"GENERATIONAL" INNOVATION:
THE RECONFIGURATION OF EXISTING SYSTEMS
AND THE FAILURE OF ESTABLISHED FIRMS

Rebecca M. Henderson and Kim B. Clark

WP 3027-89-BPS May, 1989
"GENERATIONAL" INNOVATION:
THE RECONFIGURATION OF EXISTING SYSTEMS
AND THE FAILURE OF ESTABLISHED FIRMS

Rebecca M. Henderson and Kim B. Clark

WP 3027-89-BPS

May, 1989

Rebecca M. Henderson
Assistant Professor of Management
Massachusetts Institute of Technology

Kim B. Clark
Professor of Management
Harvard University

This research was supported by the Division of Research, Harvard Business School. Their support is gratefully acknowledged. We would also like to thank Dataquest and VLSI Research Inc. for generous permission to use their published data, the staffs at Canon, CCA, Nikon, Perkin Elmer and Ultratech, and all those individuals involved with photolithographic alignment technology who gave so generously of their time.
"GENERATIONAL" INNOVATION:
THE RECONFIGURATION OF EXISTING SYSTEMS
AND THE FAILURE OF ESTABLISHED FIRMS

ABSTRACT

The traditional categorization of innovation as either "incremental" or "radical" is incomplete and fundamentally misleading. "Generational" innovation — innovation that reconfigures a technical system without changing its elements — is qualitatively different from both incremental and radical innovation and often has important and unexpected organizational and competitive consequences.

This paper defines generational innovation and illustrates the concept's explanatory force through an empirical study of the technical and competitive history of the semiconductor photolithographic alignment equipment industry.
INTRODUCTION

"We need to know more about differences between competence destroying and competence-enhancing technological advances... (about) what distinguishes between incremental improvements and dramatic advances." (Tushman and Anderson (1986))

The power of new technology to challenge and transform organizations and the structure of industries has been an important theme of research since Schumpeter. "Radical" innovation creates great difficulties for established firms, (Bowkes, Sawers and Stillerman, 1969; Cooper and Schendel, 1976; Rothwell, 1986; Tushman and Anderson, 1986) and is often the basis for successful entry and even redefinition of the industry, while "incremental" innovation often reinforces the dominance of established firms. Radical and incremental innovations have different competitive consequences because they require quite different organizational capabilities. Incremental innovation feeds on and reinforces the existing problem solving capabilities of established organizations, while radical innovation forces them to ask a new set of questions and to employ new problem solving approaches (Burns and Stalker, 1966; Arrow, 1974; Hage, 1980; Dess and Beard, 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). Take, for example, the case of Xerox and small copiers.

Xerox, the pioneer of plain paper copiers, was confronted in the mid-1970's with competitors offering copiers that were much smaller and more reliable than the traditional product. The new products required no new scientific or
engineering knowledge, but despite the fact that the company had invented the core technologies and had enormous experience in the industry, it took Xerox almost eight years to bring a competitive product into the market. In that time Xerox lost half of its market share and suffered serious financial problems (Clark, 1987). Existing models that rely on distinctions between radical and incremental innovation provide little insight into the reasons these types of innovation should have such consequences. In this paper we develop and apply such a model.

The paper has four parts. We first present a framework that allows us to analyze the nature of technical change and to develop the links between changes in technology and changes in the information processing task that they present to an organization (section 2). We define incremental and radical change within this framework and characterize an important class of innovations as "generational." We argue that generational innovation is qualitatively different from radical and incremental innovation. While incremental innovation changes the performance of an existing set of components, and radical innovation changes the elements of that set, generational innovation changes the relationships between them. While radical innovation challenges the established organization with a wholly different system, the challenge of generational innovation is to learn about a differently configured version of an established system. Knowledge gained through experience with incremental innovation about particular component technologies often remains relevant, but "relational knowledge," or knowledge about the interactions between components, system performance and user needs is rendered obsolete.

In section 3 we briefly summarize the competitive and technical history of the photolithographic alignment equipment industry. The industry offers a useful context in which to illustrate the model and to examine its power in explaining the competitive and organizational consequences of generational innovation. Photolithographic aligners are sophisticated pieces of capital equipment used in the manufacture of integrated circuits. Aligner performance has improved
dramatically over the last twenty-five years, but the core technologies have
changed only marginally since the technique was first invented. Yet the industry
has been characterized by great turbulence. Changes in market leadership have been
frequent, entry has occurred throughout the industry's history, and incumbents
have often suffered sharp declines in market share following the introduction of a
new generation of equipment.

We believe that these events are explained by the intrusion of generational
innovation into the industry. While generational innovation required little new
scientific knowledge, it required significant shifts in relational knowledge.
Established firms in the industry had great difficulty recognizing and taking
advantage of these shifts, and as a result their development efforts were
significantly less effective than those of entrants.

We close the paper with a brief summary and with a discussion of the
implications of this research for understanding organizational responses to
technological change.

II

Conceptual Framework

The central notion in the existing literature on technical innovation is the
distinction between refining and improving an existing design, and introducing a
new concept that departs in a significant way from past practice. (Mansfield,
1968; Moch and Morse, 1977; Freeman, 1982;) "Incremental" innovation introduces
relatively minor changes to the existing product, exploiting the potential of the
established design (Etting, Bridges and O'Keefe, 1984; Dewar and Dutton, 1986;
Tushman and Anderson, 1986). Nelson and Winter, for example, have argued that
incremental innovation occurs along a series of "natural trajectories" during
which firms search "locally" for appropriate solutions (Nelson and Winter, 1982).
Building on Kuhn's work in the history of science, (Kuhn, 1970), Dosi (1982) has
suggested that incremental innovation occurs within a "technological paradigm", or
an established "...outlook, set of procedures, definition of the relevant problems
and of the specific knowledge related to their solutions."

"Radical" innovation, in contrast, is based upon a different set of
engineering and scientific principles and often opens up whole new markets and
potential applications (Ettlie, Bridges and O'Keefe, 1984; Dewar and Dutton, 1986;
Tushman and Anderson, 1986). Thus the commercial and technical environment
associated with radical innovation is far more dynamic and uncertain than that
associated with incremental innovation (Dess and Beard, 1984).

Innovation's impact on competition depends on its effects on both the supply
and demand sides of the market. On the supply side, the two patterns of innovation
are associated with very different organizational capabilities. Radical and
incremental innovation require not only a difference in knowledge base, but also
in organizational structure and process. In general it appears that more formal,
"hierarchical" organizations that rely on structured patterns of communication,
problem solving routines and procedures are best equipped to undertake incremental
innovation, while more "organic," "entrepreneurial" organizations in which the
information flow is relatively unstructured and which are more responsive to
dramatic change are best equipped to exploit radical innovation (Burns and
Stalker, 1966; Moch and Morse, 1977; Hage, 1980; Kimberly and Evanisko, 1981;
Ettlie, 1983; Ettlie, Bridges and O'Keefe, 1984; Dewar and Dutton, 1986).

