceramics industries that helps to explain how minor innovations can have great competitive consequences.

CONCEPTUAL FRAMEWORK

Component and Architectural Knowledge

In this paper, we focus on the problem of product development, taking as the unit of analysis a manufactured product sold to an end user and designed, engineered, and manufactured by a single product-development organization. We define innovations that change the way in which the components of a product are linked together, while leaving the core design concepts (and thus the basic knowledge underlying the components) untouched, as “architectural” innovation. This is the kind of innovation that confronted Xerox and RCA. It destroys the usefulness of a firm’s architectural knowledge but preserves the usefulness of its knowledge about the product’s components.

1 In earlier drafts of this paper we referred to this type of innovation as “generational.” We are indebted to Professor Michael Tushman for his suggestion of the term architectural.
Architectural Innovation

This distinction between the product as a whole—the system—and the product in its parts—the components—has a long history in the design literature (Marple, 1961; Alexander, 1964). For example, a room fan’s major components include the blade, the motor that drives it, the blade guard, the control system, and the mechanical housing. The overall architecture of the product lays out how the components will work together. Taken together, a fan’s architecture and its components create a system for moving air in a room.

A component is defined here as a physically distinct portion of the product that embodies a core design concept (Clark, 1985) and performs a well-defined function. In the fan, a particular motor is a component of the design that delivers power to turn the fan. There are several design concepts one could use to deliver power. The choice of one of them—the decision to use an electric motor, for example, establishes a core concept of the design. The actual component—the electric motor—is then a physical implementation of this design concept.

The distinction between the product as a system and the product as a set of components underscores the idea that successful product development requires two types of knowledge. First, it requires component knowledge, or knowledge about each of the core design concepts and the way in which they are implemented in a particular component. Second, it requires architectural knowledge or knowledge about the ways in which the components are integrated and linked together into a coherent whole. The distinction between architectural and component knowledge, or between the components themselves and the links between them, is a source of insight into the ways in which innovations differ from each other.

Types of Technological Change

The notion that there are different kinds of innovation, with different competitive effects, has been an important theme in the literature on technological innovation since Schumpeter (1942). Following Schumpeter’s emphasis on creative destruction, the literature has characterized different kinds of innovations in terms of their impact on the established capabilities of the firm. This idea is used in Figure 1, which classifies innovations along two dimensions. The horizontal dimension captures an innovation’s impact on components, while the vertical captures its impact on the linkages between components. There are, of course, other ways to characterize different kinds of innovation. But given the focus here on innovation and the development of new products, the framework outlined in Figure 1 is useful because it focuses on the impact of an innovation on the usefulness of the existing architectural and component knowledge of the firm.

Framed in this way, radical and incremental innovation are extreme points along both dimensions. Radical innovation establishes a new dominant design and, hence, a new set of core design concepts embodied in components that are linked together in a new architecture. Incremental innovation refines and extends an established design. Improvement occurs in individual components, but the underlying core design concepts, and the links between them, remain the same.

2 We are indebted to one of the anonymous ASQ reviewers for the suggestion that we use this matrix.
Figure 1. A framework for defining innovation.

<table>
<thead>
<tr>
<th>Linkages between Core Concepts and Components</th>
<th>Core Concepts</th>
<th>Overturned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unchanged</td>
<td>Incremental Innovation</td>
<td>Modular Innovation</td>
</tr>
<tr>
<td>Changed</td>
<td>Architectural Innovation</td>
<td>Radical Innovation</td>
</tr>
</tbody>
</table>

Figure 1 shows two further types of innovation: innovation that changes only the core design concepts of a technology and innovation that changes only the relationships between them. The former is a modular innovation, such as the replacement of analog with digital telephones. To the degree that one can simply replace an analog dialing device with a digital one, it is an innovation that changes a core design concept without changing the product’s architecture. Our concern, however, is with the last type of innovation shown in the matrix: innovation that changes a product’s architecture but leaves the components, and the core design concepts that they embody, unchanged.

The essence of an architectural innovation is the reconfiguration of an established system to link together existing components in a new way. This does not mean that the components themselves are untouched by architectural innovation. Architectural innovation is often triggered by a change in a component—perhaps size or some other subsidiary parameter of its design—that creates new interactions and new linkages with other components in the established product. The important point is that the core design concept behind each component—and the associated scientific and engineering knowledge—remain the same.

We can illustrate the application of this framework with the example of the room air fan. If the established technology is that of large, electrically powered fans, mounted in the ceiling, with the motor hidden from view and insulated to dampen the noise, improvements in blade design or in the power of the motor would be incremental innovations. A move to central air conditioning would be a radical innovation. New components associated with compressors, refrigerants, and their associated controls would add whole new technical disciplines and new interrelationships. For the maker of large, ceiling-mounted room fans, however, the introduction of a portable fan would be an architectural innovation. While the primary components would be largely the same (e.g., blade, motor, control system), the architecture of the product would be quite different. There would be significant changes in the
Architectural Innovation

interactions between components. The smaller size and the co-location of the motor and the blade in the room would focus attention on new types of interaction between the motor size, the blade dimensions, and the amount of air that the fan could circulate, while shrinking the size of the apparatus would probably introduce new interactions between the performance of the blade and the weight of the housing.

