



5

Resource-Based Theory, Dynamic Capabilities, and Real Options

Although early contributions to resource-based theory and dynamic capabilities came from the discipline of economics (e.g., Demsetz, 1973; Gort, 1962; Marris, 1964; Penrose, 1959; Richardson, 1960, 1972; Rubin, 1973; Slater, 1980), during the last 20 years the business field of strategic management has made significant contributions to resource-based theory and dynamic capabilities (e.g., Foss, 1997; Heene & Sanchez, 1997; Volberda & Elfring, 2001). Logic dictates that (organizational) economic theory will continue to play an important role in the study of economic value creation and sustainable competitive advantage. After all, sustainable competitive advantage requires an understanding of market frictions, and there is a large and well-developed economics research literature on market failures that students studying the economics of organization can draw on.

Although the market-failures literature is well developed, the organizational-failures literature is comparatively less developed, thereby providing research opportunities for students studying the economics of organization. Furthermore, resource-based theory and dynamic capabilities and real options research may develop into a paradigmatic approach to strategic management, an important contribution to the evolving science of organization. Clearly, there is a need for rigorous empirical research to establish both the nature and the impact of dynamic capabilities on sustainable competitive advantage. Capabilities that can prove especially useful in dynamic business environments are operational and strategic flexibility.

I begin this chapter with the seminal work of Penrose (1959), who provides (1) a general theory of the growth of the firm, (2) a theory of entrepreneurship based on the subjective opportunity set of the firm, (3) expansion based on indivisibility and the balance of processes, (4) a

resource-based theory of diversification, and (5) a theory of expansion through acquisition and merger. In addition, Penrose provides a theory of the limits to the rate of the growth of the firm, in particular, arguing that the binding constraint on the firm's rate of the growth is provided by the capacities of its existing management—the so-called Penrose effect.

Chandler's (1990) *Scale and Scope* represents a culmination of Chandler's long quest to chart the evolution of modern industrial enterprise. The book provides the reader with an extraordinary breadth and depth of knowledge concerning the development of managerial capitalism. The essence of successful firm strategy, Chandler argues, is the making of three interrelated investments: (1) investment in production to achieve the cost advantages of scale and scope; (2) investment in product-specific marketing, distribution, and purchasing networks; and (3) investment in managerial talent and management structure to plan, coordinate, and monitor the firm's often dispersed operations. Chandler argues that such three-pronged investment enables firms to develop organizational capabilities, which, in turn, provides the dynamic for the continuing growth of the enterprise.

Itami and Roehl (1987) emphasize the dynamic fit between resources and the environment. Itami and Roehl build on the work of Penrose (1959) concerning corporate growth and move the arguments forward by emphasizing the role of invisible assets of a firm, which are based on information. Invisible assets include intellectual property rights of patents and trademarks, trade secrets, proprietary data files, personal and organizational networks, reputation, and culture. Itami and Roehl argue that these invisible assets are often a firm's only real source of sustainable competitive advantage.

Nelson and Winter (1982) consider the promise and the problems of evolutionary modeling of economic change. Among the many benefits that may be derived from Nelson and Winter's theoretical approach that reconciles economic analysis with real-world business firm decision making, the most important relate to improved understanding of technological change and the dynamics of the competitive process. Nelson and Winter's evolutionary theory is intrinsically dynamic, in which the heterogeneity of firms is a key feature.

This chapter on dynamic resource-based theory concludes with a research book that some in the strategic management field may find to be a curious choice. Over time, however, I anticipate that it will become abundantly clear that a key category in developing dynamic capabilities involves strategies that enhance adaptability and strategic flexibility.

Viewed in this light, Trigeorgis's (1996) research book is an important contribution to the dynamic capabilities research literature. Real options research has the potential to make a significant difference to our understanding of resource accumulation and capability-building processes and investment decision making under uncertainty. Finally, supplementing the real options analysis with game-theoretic tools that capture competitive dynamics is promising for future research by students pursuing the evolving science of organization.

The Theory of the Growth of the Firm (*Penrose, 1959*)

Penrose (1959) is concerned with the growth of firms and only incidentally with the size of the firm. Penrose argues that firm size is only a by-product of the process of growth and that there is no optimum, or even most profitable, size of the firm. Penrose is primarily concerned with a theoretical analysis of the growth process of the firm.

Penrose (1959) emphasizes the internal resources of a firm—on the productive services available to a firm from its own resources, particularly the productive services available from management with experience within the firm. The (firm-specific) experience of management affects the productive services that all its other resources are capable of rendering. As management tries to make the best use of the available resources, a dynamic interacting process occurs that encourages a continuous, but limited, rate of growth of the firm. To focus attention on the crucial role of the firm's inherited resources, the environment is treated, in the first instance, as an image in the entrepreneur's mind of the possibilities and restrictions with which it is confronted. For it is, after all, such an image that in fact determines a person's behavior. Whether experience confirms expectations is another story.

The Firm in Theory. Penrose (1959) notes that in a private enterprise industrial economy the business firm is the basic unit for the organization of production. Because of its complexity and diversity, a firm can be approached with many different types of analysis—sociological, organizational, engineering, or economic—and from whatever point of view that seems appropriate to the business problem at hand.

The theory of the firm—as it is called in the neoclassical economics literature—was constructed for the purpose of assisting in the theoretical investigation of one of the central problems of economic analysis—the way in which prices and the allocation of resources among different

uses are determined. It is but part of the wider theory of economic value. The equilibrium of the firm is, in essence, the equilibrium output. As Boulding (1950) notes, "The firm is a strange bloodless creature without a balance sheet, without any visible capital structure, without debts, and engaged apparently in the simultaneous purchase of inputs and the sale of outputs at constant rates" (p. 24).

Penrose (1959) points out that if we become interested in other aspects of the firm, then we ask questions that the theory of the firm is not designed to answer. Penrose wants to deal with the firm as a growing organization, not as a price-and-output decision maker for given products. Penrose argues that the essential difference between economic activity inside the firm and economic activity in the market is that economic activity in the firm is carried out within an administrative organization (see Simon, 1947), whereas economic activity in the market is not. Penrose refers to this court of last resort in the firm as central management. It is the area of coordination—the area of authoritative communication (Barnard, 1938; Simon 1947)—which defines the boundaries of the firm, and, consequently, it is a firm's ability to maintain sufficient administrative coordination to satisfy the definition of an industrial firm that sets the limit to its size as an industrial enterprise.

The Firm as a Collection of Productive Resources. According to Penrose (1959), a firm is more than an administrative unit; a firm is also a collection of productive resources, where the choice of different uses of these resources over time is determined by administrative decision. The physical resources of a firm consist of tangible things—plant, equipment, land, and natural resources; raw materials; semifinished goods; waste products and by-products; and even unsold stocks of finished goods. There are also human resources available in a firm—unskilled and skilled labor and clerical, administrative, financial, legal, technical, and managerial staff.

Penrose (1959) argues that, strictly speaking, it is never resources themselves, but only the services that the resource can render, that are the inputs in the production process. Resources consist of a bundle of potential services and can, for the most part, be defined independently of their use, whereas services cannot be so defined, the very word *service* implying a function, an activity. *It is largely in this distinction that we find the source of the uniqueness of each individual firm.*

The business firm, as Penrose (1959) defines it, is both an administrative organization and a collection of productive resources. The general

purpose of the business firm is to organize the use of its own resources together with other resources acquired from outside the firm for the production and sale of goods and services at a profit.

The term *entrepreneur* refers to individuals or groups within the firm providing the entrepreneurial services, whatever their position or occupational classification may be. Entrepreneurial services are those contributions to the operations of a firm that relate to the introduction and acceptance on behalf of the firm of new ideas, particularly with respect to products, location, and significant changes in technology; to the acquisition of new managerial personnel; to fundamental changes in the organization of the firm; to the raising of capital; and to the making of plans for expansion, including the strategic choice of expansion method.

Penrose (1959) submits that a versatile type of executive service is needed if expansion requires major efforts on the part of the firm to develop new markets or entails branching out into new lines of production. Here, the imaginative effort, the sense of timing, the instinctive recognition of what will catch on or how to make it catch on, become of overwhelming importance. These services are not likely to be equally available to all firms. Firms not only alter the environmental conditions necessary for the success of their actions, but, even more important, they also know that they can alter these conditions and that the environment is not independent of their own activities.