It is this connection between innovation and organizational capability that
gives radical innovation its ability to transform the competitive environment.
Organizational capabilities are difficult to create, and costly to adjust (Nelson
and Winter, 1982; Hannan and Freeman, 1984). An established firm faced with
radical innovation must make very different decisions and must make them very
differently. The number and magnitude of the changes that radical innovation
requires in the firm's existing stock of knowledge and in the way in which it
processes information may be daunting. Thus several studies have shown that radical innovations tend to originate outside of established firms, and that established firms often have great difficulty in making the shift to a new "trajectory" (Jewkes, Sawers and Stillerman, 1969; Cooper and Schendel, 1976; Rothwell, 1986; Tushman and Anderson, 1986).

Differences in organizational capability help to explain changes in the "suppliers" of innovation. But competition is also shaped by an innovation's effects on demand. The cumulative economic effects of incremental innovation are often significant, and incremental innovation often broadens a product's appeal and expands its market. Radical innovation, in contrast, offers customers a wholly new set of possibilities, or meets established needs in a wholly different way. Although the new concept may bear some relationship to the established design, the change in price/performance is usually so dramatic, and the new needs it meets are so different, that the old product becomes at best a very poor substitute for the new design. Indeed, in the economics literature, radical innovation is defined as innovation that creates such changes in performance or cost that the old product cannot act as a substitute for the new even at competitive prices (Gilbert and Newbery, 1982; Reinganaum, 1983).

As this brief summary suggests, research on innovation, organization and competition over the last several years has added important insights into incremental and radical innovation and the organizational characteristics and competitive consequences associated with them. However there has been strikingly little discussion about intermediate classes of innovation.

We believe that this neglect flows from a reluctance to break open the "black box" that is a specific technology in order to understand its competitive and organizational implications. Although some students of technology and economic history (See for example, work by Rosenberg, (1976,1982); Abernathy, (1978); Abernathy and Clark, (1985); Clark, (1985); and Sahal, (1986)) have described
technologies and examined their inner structure to some extent, few studies have
developed an understanding of how the characteristics of the technology affect the
innovation process. There is an intuitive sense in the literature, and some empirical evidence, that innovations in very different technologies may be
different in character because of the underlying structure of the technology (Moch
and Morse, 1977; Ettlie, Bridges and O'Keefe, 1984; Sahal, 1986; Clark, 1987). But
we have few ways to talk about these differences beyond economic concepts like
minimum efficient scale, returns to scale and appropriability.

Developing a framework that allows one to define and analyze intermediate
classes of innovation may help us to understand an important class of
technological changes, and may also deepen our understanding of innovation,
or ganization and competition in general.

The Framework

The framework we present in this section is summarized in Figure (1). In order to focus the discussion, we develop it in terms of product technology, but the concepts are general; they apply to processes as well. Our starting point is the distinction between the product as a whole — the "system" — and the product in its parts — the "components." (This is an old and well established distinction. See for example, work by Marples (1961) and Alexander (1964).) We conceive of the product as a set of components, (X). Take, for example, a room fan. Its major components include the blade, the motor that drives it, the blade guard, a simple control system and a mechanical housing. For the moment we will use the intuitive

1 Savioit and Metcalfe (1984) have proposed a framework that is related to the one we propose here. They distinguish between three sets of product characteristics: those that describe its technical features, those that describe the services that it performs and those that describe the method of its production.
Figure (1): A Basic Framework.

<table>
<thead>
<tr>
<th>COMPONENTS</th>
<th>SYSTEM PARAMETERS</th>
<th>USER NEEDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>$Y_1 = f_1(x^*)$</td>
<td>$Z_1 = g_1(y^*)$</td>
</tr>
<tr>
<td>$x_1^1 ... x_1^n$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_2$</td>
<td>$Y_2 = f_2(x^*)$</td>
<td>$Z_2 = g_2(y^*)$</td>
</tr>
<tr>
<td>$x_2^1 ... x_2^n$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_i$</td>
<td></td>
<td>$Z_i = g_i(y^*)$</td>
</tr>
<tr>
<td>$x_i^1 ... x_i^j$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Y_m = f_m(x^*)$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
notion that each component fills a distinct function and is physically distinct from the other components.

Each component is characterized by a set of "component parameters" \( \{x\} \), which describe some physical attribute of the component. For the fan blade, for example, the component parameters include the weight of the blade and its dimensions. A complete set of component parameters would constitute an exhaustive description of each component.

At the system level, the product is characterized by a set of parameters \( \{Y\} \) that describe the physical properties of the system as a whole. In our fan example the system properties include things like total weight, the volume of air moved per minute, and resistance to impact. Any individual system property like the volume of air moved per minute is determined by the interaction between several component parameters. For example, the amount of air a fan puts out is a function of the size and shape of the blade, the power of the motor, and the efficiency with which the motor drives the blade. We represent this in the model by showing each system parameter as a function of some subset of the component parameters.

Thus, we have:

\[
Y_m = f_m(x^*),
\]

where \( m \) is an index of the \( M \) system parameters, and the argument of \( f_m() \), \( x^* \), is a subset of the component parameters. The function \( f() \) is a statement of the physical laws governing the relationships between the component and system parameters. Sometimes it can be approximated by mathematical formulae derived from an understanding of the physics involved or estimated empirically.

A single component parameter can affect several system parameters. For example, the dimensions of a fan blade affect not only the amount of air that it puts out but also its esthetic appeal and its portability. This restricts the available set of system parameters. It is very difficult, for example, to design a fan that weighs as little as an alarm clock but that puts out enough air to keep a
hundred people cool. Similarly, component parameters often interact with each
other in determining a system parameter. The effect of a change in blade shape on
the amount of air moved, for example, depends on the power of the fan's motor.
Thus the design of a product is a complex process of simultaneously determining
component and system parameters.

