3.1 The Evolution of Innovation

Clayton M. Christensen

Scholars have searched for patterns in the evolution of technological innovation for decades, hoping that regularities, frameworks, or theories might emerge to give guidance to managers charged with the complex and uncertain task of managing innovation. The character of these studies has varied. They range from deep historical studies of individual industries, to cross-industry surveys that search for frameworks that are robust enough to describe how innovation has evolved in very different competitive environments, to studies that focus on specific frameworks as being particularly useful. These studies typically search not just for patterns in what types of innovation are likely to occur at different stages of a product or industry’s maturity but for insights about what types of companies are most likely to succeed or fail at these innovations, in different circumstances.

Thomas Kuhn, the historian of science, observed that in the building of bodies of knowledge, paradigms eventually emerge [Kuhn, 1970]. In Kuhn’s model, a paradigm can never explain everything — the world is too complex. Nor can a paradigm explain anything perfectly. However, it is a way of organizing and explaining enough about a class of phenomena that subsequent researchers find it to be a useful starting point for their own work.

No single paradigm has emerged in the study of patterns of innovation that would enable all researchers or managers to predict with certainty how technology is likely to evolve or what types of companies are likely to emerge victorious from innovative battles of various sorts. At this point in our understanding of the field, it seems unlikely that such a comprehensive, overarching theory will ever emerge, and Kuhn probably would not expect such a paradigm to develop. Indeed, the problems of managing technological

---

1Four studies exemplifying the tradition of deep, single-industry historical studies are Abernathy [1978], Constant [1980], Henderson [1989], and Christensen [1993]. Examples of cross-industry surveys are Sahal [1981], Tushman and Anderson [1986], Anderson and Tushman [1990], Rosenbloom and Christensen [1994], and Christensen [1997b]. Research that promotes the usefulness of specific frameworks includes Roussel [1984] and Foster [1986].
innovation are so varied and complex that multiple bodies of knowledge are likely to be required to understand how to manage the evolution of innovation.

There now appear to be potential paradigms emerging in the study of how four particular dimensions or aspects of technological evolution occur. They are potential paradigms in the sense that each has helped make sense of what previously appeared to have been random or contradictory phenomena and because each has spawned streams of sustained, credible subsequent scholarship. These bodies of theory might be labeled as follows:

1. The dominant design theory, which asserts that the nature of innovation shifts markedly after a dominant design has emerged.
2. The technology s-curve theory, which states that the pace of performance improvement utilizing a particular technological approach will follow an S-curve pattern, flattening as certain natural limits to further improvement are reached. Theories of punctuated equilibrium are related to movement along a technology S-curve, intersected occasionally by a new S-curve.
3. The theory that patterns of innovation are determined by intersecting trajectories of performance demanded in the market, vs. performance supplied by technologists.
4. The study of how modularization of design can create options for the future, how it affects the optimal scope of the firm, and how it changes the nature of the competitive advantages that can and cannot be developed.

The following sections summarize each of these viewpoints, characterize the sorts of innovative problems to which the theories seem to have relevance, and describe the sequence of studies that have contributed most strongly to the building of each body of knowledge.

The Dominant Design Theory

The notion that a dominant design powerfully impacts the nature of innovation was articulated by Abernathy and Utterback [1978]. They proposed that in most industries there would be an initial period of product design ferment. In the early years of the automobile industry, for example, fundamental questions such as whether the power source should be a steam-, electric-, or gasoline-powered engine were not yet resolved. How the body would be supported over the drive train, what a transmission was, and how it would interface with the driver and with the engine were characteristic of the design issues that engineers in different firms approached differently from one product generation to the next.

Eventually, however, a dominant design emerged — a general acceptance of how the principal components would interface with others in the architecture of the automobile's design. The dominant design was not necessarily the optimal design. However, it became a standard architecture, with accepted metrics for determining the way in which components and modules would interact. This gave organizationally independent suppliers a well-defined technological context within which they could work to improve their pieces of the system. Abernathy and Utterback noted that the magnitude and rate of technological innovation directed at innovative product design declined markedly after the emergence of the dominant design.