Expansion Without Merger: The Receding Managerial Limit. Penrose (1959) notes three classes of explanation for why there may be a limit to the growth of firms—managerial ability, product or factor markets, and uncertainty and risk. The first explanation refers to conditions within the firm, the second explanation refers to conditions outside the firm, and the third explanation is a combination of internal attitudes and external conditions. The capacities of the existing managerial personnel of the firm necessarily set a limit to the expansion of that firm in any given period of time, and such management possessing firm-specific abilities cannot be hired in the marketplace.

Penrose (1959) argues that an administrative group is something more than a collection of individuals; an administrative group is a collection of individuals who have had experience in working together—only in working together can teamwork be developed. Experiences these individuals gain from working within the firm, and with each other, enable them to provide services that are uniquely valuable for the operations of the particular group with which they are associated. Existing management limits the amount of new management that can be hired

at any point in time (after all, the services of existing management are required to instruct the new personnel).

Penrose (1959) submits that if a firm expands its organization more rapidly than the individuals in the expanding organization can obtain the experience with each other they need for effective operation of the group, the efficiency of the firm will suffer. Because the services from current managerial resources control the amount of new managerial resources that can be absorbed, they create a fundamental limit to the amount of expansion a firm can undertake at any point in time. The amount of activity that can be planned at a given time period limits the amount of new personnel that can be profitably absorbed in the next period. This idea over subsequent years is known as the *Penrose effect*. Moreover, as plans are completed and put into operation, managerial services absorbed in the planning processes will be gradually released and become available for further planning.

Penrose (1959) argues that knowledge comes to people in two different ways: It can be formally taught, and it can be achieved via learning-by-doing in the form of personal experience. Experience produces increased knowledge and contributes to objective knowledge insofar as its results can be transmitted to others. But experience itself can never be transmitted; experience produces a change—frequently a subtle change—in individuals and cannot be separated from them.

Increasing experience shows itself in two ways—changes in knowledge acquired and changes in the ability to use knowledge. There is no sharp distinction between these two forms because to a considerable extent the ability to use old knowledge is dependent on the acquisition of new knowledge. But it is not exclusively so dependent; with experience a person may gain wisdom, sureness of movement, and confidence—all of these become part of his or her very nature, and they are all qualities that are relevant to the kind and amount of services a person can give to the firm. Much of the experiences of business personnel are frequently so closely associated with a particular set of external circumstances that a large part of a personnel's most valuable services may be available only under these circumstances.

A person whose past productive activity has been spent within a particular firm, for example, can, because of his or her intimate knowledge of the resources, structure, history, operations, and personnel of the firm, render services to that firm that person could give to no other firm without acquiring additional experience. Penrose (1959) submits that, after it is recognized that the very processes of operating and of

expanding are intimately associated with a process by which knowledge is increased, it becomes clear that the productive opportunity of a firm will change, even in the absence of any change in external circumstances or in fundamental technological knowledge. New opportunities open up that did not exist at the time expansion plans were made. That is, the subjective opportunity set of the firm will change.

Penrose (1959) hastens to add that management not only is the source for expansion but also is a brake on expansion. A firm has a given amount of experienced managerial services available at any one time. Parts of these managerial services are needed for ordinary operation, and the rest of these managerial services are available for planning and executing expansion programs. The effect of uncertainty is to require that some of these available services be used to gather information, process the information, and reach conclusions about the possibilities of action in which the management team has confidence.

Inherited Resources and the Direction of Expansion. Penrose (1959) maintains that the external inducements to expansion include growing demand for particular products, changes in technology that call for production on a larger scale than before, discoveries and inventions with particularly promising uses, and opportunities to obtain a better market position. Inducements to expansion also include backward integration to control sources of supply, diversification of final products to spread risk, or expansion of existing or allied products to preclude the entry of new competitors. External obstacles to expansion include keen competition in markets for particular products that makes profitable entry or expansion in those markets difficult.

Penrose (1959) argues that whereas external inducements and obstacles have been widely discussed, little attention has been paid, in a systematic way at least, to the equally important internal influences on the direction of expansion. Internal obstacles arise when some of the important types of specialized services required for expansion in particular directions are not available in sufficient amounts within the firm. In particular, internal obstacles arise when not enough of the managerial capacity and the technical skills required for the planning, execution, and efficient operation of a new program can be obtained from among existing experienced personnel. Internal inducements to expansion arise largely from the existence of a pool of unused productive services, resources, and specialized knowledge, all of which will always be found within any firm. To Penrose, a resource can be viewed as a bundle of possible services. As long as resources are not used fully

in current operations, there is an economic incentive for a firm to find a way of using them more fully.

Penrose (1959) maintains that three significant obstacles preclude the attainment of a state of rest:

1. Those arising from the familiar difficulties posed by the indivisibility of resources, which Penrose (1959) calls “the jig-saw puzzle” (p. 69)
2. Those arising from the fact that the same resources can be used differently under different circumstances
3. Those arising because, in the ordinary processes of operation and expansion, new productive services are continually being created

Penrose (1959) then discusses how the division of labor (specialization) can lead to the growth of the firm and diversification. This process has been called the *virtuous circle*, in which specialization leads to higher common multiples and higher common multiples lead to greater specialization. Penrose also argues that diversification strategy can be driven by the desire to achieve multiproduct economies of scale (which in modern strategic management language is called *economies of scope*; Teece, 1980).

Penrose (1959) observes that for many purposes it is possible to deal with rather broad categories of resources, overlooking the lack of homogeneity in the members of the category. Economists usually recognize this simplification, stating that for convenience alone resources are grouped under a few headings—for example, land, labor, and capital—but Penrose (1959) points out that the subdivision of resources may proceed as far as is useful, according to whatever principles are most applicable for the business problem at hand. The heterogeneity of the productive services available or potentially available from its resources gives each firm its unique character.

Furthermore, the possibilities of using services change with changes in knowledge. Consequently, there is a close connection between the type of knowledge possessed by the personnel of the firm and the services obtainable from its material resources. The firm, then, is viewed as a collection of resources. Unused productive services shape the scope and direction of the search for knowledge. Knowledge and an economic incentive to search for new knowledge are built into the very nature of firms possessing entrepreneurial resources of even average initiative. Physically describable resources are purchased in the strategic factor markets for their known services, but, as soon as these resources

become part of a firm, the range of services they are capable of yielding starts to change. The services that resources yield depends on the capacities of the people using them, but the development of the capacities of people is partly shaped by the resources they deal with. The two together create the distinctive, subjective, productive opportunity set of a particular firm.

If resources were completely nonspecific, a firm could in principle produce anything. The selection of the relevant product markets is necessarily determined by the inherited resources of the firm—the productive services it already has. To be sure, the anticipation of consumer acceptance is a necessary condition of entrepreneurial interest in any product, but the original economic incentive to a great deal of innovation can be found in a firm's desire to use its existing resources more efficiently. There is a close relation between the various kinds of resources with which a firm works and the development of ideas, experience, and knowledge of its managers and entrepreneurs. Changing experience and knowledge affect not only the productive services available from resources but also the demand as seen by the firm.

Penrose (1959) further elaborates, noting that unused productive services are, for the enterprising firm, at the same time a challenge to innovate, an economic incentive to expand, and a source of sustainable competitive advantage. Unused productive services facilitate the introduction of new combinations of resources—innovations—within the firm. Unused productive services are a selective force in determining the direction of expansion. Therefore, analysis is required of internal and external inducements and internal and external obstacles for expansion.

The Economies of Diversification. Penrose (1959) argues that of all the outstanding characteristics of business firms, perhaps the most inadequately treated in economic analysis is the diversification of their activities. *Anticipating Teece (1982), Penrose argues that market imperfections are an important explanation of diversification strategy.* Diversification that involves a departure from the firm's existing areas may be one of three kinds:

- The entry into new markets with new products using the same production base
- Expansion in the same market with new products based on a different area of technology
- Entry into new markets based on a different area of technology

Penrose (1959) observes that a firm's opportunities are necessarily widened when the firm develops a specialized knowledge of a technology that is not very specific to any particular kind of product, for example, knowledge of different types of engineering or industrial chemistry. Diversification and expansion based primarily on a high degree of competence and technical knowledge in specialized areas of manufacture are characteristic of many of the largest firms in the economy. Penrose submits that this type of competence, together with the market position such competence and technical knowledge ensures, is the strongest and most enduring position a firm can develop.

Diversification through both internal and external expansion is likely to be extensive because of the wide variety of productive services generated within such firms and because the competitive advantages these firms possess are conducive to expansion. Opportunities for expansion both within existing resource bases and through the establishment of new resource bases are likely to be so prevalent that the firm has to choose carefully among many different courses of action.