So far the framework gives us the basis for talking about changes in the
internal structure of the technology. But we also need to understand the
connection between the technology and customer needs. We begin by assuming
customers evaluate the product against some set of "needs," (Z). In the case of a
room fan the relevant criteria or needs might include ease of operation,
reliability, transportability and aesthetics (brand image, appearance). The extent
to which a product meets a particular need, that is, the value of a given element
of Z, depends on some subset of the system parameters of the product. A fan's
degree of reliability, for example, depends on its resistance to being dropped,
hit or twisted and the mean time between failure of the entire system. We
represent these relationships in the model as:

\[ Z_n = g_n(Y^*) \]  \hspace{1cm} (2)

where n indexes the N user needs, and the argument of \( g_n(\cdot) \) is a subset of the M
system parameters, \( Y^* \). \( g_n(Y^*) \) is a function of the nature of the customer and of
the way in which the product is used as well as of the physical relationships
between the product's system parameters and the criteria that the customer uses to
evaluate the product. Thus, for example, the extent to which a user's need for
reliable operation is met depends on a set of physical relationships (e.g. system
reliability, frequency of drops) that are partly inherent in the fan's design and
partly dependent on how the fan is used. Meeting the need for a fashionable or
"attractive" fan depends on the interaction between the fan's design, its local
environment and the customer's expectations of what an "attractive" fan would be.
Such expectations may be less stable than the physical relationships underlying
reliability. Thus, $g_n(Y^*)$ captures both the physical and the behavioral and social processes that determine the ways in which the use of the product maps system parameters into user needs. If customers differ in the way in which they use the product, then the function $g_n(Y^*)$ may differ between customers.

Just as the design of a product implies a choice of tradeoffs between component and system parameters, so product design also embodies either implicit or explicit tradeoffs between user needs. The users' own preferences across these tradeoffs can be represented through a "utility function" that captures the value to the customer of some particular set of levels of user needs.

$$u_\alpha = h(Z)$$

(3)

Where $\alpha$ indexes the customers with different preferences. The function $h(Z)$, like the functions $h_m(x^*)$ and $g_n(Y^*)$ may embody interactions between its arguments. For example, additional aesthetic appeal may not enhance the value of a fan if it does not meet some minimum standard of reliability, and quiet operation may be much more highly valued if the fan is a relatively small one.

Types of Technological Change

This framework allows us to characterize "incremental" and "radical" innovation and provides us with a useful framework for analyzing intermediate classes of innovation. In the context of the model, the emergence of a "dominant design" (Abernathy, 1978; Sahal, 1986) is equivalent to the emergence of a stable set of components, system parameters and customer needs. "Incremental" innovation improves individual component performance, changing the values of one or more component parameters, and thus changing the values of some of the system parameters and user needs. But it leaves the elements of the sets $X, Y$ and $Z$ and the relationships between them $h_m(x^*)$ and $g_n(Y^*)$, unchanged.
"Radical" innovation occurs when a new design changes the set of components, system parameters and user needs, or the elements of the sets X, Y, and Z, thereby also changing the relationships between them, or the form of the functions \( h_n(X') \) and \( g_n(Y') \). In its extreme form, radical innovation introduces a completely different set of components and a wholly new system. But in general, "more radical" innovations introduce more new components and consequently affect more system parameters and more user needs.

We define a generational innovation as innovation that changes the relationships between the components, system parameters and user needs of the technology but that leaves the set of components, system parameters and user needs themselves relatively unchanged. While incremental innovation changes the value of the sets X, Y, and Z, and radical innovation changes the elements of the sets X, Y, and Z and the relationships that link them \([h_n(X'), g_n(Y')]\), generational innovation changes the values and relationships, but leaves the elements largely unchanged. Generational innovation is sometimes triggered by changes in a particular component technology, but its essence is a fundamental reconfiguring of the technological system around an essentially stable set of components, system parameters, and customer needs.

The room air fan example can illustrate the distinctions that the model allows us to draw. Suppose, for example, that the established product concept is a large, electric powered fan, mounted in the ceiling, with the motor hidden from view and insulated to dampen the noise. The control system is an on-off switch mounted on the wall and connected to the fan through a set of wires that run inside the wall and ceiling.

An incremental innovation could involve an improvement in blade design and in the power of the motor to achieve a higher rate of cooling. The introduction of a central air conditioner would be a radical innovation. Not all of the established knowledge base would become obsolete. The new technology would involve using fans to move air and would require knowledge of electric motors and fan design. But new
components associated with compressors, refrigerants and their associated controls would add whole new technical disciplines and new inter-relationships. Furthermore, the product would be used in a very different way and might be sold to wholly different customers.

The distinction between radical and incremental innovation is therefore clear. What of generational innovation? For the maker of large ceiling mounted room fans, the introduction of a portable fan would be a generational innovation. While the basic components would be largely the same (e.g. blade, motor, control system), the design choices would be different (e.g. different materials) and the values of the component parameters would change (e.g. smaller dimensions). In this sense, the change from ceiling to portable fan has the character of an incremental innovation. But there would also be significant changes in the interactions between components and in the menu of user needs the product could meet. 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 generate. Shrinking the size of the apparatus would probably require new materials with new properties, as well as new interactions between performance and weight. On the user side, there could be new tradeoffs between ease of use, safety and reliability. It is the need to explore these new patterns of interaction and these new sets of tradeoffs that sets apart this kind of innovation.

Radical innovation obsoletes much of the existing knowledge and the knowledge processing capabilities of established firms, but this obsolescence is usually immediately evident. The obsolescence of established firm knowledge triggered by generational innovation is much harder to observe and may be more difficult to correct. We can best illustrate these differences, and explore their competitive significance, through a discussion of the knowledge required to undertake successful product design.
We have chosen to concentrate our attention on the problem of product design since it is one that explicitly requires the integration of technological knowledge with knowledge of the market and of customer requirements. Moreover it is a problem that several researchers have suggested is critical to successful organizational response to innovation. (Clark (1985), (1987), Freeman (1982)).

In the discussion that follows we model the organization undertaking product development as a boundedly rational individual with limited knowledge and with limited information processing abilities. This abstraction is in the tradition of a number of researchers who have studied the competitive implications of technical change, including Arrow (1974), Nelson and Winter (1982) and Simon (1969), and we believe that it provide a useful "first cut" at the problem. The implications of any innovation for a particular organization will clearly depend upon the way in which knowledge inside the organization is managed, but we have abstracted from this issue in order to focus clearly upon the differences in the information processing tasks presented to the organization by different types of innovation.

Consider the information that is required to design a new product. A designer must know enough about the technologies of each of the components to be able to generate some set of component parameters \( x \) (Marples, 1961; Ramstrom and Rhenman 1965). They must understand how these component parameters interact with each other to produce a product characterized by a given set of system parameters, or something of the relationships \( h(x) \), and they must understand how these system parameters fill a particular set of user needs, or understand some portion of the relationships \( g(Y) \). Finally, in order to design a new product successfully a designer must be able to roughly assess the relative economic value of different

---

2 We deliberately avoid a discussion of "invention," or of the original source of the new idea. While the problem of where inventions are likely to come from is important and interesting, we believe that a study of the implementation of that new idea, or its translation into a saleable product is at least as important, and provides more insight into the nature of the information processing task that an organization must undertake in order to be commercially successful.
clusters of user needs: that is they must know something about the customers' utility functions $U_i(Z)$.