The emergence of dominant designs enabled a surge of innovation in process technology, as suggested in Fig. 3.1. When designs were in flux, processes could not be standardized: volumes per process sequence were low, equipment had to remain flexible, and product costs were high. The dominant design enabled engineers to focus their innovative energies on process technology improvements. This enabled significant cost reductions in the product, allowing further volume growth and even further process refinement and cost reduction.

Other scholars have subsequently built upon the dominant design paradigm, articulating how it operates and what its impact on patterns of innovation can be. For example, Suarez and Utterback [1995] and Christensen et al. [1997] found that, because the existence of dominant designs restricts engineers' freedom to differentiate their products through innovative design, there are fewer opportunities for small or entrant
firms to find refuge in niche markets in the post–dominant design era. They found that firms entering a range of industries after dominant designs were defined faced lower posterior probabilities of survival. They further concluded that dominant designs are architectural in character and that their elements coalesce one by one over time. They found that certain components came to be used in most models, but whether or not these components were adopted had little impact on survival. What mattered was adoption of the dominant elements of architectural design.

Finally, they noted that there appeared to have been a unique “window of opportunity” for entry into fast-moving industries. Not only did firms that entered after the dominant design emerged have low probabilities of survival, but firms that entered too early had low probabilities of survival as well — presumably because they had honed capabilities in that period of high design turbulence that were inappropriate for survival in the process-intensive post–dominant design era.

The Technology S-Curve Theory

The notion that the pace of technological progress follows an S-curve pattern has been featured in literature on technological innovation for years; Foster (1986) summarized the arguments most comprehensively. S-curve theory suggests that the rate of technological progress ultimately is subject to decreasing returns to engineering effort because trajectories of technological progress are eventually constrained by natural limits of some sort: they get too small, too large, too complex, or push the intrinsic properties of available materials to their theoretical maxima [Sahal, 1981]. As these limits are approached, it requires increasing amounts of effort to wring out additional performance improvement.

S-curve theorists argue that, when the rate of technological progress has begun to decline, the technology and its practitioners are vulnerable to being overtaken by a new technological approach, following its own S-curve pattern, as shown in Fig. 3.2. A key management task is therefore to monitor a company’s position on its S-curve and, when it has passed its point of inflection, to find and develop the new technology that might overtake the present approach. S-curve patterns of technological progress and technology substitution have been shown to have occurred in foam rubber, aircraft engines, magnets, and disk drive components.²

Subsequent research has substantially refined the usefulness and limits of the S-curve paradigm. While it can be used to describe overall industry trajectories, it has more limited decision-making usefulness within companies, for two reasons. First, S-curves seem less relevant to performance of assembled

²A listing of these publications can be found in Christensen, [1992].
products than to componentry. Christensen [1992] showed, for example, that the perception of the recording density at which disk drive makers felt they needed to jump to a new recording head S-curve differed across companies in the same industry by an order of magnitude. This was because, in the design of most assembled products, there is more than one route to achieving performance improvement. When the head technology reached its performance limit, some companies in the industry elected to jump to the S-curve of the next-generation component technology, while other firms found design routes around the bottleneck by improving other components and architectural aspects of the design, which were not yet at the apex of their S-curves. Hence, some firms followed an S-curve switching strategy, component by component, as a means of relieving bottlenecks to performance improvement, while other firms followed an S-curve extension strategy — staying as long as possible with existing technologies, through clever innovations in system design. The result of both strategies was a steady improvement in the performance of the finished disk drive — a pace comprised of many incremental improvements and some radical S-curve leaps among the individual components comprising the drives. Iansiti [1995] measured the same phenomena in his study of the evolution of mainframe computer technology.