The Firm as a Pool of Resources. Penrose's (1959) thesis is that a firm is essentially a pool of resources, the use of which is organized in an administrative framework. In a sense, the final products being produced by a firm at any given time represent one of several ways in which the firm could be using its resources, an incident in the development of its basic potentialities. The continual change in the productive services and knowledge within a firm and the continual change in external circumstances present the firm with a continually changing productive opportunity set.

The Rate of Growth of a Firm Through Time. Penrose (1959) notes that markets and firms are interacting institutions, each being necessary to the existence of the other. Penrose emphasizes that one of the more significant characteristics of entrepreneurial and managerial services is their heterogeneity, their uniqueness for every individual firm. The factors determining the availability of managerial services and the need for these services in expansion determine the maximum rate of growth of the firm, where rate of growth is defined as the percentage rate at which the size of the firm increases per unit of time. The possibility of acquiring other firms raises enormously the maximum rate of expansion, primarily because acquisition substantially reduces the managerial services required per unit of expansion.

Concluding Comments. Kor and Mahoney (2000) suggest the following list of key ideas derived from Penrose (1959):

- Firm growth can be usefully studied as a dynamic process of management interacting with resources.
- Firms are institutions created by people to serve the purposes of people.
- Services of resources are drivers of firm heterogeneity.
- Services that material resources will yield depend on the knowledge possessed by human resources. The two together create a subjective opportunity set that is unique for each firm.
- Firm growth is a function of firm-specific experiences in teams.
- Managerial capability is the binding constraint that limits the growth rate of the firm—the so-called Penrose effect.
- Excess capacity of productive services of resources is a driver of firm growth.
- Unused productive services of resources can be a source of innovation.
- Firm diversification is often based on a firm's competencies that can lead to a sustainable competitive advantage.
- An important component of the competitive process is experimentation.

Finally, it is noted that some criticize Penrose's resources approach for ignoring the business environment. Penrose (1959), in fact, addresses this issue, arguing that whether or not we treat the resources of the firm or its environment as the more important factor explaining growth depends on the question we ask: If we want to explain why different firms see the same environment differently, why some firms grow and some do not, or, to put it differently, why the environment is different for every firm, we must take the resources approach; if we want to explain why a particular firm or group of firms with specified resources grows the way it does, we must examine the opportunities for the use of those resources. Penrose calls these opportunities for small firms the *interstices* in the economy. The productive opportunities of small firms are thus composed of those interstices left open by the large firms that the small firms see and believe they can take advantage of. Penrose concludes that management's experiments with different types of corporate structures are in themselves an important aspect of competition.

In my judgment, Penrose (1959) is the seminal work in resource-based theory that anticipates the works of Chandler (1962, 1990), which document organizational innovations and organizational capabilities

that, in turn, provide an internal dynamic for the continuing growth of the modern industrial enterprise.¹

Scale and Scope: The Dynamics of Capitalism (*Chandler, 1990*)

I turn now to Chandler (1990), who provides a detailed but highly generalized description and analysis of the creation and dynamic evolution of the central institution of managerial capitalism—the modern industrial enterprise. These concepts and generalizations are then used to develop an explanatory theory of the evolution of the modern industrial enterprise. The richness of information provided in this research book can be helpful for students in the evolving science of organization in answering questions that have long concerned economists and business historians—questions about changes in internal organization and management, competition and cooperation among firms, growth through horizontal acquisitions and mergers, vertical integration, expansion into international markets, diversification into new product lines, and the effect of legal requirements, government rulings, and cultural values on firm growth and economic performance.

Chandler (1990) observes that in the last half of the 19th century a new form of capitalism appeared in the United States and Europe. The building and operating of rail and telegraph systems called for the creation of a new type of business enterprise. The massive investment required in constructing those systems, and the complexities of their operations, brought the separation of ownership and management. The

¹For further readings see Penrose (1955, 1960). Penrose (1960) provides a case study of the Hercules Powder Company to illustrate that growth is governed by a creative and dynamic interaction between a firm's productive resources and its market opportunities. Richardson (1972), Rubin (1973), Slater (1980), and Teece (1982) are influential journal articles in the economics research literature that build on Penrose (1959). For a recent assessment of Penrose (1959), see Kor and Mahoney (2000), which focuses on (1) the research process that led to Penrose's (1959) classic, (2) Penrose's (1959) contributions to the discipline of strategic management, (3) the generative nature of Penrose's (1959) research for current resource-based theory, and (4) suggested future research building on Penrose's (1959) resource approach. In addition, Kor and Mahoney (2004) and Rugman and Verbeke (2002) consider Penrose's (1959) direct and indirect contributions to the modern resource-based view within strategic management. Pettus (2001) studies the Penrose effect in the deregulated trucking industry, whereas Tan (2003) and Tan and Mahoney (2003) provide empirical tests of Penrose (1959) in the context of multinational firms. Finally, Pitelis (2002) provides an excellent edited collection of recent writings that document the legacy of Penrose (1959) on contemporary research on the growth of the firm.

enlarged enterprises came to be operated by teams of salaried managers who had little or no equity in the firm.

Chandler (1990) notes that the new forms of transportation and communication, in turn, permitted the rise of modern mass marketing and modern mass production. The unprecedented increase in the volume of production, and in the number of transactions, led the entrepreneurs who established the new mass-producing and mass-distributing enterprises—like the railroad personnel before them—to recruit teams of salaried managers.

Chandler (1990) examines the beginning and growth of global managerial capitalism, focusing on the history of its basic institution—the modern industrialized enterprise—in the world's three leading industrial nations. They each had been rural, agrarian, and commercial and each became industrial and urban, a transformation that, in turn, brought the most rapid economic growth in business history. At the center of the transformation were the United States, Great Britain, and Germany, which together accounted for just over two thirds of the world's industrial output in 1870. The industrial sector grew significantly in the United States and Germany; in Great Britain the development was slower but sustained. Further, whereas Great Britain experienced only a moderate change of employment structure after the 1880s, the United States—and Germany to a lesser degree—showed a dramatic transformation from an agrarian to a modern economy, in which almost half of the country's employment centered in industry.

Chandler (1990) maintains that as a result of the regularity, increased volume, and greater speed of the flows of goods and materials made possible by the new transportation and communication systems, new and improved processes of production developed that for the first time in business history achieved substantial economies of scale and scope. Large manufacturing works applying the new technologies could produce at lower costs than could the smaller manufacturing works.

Chandler (1990) observes that for entrepreneurs to benefit from the cost advantages of these new, high-volume technologies of production, the entrepreneurs had to make three sets of interrelated investments:

1. An investment in production facilities large enough to use a technology's potential economies of scale and scope
2. An investment in a national and international marketing and distribution network so that the volume of sales might keep pace with the new volume of production

3. Investment in management, which required entrepreneurs to recruit and train managers not only to administer the enlarged facilities and increased personnel in both production and distribution but also to monitor and coordinate those two basic functional activities and to plan and allocate resources for future production and distribution

Chandler (1990) submits that *this three-pronged investment in production, distribution, and management brought the modern industrial enterprise into being*. The first entrepreneurs to create such enterprises acquired substantive competitive advantages. Their industries quickly became oligopolistic—that is, dominated by a small number of first movers. These first-mover firms, along with a few challengers that subsequently entered the industry, no longer competed primarily on the basis of price. Instead, these firms competed through functional and strategic effectiveness. These firms did so functionally by improving their product, marketing, purchasing, and labor relations, and these firms did so strategically by moving into growing markets more rapidly and by divesting out of declining markets more quickly and effectively than did their competitors.

Such rivalry for market share and profitability honed the enterprise's functional and strategic capabilities. These organizational capabilities, in turn, provided an internal dynamic for the continuing growth of the enterprise. In particular, these organizational capabilities stimulated its owners and managers to expand into more distant markets in their own country and then to become multinational by moving abroad. These organizational capabilities also encouraged the firm to diversify by developing products competitive in markets other than the firm's original market, becoming a multiproduct enterprise.

Scale, Scope, and Organizational Capabilities. Chandler (1990) argues that the modern industrial enterprise can be defined as a collection of operating units, each with its own specific facilities and personnel, whose combined resources and activities are coordinated, monitored, and allocated by a hierarchy of middle and top managers. This hierarchy makes the activities and operations of the whole enterprise more than the sum of its operating units. The manufacturing enterprises became multifunctional, multiregional, and multiproduct because the addition of new units permitted these enterprises to maintain a long-term rate of return on investment by reducing overall costs of production and distribution, by providing products that satisfied existing demands, and by transferring facilities and capabilities to more profitable markets when economic returns were reduced by competition, changing technology, or altered market demand.