If a designer had perfect information about the customer and the technology and was faced with no significant uncertainty, designing a new product would be a trivial exercise. The designer would have complete knowledge of the way in which different sets of component parameters ($x$) could be generated and of the functions $h_m(x^*)$, $g_n(Y^*)$ and $U_i = h(Z)$ and could develop an "optimal" product. However a designer of limited information processing abilities with only limited experience is unlikely to be in this position. Their knowledge will be incomplete, and will be a function of the recent history of innovation and of the marginal value of different types of information at different stages in the product's evolution (Simon, 1969; Newell and Simon, 1972).

As a technology evolves both the designer's stock of knowledge and the routines or procedures that they use for acquiring new knowledge change. The process of design always requires the development of new knowledge, but the types of new knowledge that are required differ dramatically with the nature of the innovation. During periods of radical innovation there is widespread experimentation in product design. Products are characterized by widely different sets of components, system parameters and user needs. For any single product, knowledge of the relationships between components, system parameters and user needs is likely to be tacit and incomplete - a designer is likely to have very little knowledge of the relationships $h_m(x^*)$ and $g_n(Y^*)$, since every new product embodies a widely different set. Knowledge about the values that customers place on different configurations of user needs, and of the most cost effective way of meeting them will also be scarce. The process of radical innovation can be thought of as a process of exploring the market and the technology "space" in order to better understand both the customer needs and the technological options. The limited information gathering and processing capabilities of the designer are most
effectively orientated towards learning about new technologies and new needs, and towards exploring the relationships between them in a relatively experimental or tacit way. In terms of our framework, the designer develops knowledge about a large number of possible members of the sets \((X), (Y)\) and \((Z)\), and the information processing routines that they develop to search for new knowledge are orientated towards learning about potential new elements of the set. Their knowledge of the universe of possible relationships between them is likely to be tacit and incomplete.

The transition to incremental innovation changes the types of knowledge that are most useful to the designer. While competitive advantage during periods of radical innovation is gained by the introduction of quite different products incorporating quite different component technologies, during periods of incremental innovation it is gained by the more effective exploitation of a limited set of components and user needs within the context of a stable set of interactions. It becomes both possible and valuable to develop a more detailed understanding of the relationships between the existing components, system parameters and customer needs. As the technology matures this knowledge is likely to become widely diffused or relatively "cheap" and the designer's time is likely to acquire detailed knowledge about the performance of particular components and about a limited set of component/parameter/customer need interactions. In terms of our framework, the designer develops very detailed knowledge about particular members of the sets \((X), (Y)\) and \((Z)\) and, since these elements remain stable, detailed knowledge about some known subset of relationships \(h_{n}(x^{*})\) and \(g_{n}(Y^{*})\). Moreover they will develop strategies for acquiring new knowledge that are tightly focused on these particular component technologies and on a limited set of particularly critical interactions.

Given this context, we can begin to understand the implications of "generational" innovation for the knowledge and knowledge processing capabilities
of a designer. During generational innovation the sets (X), (Y) and (Z) remain relatively stable, but the relationships between them change. This obsolescence of the "relational knowledge" of the designer may be just as significant as the obsolescence that is associated with radical innovation. In some situations it may be even more significant because it is more difficult to identify.

The advent of radical innovation - of completely different component technologies or quite different types of user needs - is usually unmistakable, and a designer runs little risk of assuming that their historical knowledge is relevant. But generational innovation is more subtle, and it may be much more difficult to notice that one's relational knowledge has become obsolete. The set of components, system parameters and user needs remains stable, and much of the designer's knowledge remains relevant. There are fewer signals to alert the designer to the nature of the innovation, and a designer may attempt to meet the threat of a competitive product with a design that is based upon his or her historically derived relational knowledge and the routines that they have developed to gather new knowledge that were appropriate to the previous generation of the technology. As a result they run the risk of producing a significantly inferior product and "failing" in the face of generational innovation.

A characterization of the information processing task that is performed within an organization designing a new product is a much more complex problem. But to the extent that the dynamics that we have described in the case of a single boundedly rational individual also characterize the dynamics of the evolution of the knowledge and knowledge processing capabilities of an organization, this framework gives us insight into the reasons that generational innovation may have such dramatic organizational and competitive implications.
III

Empirical Section

As a means to deepen our discussion we now turn to a description of the history of the semiconductor optical photolithographic equipment industry. We have suggested that in order to adequately characterize technological innovation it is necessary to break open the "black box" of a technology to examine in some detail the relationships between its components, system parameters and customer needs. Consequently we present in this section a detailed analysis of semiconductor photolithographic alignment technology focused at the component level. We use this analysis to characterize the history of innovation in photolithography and then explore the extent to which it is a source of insight into the industry's competitive history. Our empirical results are suggestive, but they are presented here only as an illustration of the explanatory power of our framework. The rigorous formulation and testing of appropriate hypotheses remains to be done.

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." Figure (2) schematically illustrates the lithography process. 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 is

---

3 Our analysis draws on the results of a much larger study of the technical and competitive history of the industry (Henderson, 1988). This study included the construction of comprehensive technical, managerial and sales histories for every product development project undertaken in the industry's history since 1965. These histories drew from field interviews with over one hundred individuals, including product designers and leading engineers and scientists, and from intensive study of internal firm documents, the trade press and the major scientific journals. An important element of the work was the use of an iterative validation process to ensure its accuracy. At each stage of the research written summaries of the results and the preliminary conclusions were circulated to key individuals and confirmed through follow up interviews.
Figure (2): The Lithographic Process

1. Expose Resist
   
   ![Diagram ofExpose Resist]
   
   Mask
   Resist
   Wafer

2. Develop Resist
   
   ![Diagram ofDevelop Resist]
   
   Resist
   Wafer

3. Deposit Material

   ![Diagram ofDeposit Material]
   
   Material
   Resist
   Wafer

4. Remove Remaining Resist

   ![Diagram ofRemove Remaining Resist]
   
   Pattern formed on wafer

This is a highly schematized version of a very complex process. For more detail, see the book edited by Watts and Einspruch (1987).
falls onto the resist, so that only those portions of the resist defined by the mask are exposed to light, or "exposed." The light chemically transforms the resist so that it can be stripped away, leaving the unexposed areas available as a basis for further processing.\(^4\) 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.

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 (1) describes the principal components of a generic optical photolithographic aligner. The core technologies of photolithographic alignment have remained stable since the technology was first developed in the middle sixties. Despite this stability, the competitive history of the industry has been strikingly turbulent. A sequence of dominant firms have each in turn failed to maintain their position in the industry. Table (1) shows the sales histories of the leading firms. The first commercially successful aligner was introduced by Kulicke and Soffa in 1965. They were extremely successful and held nearly 100\% of the (very small) market for the next nine years, but by 1974 Cobilt and Kasper had replaced them and held approximately half of the market for contact aligners each. In 1974 Perkin Elmer entered the market, and immediate became the largest firm in the industry. Further entry followed in the late 1970s, and by 1981 GCA and Canon were the leading players. As of this writing GCA has also lost its dominant position and while Canon remains an important player Nikon is probably the largest firm in leading edge photolithographic equipment today.