The distinctions between radical, incremental, and architectural innovations are matters of degree. The intention here is not to defend the boundaries of a particular definition, particularly since there are several other dimensions on which it may be useful to define radical and incremental innovation. The use of the term architectural innovation is designed to draw attention to innovations that use many existing core design concepts in a new architecture and that therefore have a more significant impact on the relationships between components than on the technologies of the components themselves. The matrix in Figure 1 is designed to suggest that a given innovation may be less radical or more architectural, not to suggest that the world can be neatly divided into four quadrants.

These distinctions are important because they give us insight into why established firms often have a surprising degree of difficulty in adapting to architectural innovation. Incremental innovation tends to reinforce the competitive positions of established firms, since it builds on their core competencies (Abernathy and Clark, 1985) or is “competence enhancing” (Tushman and Anderson, 1986). In the terms of the framework developed here, it builds on the existing architectural and component knowledge of an organization. In contrast, radical innovation creates unmistakable challenges for established firms, since it destroys the usefulness of their existing capabilities. In our terms, it destroys the usefulness of both architectural and component knowledge (Cooper and Schendel, 1976; Daft, 1982; Tushman and Anderson, 1986).

Architectural innovation presents established firms with a more subtle challenge. Much of what the firm knows is useful and needs to be applied in the new product, but some of what it knows is not only not useful but may actually handicap the firm. Recognizing what is useful and what is not, and acquiring and applying new knowledge when necessary, may be quite difficult for an established firm because of the way knowledge—particularly architectural knowledge—is organized and managed.

The Evolution of Component and Architectural Knowledge

Two concepts are important to understanding the ways in which component and architectural knowledge are managed inside an organization. The first is that of a dominant design. Work by Abernathy and Utterback (1978), Rosenberg (1982), Clark (1985), and Sahal (1986) and evidence from studies of several industries show that product technologies do not emerge fully developed at the outset of their commercial lives (Mansfield, 1977). Technical evolution is usually characterized by periods of great experimentation followed by the acceptance of a dominant design. The second concept is that organizations build knowledge and capability around the recurrent tasks that they perform (Cyert and March, 1963; Nelson and
Winter, 1982). Thus one cannot understand the development of an organization’s innovative capability or of its knowledge without understanding the way in which they are shaped by the organization’s experience with an evolving technology.

The emergence of a new technology is usually a period of considerable confusion. There is little agreement about what the major subsystems of the product should be or how they should be put together. There is a great deal of experimentation (Burns and Stalker, 1966; Clark, 1985). For example, in the early days of the automobile industry, cars were built with gasoline, electric, or steam engines, with steering wheels or tillers, and with wooden or metal bodies (Abernathy, 1978).

These periods of experimentation are brought to an end by the emergence of a dominant design (Abernathy and Utterback, 1978; Sahal, 1986). A dominant design is characterized both by a set of core design concepts that correspond to the major functions performed by the product (Marbles, 1961; Alexander, 1964; Clark, 1985) and that are embodied in components and by a product architecture that defines the ways in which these components are integrated (Clark, 1985; Sahal, 1986). It is equivalent to the general acceptance of a particular product architecture and is characteristic of technical evolution in a very wide range of industries (Clark, 1985). A dominant design often emerges in response to the opportunity to obtain economies of scale or to take advantage of externalities (David, 1985; Arthur, 1988). For example, the dominant design for the car encompassed not only the fact that it used a gasoline engine to provide motive force but also that it was connected to the wheels through a transmission and a drive train and was mounted on a frame rather than on the axles. A dominant design incorporates a range of basic choices about the design that are not revisited in every subsequent design. Once the dominant automobile design had been accepted, engineers did not reevaluate the decision to use a gasoline engine each time they developed a new design. Once any dominant design is established, the initial set of components is refined and elaborated, and progress takes the shape of improvements in the components within the framework of a stable architecture.

This evolutionary process has profound implications for the types of knowledge that an organization developing a new product requires, since an organization’s knowledge and its information-processing capabilities are shaped by the nature of the tasks and the competitive environment that it faces (Lawrence and Lorsch, 1967; Galbraith, 1973).

In the early stages of a technology’s history, before the emergence of a dominant design, organizations competing to design successful products experiment with many different technologies. Since success in the market turns on the synthesis of unfamiliar technologies in creative new designs, organizations must actively develop both knowledge about alternate components and knowledge of how these components can be integrated. With the emergence of a dominant design, which signals the general acceptance of a single architecture, firms cease to invest in learning about alternative configurations of the established set of components. New component knowledge becomes more valuable to a firm than
new architectural knowledge because competition between designs revolves around refinements in particular components. Successful organizations therefore switch their limited attention from learning a little about many different possible designs to learning a great deal about the dominant design. Once gasoline-powered cars had emerged as the technology of choice, competitive pressures in the industry strongly encouraged organizations to learn more about gasoline-fired engines. Pursuing refinements in steam- or electric-powered cars became much less attractive. The focus of active problem solving becomes the elaboration and refinement of knowledge about existing components within a framework of stable architectural knowledge (Dosi, 1982; Clark, 1985).