Chandler (1990) submits that whatever the initial motivation for its investment in new operating units, the modern industrial enterprise has rarely continued to grow or maintain its competitive position over an extended period of time unless the addition of new units (and to a lesser extent the elimination of old units) has actually permitted the visible hand of its managerial hierarchy to reduce costs, to improve functional efficiency in marketing and purchasing as well as production, to improve existing products and processes and to develop new ones, and to allocate resources to meet the challenges and opportunities of ever-changing technologies and markets. Such a process of growth has provided organizations with the internal dynamic that has enabled them to maintain their position of dominance as markets and technologies have changed. Chandler further argues that it was the development of new technologies and the opening of new markets that resulted in economies of scale and scope and reduced transaction costs, that made the large, multiunit enterprise come when it did, where it did, and in the way it did.

Chandler (1990) maintains that coordination demanded the constant attention of a managerial team or hierarchy. The potential economies of scale and scope are a function of the physical characteristics of the production facilities. However, the actual economies of scale and scope, as measured by throughput, are a function of organizational capabilities. The full fruition of economies of scale and scope depend on knowledge, skill, experience, and teamwork—on the organizational capabilities essential to use the full potential of technological processes. Further, in many instances, Chandler finds that the first company to build a plant of minimum efficient scale, and to recruit the essential management team to enable the enterprise to reach its full potential, often remained the leader in its industry for decades.

Chandler (1990) indicates that organizational capabilities included those of lower management and the workforce, in addition to the skills of middle and top management. Organizational capabilities also included the facilities for production and distribution acquired to use fully the economies of scale and scope. Such organization capabilities provided the economic profits that, in large part, financed the continuing growth of the enterprise. Highly product specific and process specific, these organizational capabilities affected—indeed, often determined—the direction and pace of first movers and challengers and of the industries and even the national economies in which they operated (Collis, 1994).

Chandler (1990) emphasizes that only if these facilities and organizational capabilities were carefully coordinated and integrated could the enterprise achieve the economies of scale and scope that were needed to

compete in national and international markets and to continue to grow. Middle managers not only had to develop and apply functional-specific and product-specific managerial skills, but they also had to train and motivate lower level managers and to coordinate, integrate, and evaluate their work. Such organizational capabilities, of course, had to be created, and, once established, these capabilities had to be maintained. Their maintenance was as great a challenge as their creation because facilities depreciate, individual skills atrophy, and organizational capabilities can diminish. Moreover, changing technologies and markets constantly make existing facilities, individual skills, and organizational capabilities obsolete. One of the more critical tasks of the top management team has always been to maintain these organizational capabilities and to integrate these facilities and skills into a coherent, unified organization—so that the whole becomes more than the sum of its parts. Such organizational capabilities, in turn, have provided the source—the dynamic—for the continuing growth of the enterprise. Organizational capabilities have made possible the earnings that supplied much of the funding for such growth.

As Chandler (1990, 1992) repeatedly emphasizes, in the collective individual industries that are so aptly documented, the first movers' initial, interrelated, three-pronged investments in manufacturing, marketing, and management created powerful barriers to entry (see also Porter, 1980). Challengers had to make comparable (sunk-cost) investments at a greater risk, precisely because the first movers had already learned the ways of the new processes of production, were already dominating the markets for the new or greatly improved products, and were already reaping substantial economic returns from their initial investments. As the first movers' functional and organizational capabilities were honed, the difficulties of entry by newcomers became even more formidable. In the sale of consumer products, particularly branded, packaged goods, barriers to entry were reinforced by advertising, vertical tying contracts, and exclusive franchising. In the more technologically advanced producer-goods industries, patents reinforced these entry barriers. In Europe, first movers strengthened their strategic positioning still further by arranging interfirm agreements as to price, output, and market territories.

A New Era of Managerial Capitalism? Chandler (1990) notes that the historian who has studied the past experience of the business enterprise is in a better position than most analysts to identify current business practices that are truly new. Chandler observes that of the many recent

changes in the growth, management, and financing of the modern industrial enterprise, the following six have no historical precedents:

- The adaptation of a new corporate strategy of growth—that of moving into new markets where the organizational capabilities of the enterprise do not provide competitive advantages
- The separation of top management in the corporate office from middle management in the operating divisions
- The extensive and continuing divestiture of operating units
- The buying and selling of corporations as a distinct business in its own right
- The role played by portfolio managers in the capital markets
- The evolution of those capital markets to facilitate the coming of what has been termed a market for corporate control

Chandler (1990) concludes that his research book has only begun to map the evolution of the industrial enterprise in the United States, Great Britain, and Germany from the 1880s to the 1940s. Valid description and analysis on which generalizations can be made must await an in-depth, industry-by-industry, country-by-country historical study.² Much more work needs to be done that certainly may modify the patterns of institutional change that Chandler has outlined. Clearly, there are research opportunities for those students studying the economics of organization who combine the craft of the business historian and the analytical skills derived from the resource-based/dynamic capabilities perspective. Indeed, Chandler provides insights that connect (company-specific) organizational capabilities and the economics of organization.

Mobilizing Invisible Assets (*Itami & Roehl, 1987*)

To develop further the dynamic capabilities perspective, I turn now to Itami and Roehl's (1987) contribution to dynamic capabilities theory. Itami and Roehl emphasize the role of environmental, corporate, and internal information flow. Environmental information flow includes discovering customer preferences and maintaining competitor intelligence.

²Resource-based/dynamic capabilities theory has recently been empirically corroborated in the context of international business studies (e.g., Anand & Delios, 2002; Anand & Singh, 1997). Peng (2001) documents the extent to which resource-based theory has diffused in international business research.

Corporate information flow includes reputation, brand image, and marketing know-how. Internal information flow includes corporate culture and managerial capabilities (e.g., routines).

Itami and Roehl (1987) provide a strategic logic that is heavily influenced by Penrose (1959) and emphasizes the vital contribution of accumulated experience and information to a corporation's strategic resources. Itami and Roehl emphasize that the intangible assets, such as a particular technology, accumulated consumer information, brand name, reputation, and corporate culture, are invaluable to the firm's competitive advantage. In fact, these invisible resources are often a firm's only real source of competitive edge that can be sustained over time.

Itami and Roehl (1987) emphasize that current strategy, because it can change the level of invisible assets, is more than the basis for short-term competitive advantage; current strategy provides the foundation for future strategy and adds to or erodes the invisible resource base. The competitive success of a strategy is dependent on the firm's invisible assets, but the dynamics of invisible assets (their accumulation and depreciation over time) is also largely determined by the content of that strategy. Itami and Roehl explore how invisible assets affect, and are affected by, the firm's strategy. Decisions made today can affect a firm's long-term capabilities and adaptability because such decisions often determine the accumulation of invisible assets.

Itami and Roehl (1987) maintain that many invisible resources are quite fixed. There is no easy way to obtain a well-known brand name or advanced technical production skills in the market, nor can money buy an instantaneous change in corporate culture and employee morale. Accumulation of these invisible resources requires ongoing, conscious, and time-consuming efforts; you cannot just go out and buy these resources off the shelf. For this reason, a firm can differentiate itself from competitors through its invisible resources. If a resource can be bought, competitors with sufficient financial resources can gain access to it. And if a resource can be created quickly, competitors will have ready access to such a resource through imitation. But competitors cannot do this easily with invisible resources.

The important features of invisible resources are as follows:

- Unattainable with money alone
- Time-consuming to develop
- Capable of multiple simultaneous uses
- Able to yield multiple, simultaneous benefits

These features of invisible resources make it crucial to consider carefully the strategies for accumulating invisible resources.

Information is at the heart of invisible resources. Information-based invisible resources include not only the stock of accumulated information in the firm but also the channels that handle the flow of information of importance to the firm. Information can be classified as environmental, corporate, or internal.

Environmental information flows from the environment to the firm, creating invisible assets related to the environment. This type of information flow includes production skills, customer information, and channels for bringing in information.

Corporate information flows from the firm to the environment, creating invisible assets stored in the environment. This category of information flow includes such invisible assets as corporate reputation, brand image, corporate image, and influence over the distribution and its parts suppliers, as well as marketing know-how.

Internal information originates and terminates within the firm, again affecting the invisible asset stock. This category of information flow includes corporate culture, morale of workers, and management capabilities, as well as the firm's ability to manage information, the employees' ability to transmit and use the information in decision making, and the employees' habits and norms of effort expended. *Successful accumulation of invisible resources comes down to control of the information flow.*

In my judgment, Itami and Roehl (1987) is a seminal contribution to resource-based theory and the dynamic capability approach. Invisible assets serve as a focal point of strategy development and growth. Students studying the economics of organization are served well in examining closely this often overlooked classic.