\(^4\) Resist may be either "negative" or "positive." If a negative resist is used the unexposed areas are stripped away during processing. If a positive resist is used the exposed areas are stripped away. A more complete description of semiconductor lithography is available in Watts and Einspruch (1987).
Figure (3): The Principle Components of a "Generic" Optical Photolithographic Aligner.¹

- **SOURCE**: Provides the light necessary to expose the wafer.
- **MECHANICAL SYSTEM**: Holds other components in place, moves mask and wafer relative to each other.
- **ALIGNMENT SYSTEM**: Checks that the mask and the wafer are correctly positioned.
- **MASK**: Used to focus the image of the mask on the wafer surface.
- **LENS**: Focuses the image of the mask on the wafer surface.
- **WAFTER**: Wafer to be aligned.
- **STAGE**: Platform for the wafer.

**CONTROL SYSTEM**: Controls the operation of all the components.

¹ Specific aligners differ in the way in which these components are implemented, but they all conform to this conceptual design.
Table (1): Total aligner sales by company
(SM)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cobilt</td>
<td>6</td>
<td>7</td>
<td>10</td>
<td>18</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kasper</td>
<td>1</td>
<td>1</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kulicke &amp; Soffa</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perkin Elmer</td>
<td>21</td>
<td>31</td>
<td>48</td>
<td>99</td>
<td>108</td>
<td>159</td>
<td>127</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCA</td>
<td></td>
<td>16</td>
<td>71</td>
<td>44</td>
<td>201</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nikon</td>
<td>0</td>
<td>1</td>
<td>29</td>
<td>193</td>
<td>85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canon</td>
<td>15</td>
<td>34</td>
<td>58</td>
<td>145</td>
<td>139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total aligners: 3 9 40 51 96 255 322 866 547

Sources: Dataquest, VLSI Research Inc., Internal Firm Records.

These shifts in market dominance were triggered by the advent of generational innovation. In every case the established firms invested heavily to meet the threat of new competition, but despite their efforts the products they designed failed to gain widespread customer acceptance - they were simply not as technically successful as those introduced by entrants.

In order to explore the pattern of innovation in the industry, it is helpful to start with an analysis of the changes in the customer needs that aligners have satisfied and to work back through the system parameters that have supported them to an understanding of changes in the components of the technology. Figure (4) lays out an aligner's principal system parameters and user needs.

As is the case with most industrial products, the major concerns of the users of alignment equipment are technical performance and cost. The most critical measure of an aligner's technical performance is "minimum feature size," or the size of the smallest element that an aligner can accurately and reliably reproduce.
Figure (4): Principle elements of photolithographic technology.

<table>
<thead>
<tr>
<th>COMPONENTS</th>
<th>SYSTEM PARAMETERS</th>
<th>CUSTOMER NEEDS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SOURCE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength</td>
<td></td>
<td>MINIMUM FEATURE SIZE</td>
</tr>
<tr>
<td>Uniformity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collimation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MECHANICS</strong></td>
<td></td>
<td>YIELD</td>
</tr>
<tr>
<td>Contact accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gap uniformity</td>
<td></td>
<td>CRITICAL DIMENSION CONTROL</td>
</tr>
<tr>
<td>Gap accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stability</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OPTICS</strong></td>
<td></td>
<td>THROUGHPUT</td>
</tr>
<tr>
<td>Numerical aperture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lens distortion</td>
<td></td>
<td>OVERLAY</td>
</tr>
<tr>
<td>Lens stability</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ALIGNMENT</strong></td>
<td></td>
<td>RELIABILITY</td>
</tr>
<tr>
<td>Wavelength</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algorithms</td>
<td></td>
<td>FOOTPRINT</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Targets</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CONTROLS</strong></td>
<td></td>
<td>SETUP TIME</td>
</tr>
<tr>
<td>Structure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distribution</td>
<td></td>
<td>MAXIMUM WAFER SIZE</td>
</tr>
</tbody>
</table>


on a wafer. All other things equal, the smaller the size of a semiconductor
device, the faster it can run and the cheaper it is to manufacture, and since
historically device size has been limited by the minimum feature size capability
of the lithographic process, users have demanded aligners with smaller and smaller
minimum feature size capability. Users are also concerned about the equipment's
throughput, yield, footprint, maximum wafer size capability, reliability and
flexibility.5

An aligner is an extremely complex piece of equipment, and can be
caracterized by many system parameters. The most critical are those that support
its central function - its ability to accurately and consistently replicate an
extremely small pattern on the wafer, or to support particular minimum feature
size capability. Three system parameters are particularly critical in this
respect: "resolution," "critical line width control" and "overlay." Figure (5)
defines these three parameters and illustrates their relationship to each other.
An aligner's minimum feature size capability cannot exceed its resolution since
its resolution is the size of the smallest optical image that it can transmit to
the wafer. But superb resolution is useless if the image cannot be accurately
positioned or reliably reproduced. If the aligner is not accurate - that is if its
overlay characteristics are not very good - or if the process of image transfer is
not reliable - that is if the aligner's critical dimension control is not adequate
- then the aligner's minimum feature size capability will be less than its minimum
resolution.

5 The maximum size of the silicon wafers used in production has grown from less
than one inch in diameter in the sixties to the eight inch wafer® used today, and
users need lithographic equipment that can handle the size that they have chosen.
The throughput, yield and footprint of an aligner all drive its effective cost.
The faster the throughput and the higher the yield, or the more "good" wafers
produced per hour, the cheaper the aligner is to operate. "Footprint" is a measure
of the area that the aligner requires on the semiconductor production floor. Since
semiconductor facilities are extremely expensive, customers prefer aligners to be
as small as possible.
The basic concepts of alignment technology have remained stable since it was first developed, but the industry has seen all of the three types of innovation that we have identified: radical, incremental and generational. Commercial production has been dominated by optical photolithography, in which light is used as the exposure source, but radical alternatives that make use of alternative sources and quite different mask, alignment and image transfer technologies have been explored since the early seventies. They offer customers better minimum feature size capability, but to date both their cost and a number of unsolved technical problems have prevented them from being widely used beyond research and development. We therefore focus here on the optical systems that have dominated the industry.