Another limit of S-curves’ usefulness is that, with many types of innovations, the relevant attributes of performance offered by new technologies differ from those of the old. The new technology, while underperforming the established approach along accepted metrics in established markets, can become established in a new market segment, which values its different attributes. When its performance improves to the point that it satisfies the performance demanded in the original market as well, it then invades swiftly, knocking out the established firms in that market, as depicted in Fig. 3.3. S-curve theory, lacking a market dimension to its definition of performance, cannot account for this route of technological evolution, which subsequent scholars [Christensen, 1992, 1997; Levinthal, 1997] have shown to be quite common.

A stream of research closely related to the S-curve concept might be characterized as a punctuated equilibrium theory. Tushman and Anderson [1986], who initially articulated this point of view, noted that in a range of industries they studied, technologies improved at a relatively measured pace across most of their histories. This incremental pace of progress was sporadically interrupted, however, by bursts of radical change that created discontinuities in the otherwise smooth trajectories of improvement. Levinthal [1997] has described the similarity of this pattern in technological evolution to the patterns that characterize biological evolution.

Organizations’ technological capabilities are developed through the problems their engineers solve. The nature of the problems they confront or do not confront is determined to a significant degree by the prior choices of technology made earlier in the history of the company and industry: hence, technological understanding builds cumulatively [Clark, 1984]. Some scholars have observed that the leading
companies in an industry are most likely to lead in developing and adopting new technologies when those technologies build upon the accumulated technological competencies they have developed. When new technologies render established firms’ historically accumulated competencies irrelevant, however, entrants to the industry have the advantage [Tushman and Anderson, 1986; Henderson and Clark, 1990; Tripsas, 1997].

The Technology and Market Trajectories Theory

The third stream of research around which a significant body of insight is accumulating combines theories about how and why technological progress is achieved, with insights about rates of technological progress customers are able to absorb. Its fundamental premise is that patterns of innovation are influenced heavily by the interaction of trajectories (rates of improvement in product performance) in what customers need, compared to trajectories of improvement that innovating companies provide.

The notion that technological progress of a class of products can be mapped as a trajectory of improvement over time was articulated by Dosi [1982]. A stream of research initiated by Christensen and colleagues has extended Dosi’s notion of technology trajectories through a range of empirical studies [Christensen, 1993, 1997b; Rosenbloom & Christensen, 1994; Bower and Christensen, 1995; Jones, 1996]. They observe that the trajectory of technological progress frequently proceeds at a steeper pace than the trajectory of performance improvement that is demanded in the market. This means that a technology that squarely addresses the needs of customers in a tier in a market today may improve to overserve those needs tomorrow; a technological approach that cannot meet the demands of a market today may

FIGURE 3.3  Route through which disruptive technologies penetrate established markets. (Source: Adapted from Christensen [1992], p. 361.)
improve at such a rate that it competitively addresses those needs tomorrow, in the fashion described in Fig. 3.4.

Most technological innovations have a sustaining impact: they drive an industry upward along an established technology trajectory. Occasionally, however, disruptive technologies emerge — smaller, cheaper, simpler products that cannot be used in established markets because they perform poorly according to the attributes valued there. These disruptive products may enable the emergence of new market segments in which customers have a different rank ordering of product attributes than those of established markets. Practitioners of the disruptive technology can take root in this new segment, even while manufacturers and customers in mainstream markets ignore it.

Once commercially established in this new "low" end of their larger market, these disruptive technologists have very strong incentives to improve their products' performance at such a rapid rate that they can attack market tiers above them. This is because those market tiers typically are larger, and profit margins are more attractive in the higher-performance product models purchased there.