I turn next to a classic that almost everyone recognizes as the seminal and path-breaking book on evolutionary economics and dynamic capabilities.

An Evolutionary Theory of Economic Change (*Nelson & Winter, 1982*)

Nelson and Winter (1982) provide a wealth of strategic issues for consideration by current students who want to contribute to the evolving science of organization. Nelson and Winter provide the organization-theoretic foundations of economic evolutionary theory, the building blocks of which include individual skills and organizational capabilities. Nelson

and Winter develop an evolutionary model of economic growth and a perspective that emphasizes the role of Schumpeterian competition.

Nelson and Winter (1982) argue that much of firm behavior can be more readily understood as a reflection of general routines and strategic orientations coming from the firm's past than as the result of a detailed survey of the remote twigs of a decision tree extending into the future. Nelson and Winter acknowledge their intellectual debts to Joseph Schumpeter and Herbert Simon. Schumpeter (1934, 1950) points out the right problem—how to understand economic change—and many of the important elements of the answer. Simon (1982) provides a number of specific insights into human and organizational behavior that are reflected in Nelson and Winter's theoretical models. But, most important, Simon's (1947, 1982) works encourage Nelson and Winter in maintaining the view that there is much more to be said on the problem of rational behavior in the world of experience than can be adequately stated in the language of orthodox economic theory.

Nelson and Winter (1982) develop an evolutionary theory of the organizational capabilities and behaviors of business firms operating in a market environment. The firms in their evolutionary theory are treated as motivated by profitability and engaged in the search for ways to improve their profitability, but the firm's actions are not assumed to be profit maximizing over well-defined and exogenously given choice sets. Evolutionary theory emphasizes the tendency for the more profitable firms to drive the less profitable firms out of business. However, Nelson and Winter do not focus their analysis on hypothetical states of industry equilibrium, in which all unprofitable firms are no longer in the industry and profitable firms are at their desired size.

Relatedly, the modeling approach employed in Nelson and Winter (1982) does not use the familiar maximization calculus to derive equations characterizing the behavior of firms. Rather, firms are modeled as having, at any given time, certain organizational capabilities and decision rules. Over time, these organizational capabilities and decision rules are modified as a result of both deliberate problem-solving efforts and random events. And over time, the economic analogue of natural selection operates as the market determines which firms are and are not profitable and winnows out the unprofitable firms. Supporting Nelson and Winter's analytical emphasis on this sort of evolution by natural selection is a view of organizational genetics—the processes by which traits of organizations, including those traits underlying the capability to produce output and to make profits, are transmitted through time.

Nelson and Winter (1982) give attention to uncertainty, bounded rationality, the presence of large corporations, institutional complexity, and the dynamics of the actual adjustment process. Considerable attention is also given to imperfect information and imperfect competition, transaction costs, indivisibilities, increasing returns, and historical change.

Although Nelson and Winter (1982) stress the importance of certain elements of continuity in the economic process, they do not deny (nor does contemporary biology deny) that change is sometimes rapid. Also, some people who are particularly alert to teleological fallacies in the interpretation of biological evolution seem to insist on a sharp distinction between explanations that feature the processes of blind evolution and those that feature deliberate goal seeking. Whatever the merits of this distinction in the context biological evolution theory, such a distinction is unhelpful and distracting in the context of Nelson and Winter's theory of the business firm. It is neither difficult nor implausible to develop models of firm behavior that interweave blind and deliberate processes. Indeed, in human problem solving itself, both elements are involved and difficult to disentangle. Relatedly, Nelson and Winter describe their theory as unabashedly Lamarckian: The evolutionary economics theory of the firm contemplates both the inheritance of acquired characteristics and the timely appearance of variations under the stimulus of adversity.

Nelson and Winter's (1982) general term for all regular and predictable behavioral patterns of firms is *routine*. Nelson and Winter use this general term to include characteristics of firms that range from well-specified technical routines for producing things to procedures for hiring and firing, ordering new inventory, or stepping up production of items in high demand to policies regarding investment, research and development (R&D), or advertising to business strategies about product diversification and overseas investment. In Nelson and Winter's evolutionary theory, these routines play the role that genes play in biological evolutionary theory. They are a persistent feature of the organism and determine its possible behavior (though actual behavior is determined also by the environment).

Most of what is regular and predictable about business behavior is plausibly subsumed under the heading *routine*. That not all business behavior follows regular and predictable patterns is accommodated in evolutionary theory by recognizing that there are stochastic elements both in the determination of decisions and of decision outcomes. From

the perspective of a participant in business decision making, these stochastic elements may reflect the result of tumultuous meetings or of confrontations with complex problems under crisis conditions, but, from the viewpoint of an external observer seeking to understand the dynamics of the larger system, these phenomena are difficult to predict. Whereas in orthodox theory, decision rules are assumed to be the consequence of maximization, in evolutionary theory decision rules are treated as reflecting at any moment in time the historically given routines governing the actions of a business firm. Routine-changing processes are modeled as searches. Nelson and Winter's (1982) concept of search is the counterpart of that of mutation in biological evolutionary theory. Through the joint action of search and selection, the firms evolve over time, with the conditions of the industry in each period bearing the seeds of its condition in the following period.

Just as some orthodox microeconomic ideas seem to find their most natural mathematical expression in the calculus, the foregoing verbal account of economic evolution seems to translate naturally into a description of a Markov process—though one in a rather complicated state space. The process is not deterministic; search outcomes, in particular, are partly stochastic. Thus, what the industry condition of a particular period really determines is the probability distribution of its condition in the following period. Important antecedents of Nelson and Winter (1982) have been described in previous chapters:

- Behavioral theory of the firm (Cyert & March, 1963; Simon, 1947)
- Transaction costs theory (Williamson, 1975)
- Theory of the growth of the firm (Penrose, 1959)
- Business history (Chandler, 1962)

Chandler (1962) demonstrates that the organizational capabilities of a firm are embedded in its organizational structure, which is better adapted to certain strategies than to others. Thus, strategies at any point in time are constrained by the organization. Also, a significant change in a firm's strategy is likely to call for a significant change in its organizational structure.

Nelson and Winter (1982) build on the concept of Schumpeterian competition. Schumpeter's (1934) credentials as a theorist of bounded rationality could hardly be more incisively established than in the following passage:

The assumption that conduct is prompt and rational is in all cases a fiction. But it proves to be sufficiently near to reality, if things have time to hammer logic into men. Where this has happened, one may rest content with this fiction and build theories upon it. . . . Outside of these limits our fiction loses its closeness to reality. To cling to it there also, as the traditional theory does, is to hide an essential thing and to ignore a fact which, in contrast with other deviations of our assumptions from reality, is theoretically important and the source of the explanation of phenomena which would not exist without it. (p. 80)

Nelson and Winter (1982) observe that a consistent theme in retrospective studies is that failure occurs not because the intelligence system failed to acquire warning signals but because the intelligence system failed to process, relate, and interpret those signals into a message relevant to available choices. Intelligence analysts and decision makers have only a limited amount of time each day, limited communication channels to connect their systems, and limited assistance in the task of organizing, analyzing, and thinking about the available information. Sometimes, highly obvious and emphatic signals get lost in the noise as a result of these limitations. The events of September 11, 2001, are a compelling recent example. Nelson and Winter see no reason to think that economic decision making is any different in this regard.

There is similarly a fundamental difference between a situation in which a decision maker is uncertain about the state of X and a situation in which the decision maker has not given any thought to whether X matters or not. To treat them calls for a theory of attention, not a theory that assumes that everything always is attended to but that some things are given little weight (for objective reasons). In short, the most complex models of maximizing choice do not come to grips with the problem of bounded rationality.