Incremental innovation has been critical to optical photolithography's continuing success. The technology of each component has been significantly

6 Radical alternatives to optical photolithography include X-ray and Ion Beam aligners that use x-rays and ion-beams respectively as a source and Direct-write electron beam technology, that uses a beam of focused electrons to "write" on the wafer. (Chang et al., 1977; Brown, Venkatesan and Wagner, 1981; Burggraaf, 1982)
improved; for example modern sources are significantly more powerful and more uniform, alignment systems are much more accurate and lenses are larger and suffer less from distortion.

The industry has also seen significant generational innovation. The key relationships between components and system parameters, and between system parameters and customer needs have shifted dramatically four times over the course of the industry's history as the industry has shifted from the simple contact and proximity aligners to the more sophisticated optical systems that "scan" the mask relative to the wafer or that "step" it slowly across the wafer surface. One indication of the presence of generational innovation is visible in figures (6) and (7), which show the historical performance of each generation in terms of its throughput, yield and minimum feature size. Each arrow reflects a particular technological trajectory, and summarizes a history of incremental improvement within each generation. The movement between arrows - from one technological trajectory to another - reflects generational innovation in the underlying technology.

We can develop more insight into the nature of generational innovation in the industry by comparing the relationships between components and system parameters that underlie the user need for minimum feature size in the contact and proximity aligner. Contact and proximity aligners are relatively simple so that their technology can be easily described, and a focus on minimum feature size, the most critical dimension of the aligner's performance, allows us to grasp the essence of generational innovation without the need to present an exhaustive analysis of the determinants of every system parameter.

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. (Figure
Figure (6): Changes in the Balance Between Minimum Feature Size and Yield Across the Changing Generations of Photolithographic Equipment.¹

1. This table is designed only to give a sense for some general relationships. Minimum feature size achieved and production yield vary greatly between customers and between applications.

2. Production yields are difficult to measure and are considered very confidential. Typically a "poor" yield would be a yield of about 20% - a "good" yield might be on the order of 60-70%, although these figures would vary with the stage in the process at which the yield was measured.

Source: Field interviews, Internal firm records. (Henderson, 1989)
Figure (7): Changes in the Balance Between Minimum Feature Size and Throughput Across the Changing Generations of Photolithographic Equipment.¹

1. This table is designed only to give a sense for some general relationships. Minimum feature size achieved and throughput vary great between customers and between applications.

2. Throughput is defined as throughput of the largest wafer size that the aligner is designed to handle in wafers per hour.

Source: Field interviews, Internal firm records. (Henderson, 1989)
Figure (9) presents a schematic diagram of a contact aligner. 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. 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. (Figure (9) presents a schematic diagram of a proximity aligner.) As a result users switching to proximity aligners were forced to trade off minimum feature size capability for increased yield.

The introduction of the proximity aligner is clearly not a radical advance over the contact aligner. The conceptual change involved in going from one to the other was minor, and most proximity aligners can also be used as contact aligners. However, as figures (10) and (11) reveal, the relationships between component parameters, system parameters and minimum feature size are quite different for the two technologies. The introduction of proximity alignment introduced generational innovation into the industry.

The minimum feature size capability of a contact aligner is limited by its overlay and critical dimension characteristics. As long as contact between the mask and the wafer is "perfect" - absolutely flat and parallel across the wafer - yield loss is minimal, the image of the mask is not distorted and the aligner's minimum feature size capability is limited only by its resolution. Since the contact aligner does not rely upon either an optical system or upon a gap to transmit the image of the mask to the wafer, resolution is limited only by the source wavelength, and until recently the resolution of contact aligners has been greater than that of any of the other generations of equipment. Unfortunately in
Figure (8): Schematic Diagram of a Contact Aligner

Source

Alignment System

Mechanical System

Mask

Wafer

Stage

Control System

Figure (9): Schematic Diagram of a Proximity Aligner

Source

Alignment System

Mechanical System (Includes gap setting mechanism)

Mask

Wafer

Stage

Control System
Figure (10): THE CONTACT PRINTER

COMPONENTS

SOURCE
Wavelength
Uniformity
Collimation
Energy

MECHANICS
Contact accuracy
Gap uniformity
Gap accuracy
Stage accuracy
Stability

ALIGNMENT
Wavelength
Algorithms
Accuracy
Targets

SYSTEM PARAMETERS

RESOLUTION

CRITICAL DIMENSION CONTROL

OVERLAY

MINIMUM FEATURE SIZE

CUSTOMER NEEDS

Indicates those relationships that are critical and that limit the performance of the aligner.

Source: Field interviews, Internal firm records. (Henderson, 1989)
Indicates those relationships that are critical and that limit the performance of the aligner.

Source: Field interviews, Internal firm records. (Henderson, 1989)
production settings contact is far from perfect. In practice both the mask and wafer are often damaged or distorted, and minimum feature size is limited by the overlay and critical dimension control characteristics of the aligner. These are a function of the precision of the contact mechanism and the accuracy of the alignment system.

In contrast, the minimum feature size capability of a proximity printer is limited by its resolution and by its critical dimension control, and these are functions of quite different component parameters. Resolution is limited not by the wavelength of the source but by the size of the gap between the mask and the wafer. As a result although the minimum feature size capability of a contact aligner is rarely limited by its resolution, resolution is very much a binding constraint in a proximity aligner. The aligner's critical dimension control is limited by the uniformity of the gap between the mask and the wafer, since if the gap is wider in some places than in others the light will have a chance to "spread out further," producing an image of a different size at the wafer.

The framework that we have developed thus allows us to make the nature of generational innovation precise. Proximity alignment appears to be a minor innovation. The components, system parameters and user needs met by the aligner remained unchanged. Yet the relationships between the components, system parameters and user needs changed dramatically. An apparently small change in a particular component technology had important repercussions in the way in which the entire system operated.

We have used this framework to analyze the later stages of the industry's technical evolution (Henderson, 1989). Table (2) summarizes our results and outlines the key changes in technical relationships that underlay the generational innovations that have marked the industry's history. In each case, the introduction of a single new component or of small changes in existing components
## Table (2): A Summary of Relational Change in Photolithographic Alignment Technology

<table>
<thead>
<tr>
<th>Generational Innovation</th>
<th>Major changes in:</th>
<th>Critical Relationships</th>
<th>Customer Needs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Technology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROXIMITY</td>
<td>Mask and wafer separated during exposure.</td>
<td>MFS (Minimum feature size) a function of gap size, accuracy.</td>
<td>Yields much higher, but MFS now much greater constraint.</td>
</tr>
<tr>
<td>SCANNING PROJECTION</td>
<td>Image of mask projected onto wafer by scanning reflective optics.</td>
<td>Resolution driven by numerical aperture, wavelength. Overlay, - a function of lens distortion - is critical constraint for MFS.</td>
<td>MFS constraint relaxed, but scanner is slower.</td>
</tr>
<tr>
<td>STEP &amp; REPEAT (1)</td>
<td>Image of mask projected through refractive lens. Image &quot;stepped&quot; across wafer.</td>
<td>Throughput major constraint - function of lens field size and source energy.</td>
<td>Some customers have large increases in yield, theoretical limits to MFS rise, But throughput falls significantly.</td>
</tr>
<tr>
<td>STEP &amp; REPEAT (2)</td>
<td>Introduction of &quot;site by site&quot; alignment, larger 5x lenses.</td>
<td>Overlay and resolution constrain MFS. Resolution driven by Numerical aperture, wavelength. Overlay and reliability driven by lens distortion, interactions between lens and mechanical system.</td>
<td>Dramatic improvement in reliability, lens field size improves throughput. MFS again major constraint.</td>
</tr>
</tbody>
</table>

Source: Field Interviews, Internal Firm Records, (Henderson, 1988)
changed the core relationships that determined the functioning of the aligner and structured new tradeoffs across the customer needs.