Since in an industry characterized by a dominant design, architectural knowledge is stable, it tends to become embedded in the practices and procedures of the organization. Several authors have noted the importance of various institutional devices like frameworks and routines in completing recurring tasks in an organization (Galbraith, 1973; Nelson and Winter, 1982; Daft and Weick, 1984). The focus in this paper, however, is on the role of communication channels, information filters, and problem-solving strategies in managing architectural knowledge.

Channels, filters, and strategies. An organization’s communication channels, both those that are implicit in its formal organization (A reports to B) and those that are informal (“I always call Fred because he knows about X”), develop around those interactions within the organization that are critical to its task (Galbraith, 1973; Arrow, 1974). These are also the interactions that are critical to effective design. They are the relationships around which the organization builds architectural knowledge. Thus an organization’s communication channels will come to embody its architectural knowledge of the linkages between components that are critical to effective design. For example, as a dominant design for room fans emerges, an effective organization in the industry will organize itself around its conception of the product’s primary components, since these are the key subtasks of the organization’s design problem (Mintzberg, 1979; von Hippel, 1990). The organization may create a fan-blade group, a motor group, and so on. The communication channels that are created between these groups will reflect the organization’s knowledge of the critical interactions between them. The fact that those working on the motor and the fan blade report to the same supervisor and meet weekly is an embodiment of the organization’s architectural knowledge about the relationship between the motor and the fan blade.

The information filters of an organization also embody its architectural knowledge. An organization is constantly barraged with information. As the task that it faces stabilizes and becomes less ambiguous, the organization develops filters that allow it to identify immediately what is most crucial in its information stream (Arrow, 1974; Daft and Weick, 1984). The emergence of a dominant design and its gradual elaboration molds the organization’s filters so that they come to embody parts of its knowledge of the key relationships between the components of the technology. For instance, the relationships between the designers of motors and controllers for a room
fan are likely to change over time as they are able to express the nature of the critical interaction between the motor and the controller in an increasingly precise way that allows them to ignore irrelevant information. The controller designers may discover that they need to know a great deal about the torque and power of the motor but almost nothing about the materials from which it is made. They will create information filters that reflect this knowledge.

As a product evolves, information filters and communication channels develop and help engineers to work efficiently, but the evolution of the product also means that engineers face recurring kinds of problems. Over time, engineers acquire a store of knowledge about solutions to the specific kinds of problems that have arisen in previous projects. When confronted with such a problem, the engineer does not reexamine all possible alternatives but, rather, focuses first on those that he or she has found to be helpful in solving previous problems. In effect, an organization’s problem-solving strategies summarize what it has learned about fruitful ways to solve problems in its immediate environment (March and Simon, 1958; Lyles and Mitroff, 1980; Nelson and Winter, 1982). Designers may use strategies of this sort in solving problems within components, but problem-solving strategies also reflect architectural knowledge, since they are likely to express part of an organization’s knowledge about the component linkages that are crucial to the solution of routine problems. An organization designing fans might learn over time that the most effective way to design a quieter fan is to focus on the interactions between the motor and the housing.

The strategies designers use, their channels for communication, and their information filters emerge in an organization to help it cope with complexity. They are efficient precisely because they do not have to be actively created each time a need for them arises. Further, as they become familiar and effective, using them becomes natural. Like riding a bicycle, using a strategy, working in a channel, or employing a filter does not require detailed analysis and conscious, deliberate execution. Thus the operation of channels, filters, and strategies may become implicit in the organization.

Since architectural knowledge is stable once a dominant design has been accepted, it can be encoded in these forms and thus becomes implicit. Organizations that are actively engaged in incremental innovation, which occurs within the context of stable architectural knowledge, are thus likely to manage much of their architectural knowledge implicitly by embedding it in their communication channels, information filters, and problem-solving strategies. Component knowledge, in contrast, is more likely to be managed explicitly because it is a constant source of incremental innovation.

**Problems Created by Architectural Innovation**

Differences in the way in which architectural and component knowledge are managed within an experienced organization give us insight into why architectural innovation often creates problems for established firms. These problems have two sources. First, established organizations require significant time (and resources) to identify a particular innovation as architectural, since architectural innovation can often initially be
accommodated within old frameworks. Radical innovation tends to be obviously radical—the need for new modes of learning and new skills becomes quickly apparent. But information that might warn the organization that a particular innovation is architectural may be screened out by the information filters and communication channels that embody old architectural knowledge. Since radical innovation changes the core design concepts of the product, it is immediately obvious that knowledge about how the old components interact with each other is obsolete. The introduction of new linkages, however, is much harder to spot. Since the core concepts of the design remain untouched, the organization may mistakenly believe that it understands the new technology. In the case of the fan company, the motor and the fan-blade designers will continue to talk to each other. The fact that they may be talking about the wrong things may only become apparent after there are significant failures or unexpected problems with the design.

The development of the jet aircraft industry provides an example of the impact of unexpected architectural innovation. The jet engine initially appeared to have important but straightforward implications for airframe technology. Established firms in the industry understood that they would need to develop jet engine expertise but failed to understand the ways in which its introduction would change the interactions between the engine and the rest of the plane in complex and subtle ways (Miller and Sawyers, 1968; Gardiner, 1986). This failure was one of the factors that led to Boeing’s rise to leadership in the industry.