Skills. Nelson and Winter (1982) develop the basic postulates about behavior in evolutionary theory. Although evolutionary economics theory is concerned with the behavior of business firms and other organizations, Nelson and Winter find it useful to begin the analysis with a discussion of some aspects of individual behavior. An obvious reason for doing so is that the behavior of an organization is, in a limited but important sense, reducible to the behavior of the individuals who are members of that organization. Regularities of individual behavior must therefore be expected to have consequences, if not counterparts, at the organizational level (see Dosi, Nelson, & Winter, 2000). Nelson and Winter (1982) propose that individual skills are the analogue of organizational routines and that an understanding of the role that routinization plays in organizational functioning is therefore obtainable by considering the role of skills in

individual functioning. By *skill* Nelson and Winter (1982) mean an ability to achieve a smooth sequence of coordinated behavior that is ordinarily effective relative to its objectives, given the context in which the skill normally occurs. Thus, the ability to serve a tennis ball is a skill, as is the ability to engage in competent carpentry, drive a car, operate a computer, set up and solve a linear programming model, or judge which job candidate to hire. Important characteristics of skills are as follows:

- Skills are programmatic (i.e., a sequence of closely followed steps).
- The knowledge that enables a skillful performance is, in large measure, tacit knowledge, in the sense that the performer is not fully aware of the details of the performance and finds it difficult or impossible to articulate a full account of those details.
- The exercise of a skill often involves the making of numerous choices, but to a considerable extent the options are selected automatically and without awareness that a choice is being made.

Nelson and Winter (1982) note that these three aspects of skilled behavior are closely interrelated. Skilled human performance is automatic in the sense that most of the details are executed without conscious volition. Indeed, a welcome precursor of success in an effort to acquire a new skill is the diminishing need to attend to details. Although impressiveness is obviously a matter of degree and relative to expectation, only the most unmoving can escape being impressed, at some point, by a skillful performance.

The late scientist and philosopher Michael Polanyi (1962) wrote extensively of the central place in the general scheme of human knowledge occupied by knowledge that cannot be articulated—tacit knowledge. On the simple observation that we know more than we can tell, Polanyi built an entire philosophical system. Polanyi notes that to be able to do something, and at the same time be unable to explain how it is done, is more than a logical possibility—it is a common situation. Polanyi offers a good example early in the discussion of skills:

I shall take as my clue for this investigation the well-known fact that the aim of a skillful performance is achieved by the observance of a set of rules, which are not known as such to the person following them. For example, the decisive factor by which the swimmer keeps himself afloat is the manner by which he regulates his respiration; he keeps his buoyancy at an increased level by refraining from emptying his lungs when breathing out and by inflating them more than usual when breathing in; yet this is not generally known to swimmers. (p. 49)

Nelson and Winter (1982) note that the tacitness of a skill, or rather of the knowledge enabling a skill, is a matter of degree. Words are probably a more effective vehicle for communicating the skills of elementary algebra than for those of carpentry and more effective for carpentry than for gymnastic stunts. Also, a trait that distinguishes a good instructor is the ability to discover introspectively, and then articulate for the student, much of the knowledge that ordinarily remains tacit. Skill involves the observance of a set of rules, which are not known as such to the person following them.

What are some determinants of the degree of tacitness? First, there is a limit imposed by the feasible time rate of information transfer through symbolic communication, which may be well below the rate necessary or appropriate in the actual performance.

A second consideration that limits the articulation of the knowledge underlying a skill is the limited causal depth of the knowledge. Polanyi's (1962) swimming example illustrates the point that the possession of a skill does not require theoretical understanding of the basis of the skill. Yet this does not imply that an attempt to articulate the basis of the skill would not benefit from the availability of this terminology. Perhaps some novice swimmers would be helped by Polanyi's brief explanation of the body's buoyancy.

The third aspect of the limitation of articulation is the coherence aspect—that of the whole versus the parts. Efforts to articulate complete knowledge of something by exhaustive attention to details and thorough discussion of preconditions succeed only in producing an incoherent message. This difficulty is probably rooted to a substantial extent in the related facts of the linear character of language-based communication, the serial character of the central processor of the human brain, and the relatively limited capacity of human short-term memory. Given these facts, the possibilities of articulating both the details and the coherent patterns they form—the relationships among the details—are necessarily limited. In short, much operational knowledge remains tacit because the knowledge cannot be articulated fast enough, because the knowledge is impossible to articulate all that is necessary to a successful performance, and because language cannot simultaneously serve to describe relationships and characterize the related things.

The knowledge contained in the how-to book and its various supplements and analogues tends to be more adequate when the pace of the required performance is slow and pace variations are tolerable; when a standardized, controlled context for the performance is somehow

assured; and when the performance as a whole is truly reducible to a set of simple parts that relate to one another in some very simple ways. To the extent that these conditions do not hold, the role of tacit knowledge in the performance may be expected to be large.

Finally, it should be emphasized that economic costs matter. Whether a particular bit of knowledge is in principle articulable or necessarily tacit is not the relevant question in most behavioral situations. Rather, the question is whether the economic costs associated with the obstacles to articulation are sufficiently high so that the knowledge, in fact, remains tacit. There is in a sense a trade-off between ability and deliberate choice, a trade-off imposed ultimately by the fact that rationality is bounded. The advantages of skills are attained by suppressing deliberate choice, confining behavior to well-defined channels, and reducing option selection to just another part of the program (March & Simon, 1958). Orthodox microeconomic theory treats the skillful behavior of the businessman as maximizing choice, and choice carries connotations of deliberation. Nelson and Winter (1982), on the other hand, emphasize the automaticity of skillful behavior and the suppression of choice that this skillful behavior involves.

Organizational Capabilities and Behavior. The organizations that Nelson and Winter (1982) envisage are those that face a substantial coordination problem, typically because these organizations have many members, performing many distinct roles, who make complementary contributions to the production of a relatively small range of goods and services. Nelson and Winter provide several salient functions of routines:

1. *Routine as Organizational Memory.* The routinization of activity in an organization constitutes the most important form of storage of the organization's specific operational knowledge. Basically, Nelson and Winter (1982) claim that organizations remember by doing. Exercise of a routine serves as parsimonious organizational memory. Recall that Arrow (1974) gave particular emphasis to the internal dialectic or code of an organization as a key resource of the economies that formal organization provides and as an important cause of persistent differences among organizations.
2. *Routine as Truce.* Routine operation involves a comprehensive truce in intraorganizational conflict (Cyert & March, 1963). Adaptations that appear obvious and easy to an external observer may be foreclosed because such adaptations involve a perceived threat to internal political stability.
3. *Routine as Target: Control, Replication, and Imitation.* Nelson and Winter (1982) note that replication is often a nontrivial exercise. Polanyi (1962) observed the following:

The attempt to analyze scientifically the established arts has everywhere led to similar results. Indeed, even in modern industries the indefinable knowledge is still an essential part of technology. I have myself watched in Hungary a new, imported machine for blowing electric lamp bulbs, the exact counterpart of which was operating successfully in Germany, failing for a whole year to produce a single flawless bulb. (p. 52)

The assumption that perfect replication is possible in evolutionary models is intended primarily to reflect the advantages that favor the going concerns that attempt to do more of the same, as contrasted with the difficulties that they would encounter in doing something else or that others would encounter in trying to copy their success. There are some potential obstacles to replication that may be difficult to overcome even at very high cost. Some employees at the old plant may be exercising complex skills with large tacit components, acquired through years of experience in the firm. Others may have skills of lesser complexity and tacitness but are poor at teaching those skills to someone else—doing and teaching are, after all, different. Some members, for various reasons, may be unwilling to cooperate in the process of transferring their segment of the memory contents to someone else; they may, for example, be unwilling to disclose how easy their job really is or the extent of the shortcuts they take in doing it. Williamson (1975) addresses the question of the incentives of organization members to disclose idiosyncratic information of importance to the organization's functioning under the rubric information impactedness. Nelson and Winter (1982) note that the target routine may involve so much idiosyncratic and impacted tacit knowledge that even successful replication, let alone imitation from a distance, is highly problematic.

4. *Routines and Skills: Parallels.* Nelson and Winter (1982) note that routines are the skills of an organization. Organizations are poor at improvising coordinated responses to novel situations; an individual lacking skills appropriate to the situation may respond awkwardly, but an organization lacking appropriate routines may not respond at all.
5. *Optimal Routines and Optimization Routines.* The heart of Nelson and Winter's (1982) proposal is that the behavior of firms can be explained by the routines that these firms employ. Modeling the behavior of the firm means modeling the routines and how these firms change over time.
6. *Routines, Heuristics, and Innovation.* According to Nelson and Winter (1982), innovation involves change in routine. Similarly, Schumpeter (1934) identified innovation with the "carrying out of new combinations" (pp. 65–66). A *heuristic* is any principle or device that contributes to reduction in the average search to solution. Schumpeter (1950) proposed that sometime during the 20th century the modern corporation routinized innovation.
7. *Routines as Genes.* Nelson and Winter (1982) argue that as a first approximation, firms may be expected to behave in the future according to the routines they have employed in the past. Efforts to understand the functioning of industries and larger systems should come to grips with the fact that highly flexible adaptation to change is not likely to characterize the behavior of individual firms. Evolutionary theory does come to grips with this fact.