Generational Innovation and Organizational Capability.

Insights into the technical history of the industry are intriguing. But our framework also implies that generational innovation should have important competitive consequences. We suggested that generational innovation obsoletes the "relational knowledge" of the product designer, making it difficult for them to understand the nature of generational innovation and to respond to it appropriately. Our empirical results suggest that this insight is critical to an understanding of the competitive history of the photolithographic alignment industry. A full analysis of the industry's history should obviously consider more than simply changes in the technology. But in this type of industry, in which the performance of the product is critical to its acceptance in the market, and in which firms compete primarily on the basis of technical excellence, a careful study of the competitive implications of technical change can reveal patterns that may be important in a wide range of other industries.

Table (3) presents share of sales by generation of equipment for the leading firms. None of the established firms in the industry has been able to extend its position into the next generation of equipment, despite its experience with the technology and its ownership of an extensive installed base.
Table (3): Share of sales by generation for the leading optical photolithographic alignment equipment manufacturers.\(^7\)

<table>
<thead>
<tr>
<th></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></td>
<td>67</td>
<td>21</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Perkin Elmer</td>
<td></td>
<td>78</td>
<td></td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>GCA</td>
<td></td>
<td>69</td>
<td></td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Nikon</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55</td>
</tr>
</tbody>
</table>

Sources: Dataquest, VLSI Research Inc, Internal Firm Records.

There are a number of possible explanations for this failure. Established firms might have failed to invest enough in the development of the new technologies, either through ignorance or through fear of cannibalizing their existing base, or they may simply have been "unlucky." Neither of these explanations stands up to scrutiny. All of the established firms made significant investments in generational innovation, investing on average significantly more than new entrants (Henderson, 1988). And although one cannot rule out the role of chance, our analysis suggests that there are systematic similarities in the pattern of established firm response that persist across the industry's history.

---

\(^7\) Where "share of sales is defined as cumulative sales, in 1986 dollars, of all products produced by that firm in that generation of the technology as a percentage of total cumulative real sales of all products sold in that generation. 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."
despite the face that the identity of the established firm changes in every instance.⁸

Thus traditional economic explanations of the failure of established firms do not seem to hold here. Explanations that rely on "bureaucratic inertia," or on the difficulties that many organizations encounter in attempting to respond any type of change are similarly suspect, since established firms explicitly identified the competitive threat represented by generational innovation, attached great importance to meeting it, and used design teams that were on average composed of fewer than twenty five people (Henderson, 1988). Established firms failed to maintain their position into the next generation because the experience that they had gained with incremental innovation in the previous generation - specifically the relational knowledge that they had accumulated - made it very difficult for them to understand the nature of the generational innovations with which they were faced or to shape an appropriate response.

We illustrate this here by a description of Kasper's failure to maintain its position in proximity alignment. This failure is particularly interesting since the firm was relatively small and "organic" in structure and since the introduction of proximity alignment appears to be so similar to the introduction of an incremental innovation. It illustrates graphically the ways in which subtle changes in the interrelationships of components, system parameters and user needs can have very significant competitive implications.

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

---

⁸ A statistical analysis of the determinants of technical success in the industry's history suggests that established firms were significantly less likely to be technically successful than would be predicted by random chance. (Henderson, 1989)
onwards 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 did not occur until the introduction and general adoption of Canon's proximity aligner in the late 1970s.

Kasper's failure is initially puzzling given its established position in the market and its depth of experience in photolithography. There were approximately five key mechanical and electronic engineers at Kasper during the early 1970s. Several of them were highly skilled and imaginative designers, and the group designed a steady stream of contact aligners, each incorporating significant incremental improvements. From 1968 to 1973 the minimum feature size capability of the contact aligner improved from ten to five microns.

But Kasper's very success in designing contact aligners was a major contributor to their 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 since it incorporated a much more sophisticated understanding of the technical interrelationships that are fundamental to successful proximity alignment. Kasper failed to develop the particular specialist knowledge that would have enabled them to match Canon's design, but, even more importantly, their experience with contact aligners left them without the knowledge gathering routines that would have enabled them to understand the need to acquire it. The relational knowledge that they had developed through their experience with contact aligners had the effect of focusing their 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, and it was managed as a routine extension to the product line, just as the previous incremental improvements to the contact aligner had been. The gap setting...
mechanism that was used in the contact aligner to align the mask and wafer to each other was slightly modified and the aligner was offered on the market. In doing this, Kasper made the implicit assumption that the technology that had sustained contact printing could be incrementally extended to support proximity alignment. But, as the analysis that we presented above made clear, proximity alignment represents a generational shift in alignment technology. In particular, the interrelationships between the performance of the gap setting mechanism and the minimum feature size capability of the aligner are quite different in a proximity and contact aligner.

In a contact aligner, the gap setting mechanism is used only during alignment. Errors in its setting can be corrected manually, and its accuracy has little influence over minimum feature size capability. But in a proximity aligner the quality of the transmitted image is critically dependent upon the gap's accuracy and stability, and minimum feature size capability is a direct function of its size and accuracy. As result the successful design of a proximity aligner requires both the acquisition of some new knowledge — how to build an adequate gap setting mechanism — and an understanding of some new interactions between component performance and the system parameters of the aligner — in particular an understanding that minimum feature size capability is limited by the resolution and critical dimension control of the aligner, and that in turn resolution and critical dimension control are critically dependent upon the accuracy and uniformity of the gap between the mask and the wafer.

The successful design of an adequate gap setting mechanism is not a trivial task. In a contact aligner the designer can rely on the contact between the mask and the wafer to ensure that the mask is located accurately with respect to the wafer. In a proximity aligner, the mask must be located by "dead reckoning" at some point in space above the wafer, and the designer must rely on the accuracy of the mechanical mechanism to ensure that the gap between the mask and the wafer is
precise and consistent across the mask surface. This calls for significant expertise in the design of high precision mechanical mechanisms, and in particular, for a deep understanding of the interactions between the design of the gap setting mechanism and the other elements of the aligner.