This effect is analogous to the tendency of individuals to continue to rely on beliefs about the world that a rational evaluation of new information should lead them to discard (Kahneman, Slovic, and Tversky, 1982). Researchers have commented extensively on the ways in which organizations facing threats may continue to rely on their old frameworks—or in our terms on their old architectural knowledge—and hence misunderstand the nature of a threat. They shoehorn the bad news, or the unexpected new information, back into the patterns with which they are familiar (Lyles and Mitroff, 1980; Dutton and Jackson, 1987; Jackson and Dutton, 1988).

Once an organization has recognized the nature of an architectural innovation, it faces a second major source of problems: the need to build and to apply new architectural knowledge effectively. Simply recognizing that a new technology is architectural in character does not give an established organization the architectural knowledge that it needs. It must first switch to a new mode of learning and then invest time and resources in learning about the new architecture (Louis and Sutton, 1989). It is handicapped in its attempts to do this, both by the difficulty all organizations experience in switching from one mode of learning to another and by the fact that it must build new architectural knowledge in a context in which some of its old architectural knowledge may be relevant.

An established organization setting out to build new architectural knowledge must change its orientation from one of refinement within a stable architecture to one of active search
for new solutions within a constantly changing context. As long as the dominant design remains stable, an organization can segment and specialize its knowledge and rely on standard operating procedures to design and develop products. Architectural innovation, in contrast, places a premium on exploration in design and the assimilation of new knowledge. Many organizations encounter difficulties in their attempts to make this type of transition (Argyris and Schön, 1978; Weick, 1979; Hedberg, 1981; Louis and Sutton, 1989). New entrants, with smaller commitments to older ways of learning about the environment and organizing their knowledge, often find it easier to build the organizational flexibility that abandoning old architectural knowledge and building new requires.

Once an organization has succeeded in reorientating itself, the building of new architectural knowledge still takes time and resources. This learning may be quite subtle and difficult. New entrants to the industry must also build the architectural knowledge necessary to exploit an architectural innovation, but since they have no existing assets, they can optimize their organization and information-processing structures to exploit the potential of a new design. Established firms are faced with an awkward problem. Because their architectural knowledge is embedded in channels, filters, and strategies, the discovery process and the process of creating new information (and rooting out the old) usually takes time. The organization may be tempted to modify the channels, filters, and strategies that already exist rather than to incur the significant fixed costs and considerable organizational friction required to build new sets from scratch (Arrow, 1974). But it may be difficult to identify precisely which filters, channels, and problem-solving strategies need to be modified, and the attempt to build a new product with old (albeit modified) organizational tools can create significant problems.

The problems created by an architectural innovation are evident in the introduction of high-strength-low-alloy (HSLA) steel in automobile bodies in the 1970s. The new materials allowed body panels to be thinner and lighter but opened up a whole new set of interactions that were not contained in existing channels and strategies. One automaker’s body-engineering group, using traditional methods, designed an HSLA hood for the engine compartment. The hoods, however, resonated and oscillated with engine vibrations during testing. On further investigation, it became apparent that the traditional methods for designing hoods worked just fine with traditional materials, although no one knew quite why. The knowledge embedded in established problem-solving strategies and communication channels was sufficient to achieve effective designs with established materials, but the new material created new interactions and required the engineers to build new knowledge about them.

Architectural innovation may thus have very significant competitive implications. Established organizations may invest heavily in the new innovation, interpreting it as an incremental extension of the existing technology or underestimating its impact on their embedded architectural knowledge. But new entrants to the industry may exploit its potential much more effectively, since they are not handicapped by a legacy of embedded and partially irrelevant architectural knowledge.
Architectural innovation

We explore the validity of our framework through a brief summary of the competitive and technical history of the semiconductor photolithographic alignment equipment industry. Photolithographic aligners are sophisticated pieces of capital equipment used in the manufacture of integrated circuits. Their performance has improved dramatically over the last twenty-five years, and although the core technologies have changed only marginally since the technique was first invented, the industry has been characterized by great turbulence. Changes in market leadership have been frequent, the entry of new firms has occurred throughout the industry's history, and incumbents have often suffered sharp declines in market share following the introduction of equipment incorporating seemingly minor innovation. We believe that these events are explained by the intrusion of architectural innovation into the industry, and we use three episodes in the industry's history—particularly Canon's introduction of the proximity aligner and Kasper's response to it—to illustrate this idea in detail.

INNOVATION IN PHOTOLITHOGRAPHIC ALIGNMENT EQUIPMENT

Data

The data were collected during a two-year, field-based study of the photolithographic alignment equipment industry. The study was initially designed to serve as an exploration of the validity of the concept of architectural innovation, a concept originally developed by one of the authors during the course of his experience with the automobile and ceramics industry (Clark, 1987).

The core of the data is a panel data set consisting of research and development costs and sales revenue by product for every product development project conducted between 1962, when work on the first commercial product began, and 1986. This data is supplemented by a detailed managerial and technical history of each project. The data were collected through research in both primary and secondary sources. The secondary sources, including trade journals, scientific journals, and consulting reports, were used to identify the companies that had been active in the industry and the products that they had introduced and to build up a preliminary picture of the industry's technical history.