These scholars have found that the established companies in each industry they studied generally led their industries in developing and adopting innovations that were sustaining in character — even radical, discontinuous, competence-destroying technologies. It appears that, when the customers of established companies have demanded an innovation, the leading firms seemed somehow to find a way to get it. However, entrant firms consistently led in introducing disruptive technologies that could be used only in new or commercially unimportant markets. Strong, established companies that listened attentively to their customers and were skilled at directing their innovative investments to projects that promised the greatest profits typically found it nearly impossible to introduce disruptive technologies in a timely way, despite the fact that they typically are technologically simple. This pattern of innovation affecting the fortunes of established and entrant firms has been observed in retailing [McNair, 1958], telecommunications switching equipment [Jones, 1996], semiconductor testing equipment [Bower, 1997], commercial printing [Garvin, 1996], and the steel, mechanical excavator, computer, motorcycle, photocopier, accounting software, and executive education industries [Christensen, 1997a, 1997b].

The mechanism discovered by these scholars, through which disruptive technologies invade established markets and precipitate the decline of leading firms in those markets, is very similar to the way in which biologists have modeled the historical evolution of new species [Levinthal, 1997]. Rarely does speciation occur within established populations, in habitats to which they are well adapted. Rather, it occurs in peripheral or remote environments where the characteristics of a new (mutated) species give it an
advantage vs. established species in the competition for survival. Established species can then be wiped out when members of the new species return to the home range from the remote environment in which their species initially took root.

The way that performance improvement is defined seems to be altered significantly in each market segment after a disruptive technology trajectory intersects the performance needed in the mainstream market, as depicted in Fig. 3.4. Once the older and the disruptive technology both provide adequate performance, how do customers choose between the alternative products? Researchers [Christensen, 1997a; Trent, 1997] have observed that, when two alternatives both overshoot the functionality that customers in a tier of the market actually need, the basis of competition, or the criteria by which customers choose one product over another, tends to change in that market tier, in a relatively consistent pattern. No longer can product functionality be a competitively relevant dimension of innovation: instead, a new trajectory of performance improvement is defined — often it centers on reliability — and the successful technologists innovate along that dimension instead. Ultimately, the trajectories of improvement in reliability overshoot the market need, stimulating another trajectory to coalesce — typically along convenience dimensions. It is the interaction of trajectories of technology improvement and market need that trigger changes in the basis of competition or in the definition and trajectory of competitively meaningful innovation.

Theories about the Modularization of Design

The fourth stream of ideas around which a substantial body of research is beginning to build relates the modularization of product designs. These concepts are substantially less developed than the other three — yet they appear to have the potential to enhance substantially what is understood about the evolution of innovation.

Most complex products consist of subassemblies and components, which themselves are often assemblies of yet smaller components, in a nested fashion [Alexander, 1964; Christensen and Rosenbloom, 1995]. In the earliest phases of the development of a product category, complex products often must be designed in an integrated way, meaning that each element of the product must be designed, simultaneously or iteratively, by the same organization. This is because the ways that the subassemblies, components, and materials interact with each other are not well understood: changes in the design of one component may change the way other components in the system perform, in ways that initially cannot be anticipated [Ulrich and Eppinger, 1995].

Such integrated design problems — essentially having to redesign the entire product in every generation — can be very expensive. For this reason, designers often work to understand the interactions among components and materials, to establish standards for how they should interface, and to define what attributes of each component must be measured and how to measure them. When these things are achieved, the design has become modularized.

Once modularization has been achieved, engineers are able to change or upgrade individual modules in the product without unexpected changes occurring in the performance or reliability of other pieces of the system. New and improved modules either can be “plugged and played”, as can occur in most home stereo systems, or engineers at least can understand what changes they’ll need to make in other parts of the product’s design to account for what they’ve changed in one module, or to account for variation in the attributes of a component. Modularization makes improvements to product functionality and additions to features relatively fast and inexpensive. Modularity expands the economically viable options available to innovators and can be valued in the same way as financial options [Baldwin and Clark, 1997].