Kasper lacked this knowledge, or, in the terms of our framework, lacked the ability to control the component parameter "gap size" with sufficient accuracy. Canon, on the other hand, were able to bring to bear a repertoire of skills and knowledge about high precision mechanical assemblies derived from their experience in related industries. The design group at Kasper had not needed to develop this type of knowledge since it was not critical to the performance of the early contact aligners.

But this lack of a specific piece of technical knowledge is not the whole story. More critically, Kasper continued to make use of the problem solving and information gathering strategies that they had developed during their experience with the contact aligner, and as a result failed to understand the need to develop this critical body of knowledge. Canon invested heavily in building a deep understanding of the precise relationships between errors in the gap setting mechanism, the stability and uniformity of the source, and the performance of the alignment system in the determination of the resolution and critical dimension control of the aligner. This enabled them to make informed tradeoffs in the overall design of the system, and to understand, for example, where design effort could be most fruitfully spent. Kasper did not do this. They assumed that the introduction of a gap between the mask and the wafer during alignment was a minor, incremental change and so put little effort into understanding the new interactions that it created between the gap setting mechanism and resolution and critical dimension control. Thus they failed to understand the enormous impact that improvements in its accuracy and precision could have on the performance of
the aligner, and the proximity aligner that they designed suffered from an unreliable and insufficiently accurate gap setting mechanism.

Kasper's failure to understand the obsolescence of their relational 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 they "knew" from their experience with the contact aligner that their gap setting mechanism was adequate to the task. As a result they 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 them well historically. The second illustration is provided by their response to Canon's initial introduction. 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 they used for evaluating their own aligners - criteria that had been developed during their experience with contact aligners. The technical features that made it a significant advance, particularly the redesigned gap mechanism, were not observed because they were not considered important.

Kasper's engineers were not incompetent. They assumed that the core design concepts that underlay the design of their contact aligners could also serve as the basis for a proximity aligner, and that the relational knowledge and the knowledge processing capabilities that they had accumulated during their experience with contact aligners could be transferred to proximity aligners.

Kasper's commercial failure stemmed from more than this failure in design. The company had problems designing an automatic alignment system of sufficient
accuracy, and in managing a high volume manufacturing facility. They also suffered through several rapid changes of top management during the late nineteen seventies. But the obsolescence of relational knowledge brought about by the introduction of generational innovation was a critical factor in their decline. A similar study of the failure of the other dominant firms in the industry suggests that a reliance on relational knowledge about the technology derived from the previous generation was a significant factor in explaining the failure of all of the established firms who were unable to meet the threat of new entry (Henderson, 1989).

IV

Summary and Conclusions

We have suggested that the characterization of innovation as either "radical" or "incremental" is incomplete, and must be supplemented by the concept of "generational" innovation. While incremental innovation changes the performance of an existing set of components, and radical innovation changes the elements of that set and the relationships between them, generational innovation changes performance and relationships, but leaves the elements largely unchanged. Generational innovation is sometimes triggered by changes in a particular component technology, but its essence is a fundamental reconfiguring of the technological system around an essentially stable set of components and customer needs. It is qualitatively different from both incremental and radical innovation and has equally important competitive and organizational implications.

We used the model of a product designer as a boundedly rational individual of limited information processing capabilities to highlight the different information processing tasks that the different types of innovation pose for the product development process, and discussed the types of knowledge and knowledge acquisition routines that a designer faced with the different types of innovation is likely to develop in consequence. We suggested that to the degree that this
model is indicative of forces at work inside organizations: the generational change may have important and unexpected competitive consequences since it obsoletes the relational knowledge of the organization. This issue is an important avenue for future research.

We illustrated our framework through a description of the technical and competitive evolution of the semiconductor photolithographic alignment industry. We showed that the introduction of proximity alignment, an apparently "incremental" innovation, in fact introduced significant change into the relationships between components, and that Kasper's continued reliance on the knowledge and the information processing capabilities that they had developed through their experience with contact aligners severely handicapped the firm's ability to take advantage of proximity alignment.

This research opens up a number of important questions. For example, we need to develop metrics for generational innovation and to explore its interactions with the other economic and organizational forces that shape firms and industries. But we believe that the concept of generational innovation has potentially great power and will prove useful in the study of the organizational and competitive implications of technical change.
REFERENCES:

Abernathy, William

Abernathy, William, and Kim Clark

Alexander, C.

Arrow, K.

Brown, W.I., T.Venkatesan and A.Wagner

Burggraaf, P.

Burns, T. and G. Stalker

Chang, T.S. et al.

Clark, Kim B.

Clark, Kim B.

Cooper, A.C. and D. Schendel

Dess, G.G. and D. Beard.

Dewar, R.D. and J.E. Dutton,

Dosi, G.

Ettlie, J.E., and W.P. Bridges and R.D. O'Keefe

Ettlie, J.E.

Freeman, C.
1982 The Economics of Industrial Innovation. Cambridge, MIT Press.

Gilbert, R.J. and D.M. Newbery

Hage, J.

Hannan, M. and J. Freeman

Henderson, Rebecca M.

Henderson, Rebecca, M.

Jewkes, J., Sawers, D. and R. Stillerman
1968 The Sources of Invention. New York: Macmillan.

Kimberly, J.R. and M.J. Evanisko

Kuhn, Thomas


<table>
<thead>
<tr>
<th>Date Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov 12 1993</td>
</tr>
<tr>
<td>Sep 1 6 1993</td>
</tr>
<tr>
<td>Feb 13 1990</td>
</tr>
<tr>
<td>Jan 0 2 1990</td>
</tr>
<tr>
<td>Oct 10 1996</td>
</tr>
<tr>
<td>Mar 20 1990</td>
</tr>
<tr>
<td>Apr 0 6 1990</td>
</tr>
<tr>
<td>Mar 26 1999</td>
</tr>
<tr>
<td>Apr 1 2 1993</td>
</tr>
<tr>
<td>May 0 7 1993</td>
</tr>
<tr>
<td>May 0 7 1994</td>
</tr>
<tr>
<td>May 2 5 1991</td>
</tr>
<tr>
<td>May 2 5 1995</td>
</tr>
<tr>
<td>Jul 2 0 1991</td>
</tr>
<tr>
<td>Aug 3 1 1996</td>
</tr>
<tr>
<td>Jun 1 6 1996</td>
</tr>
<tr>
<td>Jun 1 6 1990</td>
</tr>
<tr>
<td>Mar 2 6 1999</td>
</tr>
<tr>
<td>Mar 2 6 1999</td>
</tr>
</tbody>
</table>