Data were then collected about each product-development project by contacting directly at least one of the members of the product-development team and requesting an interview. Interviews were conducted over a fourteen-month period, from March 1987 to May 1988. During the course of the research, over a hundred people were interviewed. As far as possible, the interviewees included the senior design engineer for each project and a senior marketing executive from each firm. Other industry observers and participants, including chief executives, university scientists, skilled design engineers, and service managers were also interviewed. Interview data were supplemented whenever possible through the use of internal firm records. The majority of the interviews were semistructured and lasted about two hours. Respondents were asked to describe the technical, commercial, and
managerial history of the product-development projects with which they were familiar and to discuss the technical and commercial success of the products that grew out of them.

In order to validate the data that were collected during this process, a brief history of product development for each equipment vendor was circulated to all the individuals who had been interviewed and to others who knew a firm’s history well, and the accuracy of this account was discussed over the telephone in supplementary interviews. The same validation procedure was followed in the construction of the technical history of the industry. A technical history was constructed using interview data, published product literature, and the scientific press. This history was circulated to key individuals who had a detailed knowledge of the technical history of the industry, who corrected it as appropriate.

We chose to study the semiconductor photolithographic alignment equipment industry for two reasons. The first is that it is very different from the industries in which our framework was first formulated, since it is characterized by much smaller firms and a much faster rate of technological innovation. The second is that it provides several examples of the impact of architectural innovation on the competitive position of established firms. Photolithographic equipment has been shaken by four waves of architectural innovation, each of which resulted in a new entrant capturing the leadership of the industry. In order to ground the discussion of architectural innovation we provide a brief description of photolithographic technology.

The Technology

Photolithographic aligners are used to manufacture solid-state semiconductor devices. The production of semiconductors requires the transfer of small, intricate patterns to the surface of a wafer of semiconductor material such as silicon, and this process of transfer is known as lithography. The surface of the wafer is coated with a light-sensitive chemical, or “resist.” The pattern that is to be transferred to the wafer surface is drawn onto a mask and the mask is used to block light as it falls onto the resist, so that only those portions of the resist defined by the mask are exposed to light. The light chemically transforms the resist so that it can be stripped away. The resulting pattern is then used as the basis for either the deposition of material onto the wafer surface or for the etching of the existing material on the surface of the wafer. The process may be repeated as many as twenty times during the manufacture of a semiconductor device, and each layer must be located precisely with respect to the previous layer (Watts and Einspruch, 1987). Figure 2 gives a very simplified representation of this complex process.

A photolithographic aligner is used to position the mask relative to the wafer, to hold the two in place during exposure, and to expose the resist. Figure 3 shows a schematic diagram of a contact aligner, the first generation of alignment equipment developed. Improvement in alignment technology has meant improvement in minimum feature size, the size of the smallest pattern that can be produced on the wafer surface, yield, the percentage of wafers successfully processed, and
Architectural Innovation

Figure 2. Schematic representation of the lithographic process.

STEPS
1. Expose Resist

   Mask
   ---
   Resist
   ---
   Wafer
   Light

2. Develop Resist

   Resist
   ---
   Wafer

3. Deposit Material

   Resist
   ---
   Wafer
   Material
   ---

4. Remove Remaining Resist

   Pattern formed on wafer
   ---

throughput, the number of wafers the aligner can handle in a given time.

Contact aligners were the first photolithographic aligners to be used commercially. They use the mask’s shadow to transfer the mask pattern to the wafer surface. The mask and the wafer are held in contact with each other, and light shining through the gaps in the mask falls onto the wafer surface. Contact aligners are simple and quick to use, but the need to bring the mask and the wafer into direct contact can damage the mask or contaminate the wafer. The first proximity aligner was introduced in 1973 to solve these problems.
In a proximity aligner the mask is held a small distance away from (in proximity to) the wafer surface, as shown in the simplified drawing in Figure 4. The separation of the mask and the wafer means that they are less likely to be damaged during exposure, but since the mask and wafer are separated from each other, light coming through the mask spreads out before it reaches the resist, and the mask’s shadow is less well defined than it is in the case of a contact aligner. As a result, users switching to proximity aligners traded off some minimum feature size capability for increased yield.

The basic set of core design concepts that underlie optical photolithography—the use of a visible light source to transmit the image of the mask to the wafer, a lens or other device to focus the image of the mask on the wafer, an alignment system that uses visible light, and a mechanical system that holds the mask and the wafer in place—have remained unchanged since the technology was first developed, although aligner performance has improved dramatically. The minimum-feature-size capability of the first aligners was about fifteen to twenty microns. Modern aligners are sometimes specified to have minimum feature sizes of less than half a micron.
Architectural Innovation

Radical alternatives, making use of quite different core concepts, have been explored in the laboratory but have yet to be widely introduced into full-scale production. Aligners using x-rays and ion beams as sources have been developed, as have direct-write electron beam aligners, in which a focused beam of electrons is used to write directly on the wafer (Chang et al., 1977; Brown, Venkatesan, and Wagner, 1981; Burggraaf, 1983). These technologies are clearly radical. They rely not only on quite different core concepts for the source, but they also use quite different mask, alignment, and lens technologies.