Industries can pass through cycles of modularization and demodularization as certain new technologies enter an industry. If a new technology changes the way components interact in the system design, engineers may no longer be able to predict how changes in the component or material (in which the new technology is embodied) might affect performance elsewhere in the system. They consequently will be unable to specify completely what critical attributes of the component must be met, meaning that at
least some elements of the design have become integral again. Integrated design increases the cost and
time required to create new product designs, however, so engineers will likely work to understand
how the new technology component interacts with other elements of the system in an effort to remod-
ularize the product. Hence, industries can go through repeated cycles of modularization and demodu-
larization [Christensen, 1994].

Modularity in product design also creates the potential for modularity in organization design. It is
very difficult for a firm to source an integral component from an outside supplier. Because it cannot
know how the component might interact with other elements of the system’s design, the company cannot
clearly specify what it needs its supplier to deliver. Once a modular design has been achieved in one or
more components, however, then outsourcing is indeed a very viable option; standard interfaces between
modules create a clear, “standard” way for the two organizations to interface [Sanchez, 1996]. There is
substantial evidence, in fact, that, when product designs become modular, integrated organizations that
produce components and design and assemble the final product lose market position to firms that
flexibly can mix and match the most cost-effective components from the best independent sources.
When new technology creates problems of integrality again, advantage shifts back to integrated com-
panies that not only design and assemble the final products but manufacture the integral components
as well [Christensen, 1994]. Indeed, the concept of the “virtual corporation” which moves alternately
in and out of vogue in the business press, is viable only if clear, well-defined interfaces are established
amongst all components and materials in the product, and across all elements of the value-added chain.
A virtual organization is not a viable organization in an integral, non-modular world [Chesbrough and
Teece, 1996].

Frequently, a core competence of a company lies in its processes for achieving exceptional product
performance through the design of certain subsystems in its products whose components and materials
have an integral character. If the components in this critical subsystem become modularized, this
competence in integration essentially becomes embodied in the standard interfaces of the components.
By this mechanism, proprietary competencies of early technology leaders can become diffused throughout
an industry in the form of component interface standards. Hence, if a situation of complete product
modularity were to occur in an industry, no firm could possess proprietary competence in product design:
all competitors could mix and match components equally.4

We might expect, therefore, that, if an industry moved toward total modularity in its products,
innovators would seek to establish competitive advantage by creating and maintaining new, nonstandard
ways of integrating components in their products. As products became more highly integral in character,
we would expect cost- and time-pressured engineers and marketers to always be searching for ways to
modularize their products.

Defining Terms:

**Dominant design:** An explicit or *de facto* industry-wide standard architectural configuration of the
components in an assembled product, in which the ways in which components interface with
others in the product’s architecture is well understood and established.

**Modularization:** A process by which the way that components and subsystems within an assembled
product interact with each other becomes so well understood that standards emerge, defining how
each component must interface with others in the system. When these standard interfaces exist,
components and subsystems from multiple suppliers can be mixed and matched in designing and
assembling a product, with predictable results for final system performance.

4Although I know of no studies that measure this phenomenon directly, I suspect that the industry of designing
and assembling personal computers was very nearly in this situation in the early and mid-1990s. The components
from which they were built interfaced with each other according to such well-established standards that it was difficult
for any manufacturer to sustainably assert that they offered proprietary cost-performance advantages in their products.
Punctuated equilibrium: A model of progress in which most of an industry's history is characterized by relatively steady, incremental, predictable improvement. This predictability is occasionally interrupted, or “punctuated”, by brief, tumultuous periods of radical, transformational change.

S-curves: An empirical relationship between engineering effort and the degree of performance improvement achieved in a product or process. The improvement produced by an incremental unit of engineering effort typically follows an S-curve pattern.

References


Further Information

References for each of the concepts noted in this chapter are listed below, including the articles and/or books that summarize the most important aspects within each body of scholarship.


The Technology and Innovation Management (TIM) section of the Academy of Management is an association of academics and managers whose research and practice focuses on issues of managing innovation. The activities of this organization, as well as the names of leading members of it, can be found on their web page, at http://www.aom.pace.edu/tim/.