Static Selection Equilibrium. Nelson and Winter (1982) note that in Friedman (1953) there is no hint that an evolutionary theory is an alternative to orthodoxy. Rather, the proposition is that selection forces may be the proper explanation of why orthodox theory is a good predictive engine. Alchian (1950) sets forth a perspective regarding firm behavior that resembles Nelson and Winter's in many ways, stressing the element of luck in determining outcomes, the role of learning by trial and feedback and imitation in guiding firms to do better, and of selection forces in molding what firms and industries do. Alchian states the following:

What really counts is the various actions actually tried, for it is from these that success is selected, not from some set of perfect actions. The economist may be pushing his luck too far in arguing that actions in response to changes in environment and changes in satisfaction with the existing state of affairs will converge as a result of adaptation or adoption towards the optimum action that would have been selected if foresight had been perfect. (p. 218)

This statement is not an argument that selection forces provide a reason for adherence to orthodox theory but rather a suggestion that there may be some important differences between an orthodox and an evolutionary perspective. Selection works on what exists, not on the full set of what is theoretically possible (Langlois, 1986; O'Driscoll & Rizzo, 1985).

Competition is viewed as a dynamic process involving uncertainty, struggle, and disequilibrium, not as a tranquil equilibrium state. In evolutionary theory, decision rules are viewed as unresponsive, or inappropriate, to novel situations or situations encountered irregularly and as a legacy from the firm's past and hence appropriate, at best, to the range of circumstances in which the firm customarily finds itself.

The heart of the R&D innovation problem is that reasonable people will disagree about which techniques will be best at which point. Importantly, this uncertainty is a major reason why it makes sense to have R&D largely conducted by competitive business firms who make their own entrepreneurial decisions, rather than place R&D decisions under more centralized control (see Nelson, 1996).

Dynamic Competition and Technical Progress. The market system is (in part) a device for conducting and evaluating experiments in economic behavior and organization. The meaning and merit of competition must be appraised accordingly. In Schumpeter's (1934) terms, competition involves carrying out new combinations. Schumpeter's concept of innovation was a broad one, noting five identified cases:

- The introduction of a new good
- The introduction of a new method of production
- The opening of a new market
- The opening of a new source of supply
- The carrying out of the new organization of any industry, like the creation of a monopoly position (p. 66)

Although Schumpeter (1934) is particularly noteworthy for this emphasis on experimentation, most of the great economists, from Adam Smith (1776/1937) to the onset of the modern period of formalization, gave some weight to the experimental role of competitive markets. An essential aspect of Schumpeterian competition is that firms do not know *ex ante* whether it pays to try to be an innovator or an imitator or what levels of R&D expenditures might be appropriate. Only the course of events over time will determine and reveal which strategies are the better ones. And even the verdict of hindsight may be less than clear.

Normative Organizational Economics From an Evolutionary Perspective. Nelson and Winter (1982) note that the modern advocacy of private enterprise solutions tends to suffer from vagueness or utopianism in its treatment of institutional matters. Three particularly important (and closely interrelated) ones involve the treatment of property rights, contracts, and law enforcement. In almost all formalized economic theories, property rights and contractual obligations are assumed to be costlessly delineated in unambiguous terms, and enforcement of the civil and criminal law is perfect and costless. By virtue of the combined force of these assumptions of clarity, perfection, and zero transaction costs, the problem of providing the basic institutional underpinnings of a system of voluntary exchange is assumed away. It is then not too surprising that voluntary exchange can be shown to be a largely effective economic solution to such problems as are left.

A legal system that could approach the theoretical standards of clarity and perfection in the delineation and enforcement of entitlements would be an elaborate and expensive system indeed. This is particularly obvious if the system of entitlements is supposed to be so sophisticated as to bring within its scope all of the externality problems that economists sometimes treat as merely problems in the definition and enforcement of property rights—for example, the question of whether a chemical plant is entitled to dispose of its hazardous wastes in ways that contaminate the groundwater or whether neighboring

property owners are entitled to uncontaminated groundwater. If the anatomy of market failure is a function of institutional structure, institutional structure itself evolves in part in response to perceived problems with the status quo.

Nelson and Winter (1982) conclude that the attempt to optimize and accordingly to control technological advance will, according to evolutionary theory, lead not to efficiency but to inefficiency. In terms of empirical testing of evolutionary organizational economics, Nelson and Winter note that organizations that operate many very similar establishments—for example, retail and fast-food chains—provide a natural laboratory for studying the problems of control and replication. Students with interest in the area of resources and organizational capabilities should see Foss (1997), Langlois and Robertson (1995), and Nelson and Winter (2002) for an update on recent research literature on dynamic capabilities and evolutionary economics.

Theory and Applications

Resource-based theory addresses some of the fundamental issues in strategy (Rumelt, Schendel, & Teece, 1994; Teece, 2000). Taking 1982 (when Nelson and Winter, 1982, was published) as the starting point, I now discuss some seminal contributions to resource-based theory:

- Lippman and Rumelt (1982): Causal ambiguity inherent in the creation of productive processes is modeled by attaching an irreducible *ex ante* uncertainty to the level of firm efficiency that is achieved by sequential entrants. Without recourse to scale economies or market power, the model generates equilibria in which there are stable interfirm differences in profitability. Sustainable competitive advantage results from the rich connections between uniqueness and causal ambiguity (see also Reed & DeFillippi, 1990; Rumelt, 1984).
- Teece (1982): This article outlines a theory of the multiproduct firm. Important building blocks include excess capacity and its creation, market imperfections, and the characteristics of organizational capabilities, including its fungible and tacit character. Teece both heavily acknowledges and builds on Penrose (1959) and argues that a firm's capabilities are upstream from the end

product—organizational capabilities might well find a variety of end-product applications, as Penrose's (1960) case study of the Hercules Powder Company effectively shows.

- Wernerfelt (1984, 1995): Building on the seminal work of Penrose (1959), these works argue that strategy involves a balance between the use of existing resources and the development of new resources.

- Montgomery and Wernerfelt (1988): According to resource-based theory (Teece, 1982), firms diversify in response to excess capacity of resources that are subject to market frictions. By probing into the heterogeneity of these resources, this article develops the corollary that firms that diversify most widely should expect the lowest average (Ricardian) rents. An empirical test, with Tobin's q as a measure of rents, is consistent with this resource-based theory.

- Dierickx and Cool (1989): This article draws the distinction between tradeable and nontradeable resources (e.g., reputation) and argues for a time-based view of competitive strategy (due, in part, to time compression diseconomies).

- Cohen and Levinthal (1990): The authors argue that prior related knowledge confers an ability to recognize the economic value of new information, assimilate the information, and apply the information to commercial uses. These dynamic capabilities constitute a firm's absorptive capacity. Cross-sectional data on technological opportunity and appropriability conditions in the American manufacturing sector collected for R&D lab managers and the FTC Line-of-Business data indicate that R&D both generates innovation and facilitates learning.

- Henderson and Clark (1990): This article distinguishes between the components of a product and the ways that the components are integrated into the system that is the product architecture. Data were collected during a 2-year, field-based study of the photolithographic alignment equipment industry. The core of the data is a panel data set consisting of research and development costs and sales revenue by product for every product development project

(Continued)

(Continued)

conducted between 1962, when the work on the first commercial product began, and 1986. The concept of architectural innovation provides rich resource-based connections between innovation and organizational capabilities.

- Barney (1991): In this often-cited article, Barney suggests that the search for sources of sustainable competitive advantage must focus on resource heterogeneity and immobility. Barney argues that sustainable competitive advantage is derived from resources that are valuable, rare, imperfectly imitable (due to path-dependence, causal ambiguity, and social complexity), and nonsubstitutable.
- Chatterjee and Wernerfelt (1991): This article theoretically and empirically investigates the resource-based view that firms diversify, in part, to use excess productive resources. In particular, empirical evidence corroborates that excess physical resources and most knowledge-based resources lead to more related diversification.
- Conner (1991): In this article, Conner analyzes resource-based theory as a new theory of the firm and makes insightful connections between resource-based theory and Schumpeterian (1934, 1950) competition.
- Montgomery and Hariharan (1991): Using a sample of 366 firms in the FTC's Line-of-Business database, the research in this article indicates that growth and diversification in large established firms result from a process of matching a firm's lumpy (indivisible) and ever-changing resources with dynamic market opportunities. Overall, this research provides empirical support for Penrose's (1959) theory of diversified entry: Unused productive services of resources are a selective force in determining the direction of firm-level expansion.
- Porter (1991): In this article, Porter argues that firms have accumulated differing resources because of differing strategies and configurations of (value-chain) activities. Resources and activities are, in a sense, duals of each other.
- Williamson (1991): This article suggests the possibility that the dynamic capabilities and resource-based perspectives will play

out in combination. Williamson argues that in the long run, the best strategy for firms is to organize and operate efficiently.