A constant stream of incremental innovation has been critical to optical photolithography’s continuing success. The technology of each component has been significantly improved. Modern light sources are significantly more powerful and more uniform, and modern alignment systems are much more accurate. In addition, the technology has seen four waves of architectural innovation: the move from contact to proximity alignment, from proximity to scanning projection alignment, and from scanners to first- and then second-generation "steppers." Table 1 summarizes the changes in the technology introduced by each generation. In each case the core technologies of optical lithography remained largely untouched, and much of the technical knowledge gained in building a previous generation could be transferred to the next. Yet, in each case, the industry leader was unable to make the transition.

Table 1

<table>
<thead>
<tr>
<th>Equipment</th>
<th>Major Changes</th>
<th>Critical relationships between components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximity aligner</td>
<td>Mask and wafer separated during exposure.</td>
<td>Accuracy and stability of gap is a function of links between gap-setting mechanism and other components.</td>
</tr>
<tr>
<td>Scanning projection stepper</td>
<td>Image of mask projected onto wafer by scanning reflective optics.</td>
<td>Interactions between lens and other components is critical to successful performance.</td>
</tr>
<tr>
<td>First-generation stepper</td>
<td>Image of mask projected through refractive lens. Image &quot;stepped&quot; across wafer.</td>
<td>Relationship between lens field size and source energy becomes significant determinant of throughput. Depth of focus characteristics—driven by relationship between source wavelength and lens numerical aperture—become critical. Interactions between stage and alignment system are critical.</td>
</tr>
<tr>
<td>Second-generation stepper</td>
<td>Introduction of &quot;site-by-site&quot; alignment, larger 5× lenses.</td>
<td>Throughput now driven by calibration and stepper stability. Relationship between lens and mechanical system becomes crucial means of controlling distortion.</td>
</tr>
</tbody>
</table>

Source: Field interviews, internal firm records (Henderson, 1988).

Table 2 shows share of deflated cumulative sales, 1962–1986, by generation of equipment for the leading firms. The first commercially successful aligner was introduced by Kulicke and Soffa in 1965. They were extremely successful and held nearly 100 percent of the (very small) market for the next nine years, but by 1974 Cobilt and Kasper had replaced them. In 1974 Perkin-Elmer entered the market with the
Table 2

<table>
<thead>
<tr>
<th>Firm</th>
<th>Contact</th>
<th>Proximity</th>
<th>Scanners</th>
<th>Step and repeat (1)</th>
<th>Step and repeat (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cobilt</td>
<td>44</td>
<td>&lt;1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kasper</td>
<td>17</td>
<td>8</td>
<td></td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Canon</td>
<td>67</td>
<td>21</td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Perkin-Elmer</td>
<td>78</td>
<td>10</td>
<td></td>
<td>&lt;1</td>
<td></td>
</tr>
<tr>
<td>GCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nikon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>61</td>
<td>75</td>
<td>99+</td>
<td>81</td>
<td>82+</td>
</tr>
</tbody>
</table>

* This measure is distorted by the fact that all of these products are still being sold. For second-generation step and repeat aligners this problem is particularly severe, since in 1986 this equipment was still in the early stages of its life cycle. Source: Internal firm records, Dataquest, VLSI Research Inc.

scanning projection aligner and rapidly became the largest firm in the industry. GCA, in turn, replaced Perkin-Elmer through its introduction of the stepper, only to be supplanted by Nikon, which introduced the second-generation stepper.

In nearly every case, the established firm invested heavily in the next generation of equipment, only to meet with very little success. Our analysis of the industry's history suggests that a reliance on architectural knowledge derived from experience with the previous generation blinded the incumbent firms to critical aspects of the new technology. They thus underestimated its potential or built equipment that was markedly inferior to the equipment introduced by entrants.

The Kasper Saga

The case of Kasper Instruments and its response to Canon's introduction of the proximity printer illustrates some of the problems encountered by established firms. Kasper Instruments was founded in 1968 and by 1973 was a small but profitable firm supplying approximately half of the market for contact aligners. In 1973 Kasper introduced the first contact aligner to be equipped with proximity capability. Although nearly half of all the aligners that the firm sold from 1974 onward had this capability, Kasper aligners were only rarely used in proximity mode, and sales declined steadily until the company left the industry in 1981. The widespread use of proximity aligners only occurred with the introduction and general adoption of Canon's proximity aligner in the late 1970s.

The introduction of the proximity aligner is clearly not a radical advance. The conceptual change involved was minor, and most proximity aligners can also be used as contact aligners. However, in a proximity aligner, a quite different set of relationships between components is critical to successful performance. The introduction of the proximity aligner was thus an architectural innovation. In particular, in a proximity aligner, the relationships between the gap-setting mechanism and the other components of the aligner are significantly different.
In both contact and proximity aligners, the mask and the wafer surface must be parallel to each other during exposure if the quality of the final image on the wafer is to be adequate. This is relatively straightforward in a contact aligner, since the mask and the wafer are in direct contact with each other during exposure. The gap-setting mechanism is used only to separate the mask and the wafer during alignment. Its stability and accuracy have very little impact on the aligner’s performance. In a proximity aligner, however, the accuracy and precision of the gap-setting mechanism are critical to the aligner’s performance. The gap between the mask and the wafer must be precise and consistent across the mask and wafer surfaces if the aligner is to perform well. Thus, the gap-setting mechanism must locate the mask at exactly the right point above the wafer by dead reckoning and must then ensure that the mask is held exactly parallel to the wafer. Since the accuracy and stability of the mechanism is as much a function of the way in which it is integrated with the other components as it is of its own design, the relationships between the gap-setting mechanism and the other components of the aligner must change if the aligner is to perform well. Thus, the successful design of a proximity aligner requires both the acquisition of some new component knowledge—how to build a more accurate and more stable gap-setting mechanism—and the acquisition of new architectural knowledge.

Kasper’s failure to understand the challenge posed by the proximity aligner is especially puzzling given its established position in the market and its depth of experience in photolithography. There were several highly skilled and imaginative designers at Kasper during the early 1970s. The group designed a steady stream of contact aligners, each incorporating significant incremental improvements. From 1968 to 1973, the minimum-feature-size capability of its contact aligners improved from fifteen to five microns.

But Kasper’s very success in designing contact aligners was a major contributor to its inability to design a proximity aligner that could perform as successfully as Canon’s. Canon’s aligner was superficially very similar to Kasper’s. It incorporated the same components and performed the same functions, but it performed them much more effectively because it incorporated a much more sophisticated understanding of the technical interrelationships that are fundamental to successful proximity alignment. Kasper failed to develop the particular component knowledge that would have enabled it to match Canon’s design. More importantly, the architectural knowledge that Kasper had developed through its experience with the contact aligner had the effect of focusing its attention away from the new problems whose solution was critical to the design of a successful proximity aligner.

Kasper conceived of the proximity aligner as a modified contact aligner. Like the incremental improvements to the contact aligner before it, design of the proximity aligner was managed as a routine extension to the product line. The gap-setting mechanism that was used in the contact aligner to align the mask and wafer with each other was slightly modified, and the new aligner was offered on the market. As a result, Kasper’s proximity aligner did not perform well. The
gap-setting mechanism was not sufficiently accurate or stable to ensure adequate performance, and the aligner was rarely used in its proximity mode. Kasper’s failure to understand the obsolescence of its architectural knowledge is demonstrated graphically by two incidents.

The first is the firm’s interpretation of early complaints about the accuracy of its gap-setting mechanism. In proximity alignment, misalignment of the mask and the wafer can be caused both by inaccuracies or instability in the gap-setting mechanism and by distortions introduced during processing. Kasper attributed many of the problems that users of its proximity equipment were experiencing to processing error, since it believed that processing error had been the primary source of problems with its contact aligner. The firm “knew” that its gap-setting mechanism was entirely adequate, and, as a result, devoted very little time to improving its performance. In retrospect, this may seem like a wanton misuse of information, but it represented no more than a continued reliance on an information filter that had served the firm well historically.

The second illustration is provided by Kasper’s response to Canon’s initial introduction of a proximity aligner. The Canon aligner was evaluated by a team at Kasper and pronounced to be a copy of a Kasper machine. Kasper evaluated it against the criteria that it used for evaluating its own aligners—criteria that had been developed during its experience with contact aligners. The technical features that made Canon’s aligner a significant advance, particularly the redesigned gap mechanism, were not observed because they were not considered important. The Canon aligner was pronounced to be “merely a copy” of the Kasper aligner.

Kasper’s subsequent commercial failure was triggered by several factors. The company had problems designing an automatic alignment system of sufficient accuracy and in managing a high-volume manufacturing facility. It also suffered through several rapid changes of top management during the late 1970s. But the obsolescence of architectural knowledge brought about by the introduction of architectural innovation was a critical factor in its decline.

Kasper’s failure stemmed primarily from failures of recognition: the knowledge that it had developed through its experience with the contact aligner made it difficult for the company to understand the ways in which Canon’s proximity aligner was superior to its own. Similar problems with recognition show up in all four episodes of architectural innovation in the industry’s history. The case of Perkin-Elmer and stepper technology is a case in point. By the late 1970s Perkin-Elmer had achieved market leadership with its scanning projection aligners, but the company failed to maintain that leadership when stepper technology came to dominate the industry in the early 1980s. When evaluating the two technologies, Perkin-Elmer engineers accurately forecast the progress of individual components in the two systems but failed to see how new interactions in component development—including better resist systems and improvements in lens design—would give stepper technology a decisive advantage.

26/ASQ, March 1990
Architectural Innovation

GCA, the company that took leadership from Perkin-Elmer, was itself supplanted by Nikon, which introduced a second-generation stepper. Part of the problem for GCA was recognition, but much of its failure to master the new stepper technology lay in problems in implementation. Echoing Kasper, GCA first pronounced the Nikon stepper a “copy” of the GCA design. Even after GCA had fully recognized the threat posed by the second-generation stepper, its historical experience handicapped the company in its attempts to develop a competitive machine. GCA’s engineers were organized by component, and cross-department communication channels were all structured around the architecture of the first-generation system. While GCA engineers were able to push the limits of the component technology, they had great difficulty understanding what Nikon had done to achieve its superior performance.

Nikon had changed aspects of the design—particularly the ways in which the optical system was integrated with the rest of the aligner—of which GCA’s engineers had only limited understanding. Moreover, because these changes dealt with component interactions, there were few engineers responsible for developing this understanding. As a result, GCA’s second-generation machines did not deliver the kind of performance that the market demanded. Like Kasper and Perkin-Elmer before them, GCA’s sales languished and they lost market leadership. In all three cases, other factors also played a role in the firm’s dramatic loss of market share, but a failure to respond effectively to architectural innovation was of critical importance.

DISCUSSION AND CONCLUSIONS

We have assumed that organizations are boundedly rational and, hence, that their knowledge and information-processing structure come to mirror the internal structure of the product they are designing. This is clearly an approximation. It would be interesting to explore the ways in which the formulation of architectural and component knowledge are affected by factors such as the firm’s history and culture. Similarly, we have assumed that architectural knowledge embedded in routines and channels becomes inert and hard to change. Future research designed to investigate information filters, problem-solving strategies and communication channels in more detail could explore the extent to which this can be avoided.

The ideas developed here could also be linked to those of authors such as Abernathy and Clark (1985), who have drawn a distinction between innovation that challenges the technical capabilities of an organization and innovation that challenges the organization’s knowledge of the market and of customer needs. Research could also examine the extent to which these insights are applicable to problems of process innovation and process development.

The empirical side of this paper could also be developed. While the idea of architectural innovation provides intriguing insights into the evolution of semiconductor photolithographic alignment equipment, further research could explore the
extent to which it is a useful tool for understanding the impact of innovation in other industries.

The concept of architectural innovation and the related concepts of component and architectural knowledge have a number of important implications. These ideas not only give us a richer characterization of different types of innovation, but they open up new areas in understanding the connections between innovation and organizational capability. The paper suggests, for example, that we need to deepen our understanding of the traditional distinction between innovation that enhances and innovation that destroys competence within the firm, since the essence of architectural innovation is that it both enhances and destroys competence, often in subtle ways.

An architectural innovation’s effect depends in a direct way on the nature of organizational learning. This paper not only underscores the role of organizational learning in innovation but suggests a new perspective on the problem. Given the evolutionary character of development and the prevalence of dominant designs, there appears to be a tendency for active learning among engineers to focus on improvements in performance within a stable product architecture. In this context, learning means learning about components and the core concepts that underlie them. Given the way knowledge tends to be organized within the firm, learning about changes in the architecture of the product is unlikely to occur naturally. Learning about changes in architecture—about new interactions across components (and often across functional boundaries)—may therefore require explicit management and attention. But it may also be that learning about new architectures requires a different kind of organization and people with different skills. An organization that is structured to learn quickly and effectively about new component technology may be ineffective in learning about changes in product architecture. What drives effective learning about new architectures and how learning about components may be related to it are issues worth much further research.

These ideas also provide an intriguing perspective from which to understand the current fashion for cross-functional teams and more open organizational environments. These mechanisms may be responses to a perception of the danger of allowing architectural knowledge to become embedded within tacit or informal linkages.

To the degree that other tasks performed by organizations can also be described as a series of interlinked components within a relatively stable framework, the idea of architectural innovation yields insights into problems that reach beyond product development and design. To the degree that manufacturing, marketing, and finance rely on communication channels, information filters, and problem-solving strategies to integrate their work together, architectural innovation at the firm level may also be a significant issue.

Finally, an understanding of architectural innovation would be useful to discussions of the effect of technology on competitive strategy. Since architectural innovation has the potential to offer firms the opportunity to gain significant advantage over well-entrenched, dominant firms, we might expect less
Architectural Innovation
entrenched competitor firms to search actively for opportunities to introduce changes in product architecture in an industry. The evidence developed here and in other studies suggests that architectural innovation is quite prevalent. As an interpretive lens, architectural innovation may therefore prove quite useful in understanding technically based rivalry in a variety of industries.

REFERENCES

Abernathy, William J.

Abernathy, William J., and Kim Clark

Abernathy, William J., and James Utterback

Alexander, Christopher

Argyris, Chris, and Donald Schöon

Arrow, Kenneth

Arthur, Brian

Brown, William L., T. Venkatesan, and A. Wagner

Burggraaf, Pieter

Burns, Tom, and George Stalker

Chang, T. H. P., M. Hatzakis, A. D. Wilson, and A. N. Broers

Clark, Kim B.

Cooper, Arnold C., and Dan Schendel

Cyert, Richard M., and James G. March

Daft, Richard L.

Daft, Richard L., and Karl E. Weick

David, Paul A.

Dess, Gregory G., and Donald Beard

Dewar, Robert D., and Jane E. Dutton

Dosi, Giovanni

Dutton, Jane E., and Susan E. Jackson

Ettlie, John E., William P. Bridges, and Robert D. O’Keefe

Freeman, Christopher
1982 The Economics of Industrial Innovation, 2nd ed. Cambridge, MA: MIT Press.

Galbraith, Jay

Gardiner, J. P.

Hage, Jerald

Hannan, Michael T., and John Freeman

Hedberg, Bo L. T.

29/ASQ, March 1990
Henderson, Rebecca M. 1988 "The failure of established firms in the face of technical change: A study of photolithographic alignment equipment."


Sahal, Devendra 1986 "Technological guideposts and innovation avenues." Research Policy, 14: 61–82.