- Leonard-Barton (1992): This article considers core organizational capabilities in terms of employee knowledge and skills, technical systems, managerial systems, and values and norms. Leonard-Barton maintains that managers of new product and process development projects should take advantage of core capabilities while mitigating core rigidities. Twenty case studies of new product and process development projects in five firms (e.g., Chaparral Steel, Ford Motor Company, and Hewlett Packard) provide illustrative data. (For students who find this topic of interest, Leonard-Barton [1995], is an exemplar research book.)

- Mahoney (1992c): In this article, I argue for an integrated organizational economic approach to strategic management based on the behavioral theory of the firm, transaction costs theory, property rights theory, agency theory, and resource-based theory/dynamic capabilities. Essentially, this article outlines the structure of this book.

- Mahoney and Pandian (1992): Following Rumelt (1984), the authors of this paper argue that absent government intervention, isolating mechanisms (e.g., resource position barriers, invisible assets) exist because of asset specificity and bounded rationality.

- Amit and Schoemaker (1993): This article adds behavioral decision-making biases and organizational implementation aspects as further impediments to the transferability or imitability of a firm's resources and capabilities.

- Mosakowski (1993): Using a longitudinal data set, a sample of 86 entrepreneurial firms in the computer software industry that completed an IPO in 1984 is examined. Empirical findings suggest that strategies that represent rare, inimitable and nonsubstitutable resources are a source of competitive advantage.

- Peteraf (1993): This article elucidates the organizational economics logic that is the foundation for the resource-based

(Continued)

(Continued)

theory of Ricardian rents (Ricardo, 1817) and sustainable competitive advantage. The essence of the framework developed here is that four conditions must be met for achieving sustainable competitive advantage: (1) superior resources (firm heterogeneity within an industry), (2) ex post limits to competition (i.e., isolating mechanisms), (3) imperfect resource mobility (e.g., nontradeable assets and cospecialized assets), and (4) ex ante limits to competition.

- Chi (1994): In this article, Chi develops a theoretical framework for analyzing the exchange structure in the trading of imperfectly imitable and imperfectly mobile firm resources. The article first explores the conditions for such resources to be gainfully traded between firms and then investigates the interconnections between barriers to imitation and impediments to trading. A major part of the article is devoted to developing a parsimonious and yet integrative (agency, property rights, and transaction costs) model for assessing the exchange structure between firms that are involved in the trading of strategic resources in the face of significant transaction cost problems, such as adverse selection, moral hazard, contractual cheating, and hold-up problems that are due to information asymmetry, imperfect measurement, imperfect enforcement, and resource interdependencies.

- Farjoun (1994): This article provides empirical support that unused productive services derived from human capital drive the diversification process. Unused productive services from existing human resources present a jigsaw puzzle for balancing processes.

- Henderson and Cockburn (1994): Using both qualitative and quantitative data drawn from both public sources and from the internal records of 10 major European and American pharmaceutical firms, this article attempts to measure the importance of heterogeneous, organizational capabilities. Component and architectural capabilities together explain a significant fraction of the variance in research productivity across firms.

- Godfrey and Hill (1995): This article persuasively espouses the realist philosophy of science, which states that we cannot

reject theories just because they contain key constructs that are unobservable.³ It is not enough to state that the unobservability of utility dooms agency theory, that transaction costs theory is untestable because some transaction costs cannot be measured, or that resource-based theory is invalid because key resources (e.g., invisible assets) are unobservable. To reject a theory one must be able to show that the predictions of observable phenomena that are derived from the theory do not hold up under empirical testing.

- Mahoney (1995): In this article, I argue that the resource-based approach of deductive economics, the dynamic capabilities approach of strategy process, and organization theory research on organizational learning (e.g., Argyris & Schon, 1978; Fiol & Lyles, 1985) need to be joined in the next generation of resource-based research.

- Zander and Kogut (1995): Based on their developed questionnaire distributed to project engineers knowledgeable of the history of 44 major innovations in 20 firms, the authors conclude that the transfer of manufacturing capabilities is influenced by the degree to which capabilities may be codified and taught. Empirical evidence corroborates the view that the nature of dynamic capabilities and the nature of competitive positioning matter.

- Foss (1996): The author argues that there are complementarities between a contractual approach (e.g., transaction costs theory and property rights theory) and a knowledge-based approach (e.g., resource-based theory and knowledge-based theory) to strategic management. These complementarities are argued to be particularly fruitful for analyzing the strategic issues of the boundary and internal organization of the firm.

³In addition to Godfrey and Hill's (1995) lucid discussion on realist philosophy, there are a number of works that cover various issues in philosophy of science and research methodology that are relevant to strategic management research, including Blaug (1980); Caldwell (1984); Camerer (1985); Evered and Louis (1981); Huff (1981, 2000); Kaplan (1964); Kuhn (1970); Ladd (1987); Machlup, (1967); MacKinlay (1997); Mahoney (1993); Mahoney and Sanchez (1997, 2004); McCloskey (1983, 1998); McCloskey and Ziliak (1996); Montgomery, Wernerfelt, and Balakrishnan (1989); Redman (1993); Seth and Zinkhan (1991); and Whetten (1989).

(Continued)

(Continued)

- Grant (1996): In this article, Grant argues that organizational capabilities are the outcome of knowledge integration: complex, team-based productive activities that cohesively integrate the knowledge of many individual specialists. Research in cross-functional capabilities in the context of new product development (Clark & Fujimoto, 1991) would be an exemplar.
- Miller and Shamsie (1996): This article empirically tests resource-based theory in the context of the seven major United States film studios (i.e., MGM, Twentieth Century-Fox, Warner Brothers, Paramount, United Artists, Universal, and Columbia) from 1936 through 1965. The authors find that property-based resources in the form of exclusive long-term contracts with celebrities and theaters helped financial performance in the stable environment from 1936 to 1950. In contrast, knowledge-based resources in the form of production and coordination talent boosted financial performance in the more uncertain posttelevision environment.
- Mowery, Oxley, and Silverman (1996): Examining cross-citation rates for 792 partners in bilateral alliances that involved at least one U.S. firm and were established during 1985 and 1986, this article provides empirical support for the importance of gaining capabilities through alliances. The empirical results bolster the argument that experience in related technological areas is an important determinant of absorptive capacity.
- Spender (1996): Building on Nelson and Winter (1982) and Nonaka and Takeuchi (1995), this article views the firm as a dynamic knowledge-based activity system. The author's arguments are consistent with Penrose's (1959) view of knowledge as the skilled process of leveraging resources, where that knowledge is embedded in the organization.
- Szulanski (1996): Based on 271 observations of 122 best-practice transfers in eight companies, the major barriers to internal knowledge transfer are found to be knowledge-related factors, such as the recipient's lack of absorptive capacity, causal ambiguity, and an arduous relationship between the source and the recipient.
- Helfat (1997): This empirical investigation of dynamic R&D capabilities examines the role of complementary know-how and

other resources in the context of changing conditions in the U.S. petroleum industry during the 1970s and early 1980s. The empirical analysis indicates that in response to rising oil prices, firms with larger amounts of complementary technological knowledge and physical resources also undertook larger amounts of R&D on coal conversion (a synthetic fuel process).

- Powell and Dent-Micallef (1997): This article examines the information technology literature, develops an integrative resource-based theoretical framework, and presents results from an empirical study of the retail industry. The empirical results support the view that information technology creates economic value by leveraging and using complementary human and physical resources.

- Teece, Pisano, and Shuen (1997): This article views the dynamic capabilities perspective as building on Schumpeter (1934, 1950), Nelson and Winter (1982), and Teece (1982). Focal concerns are resource accumulation, replicability, and inimitability of organizational capabilities.

- Tripsas (1997): This article analyzes the technological and competitive history of the global typesetter industry from 1886 to 1990. Key success factors include investment, technical capabilities, and appropriability through specialized complementary assets.

- Bogner, Mahoney, and Thomas (1998): In this article, following Machlup (1967), the authors argue that resource-based theory needs to move beyond (1) theoretical construction that abstracts from historical time, (2) theory that focuses only on the stationary state, (3) theory where taxonomic and tautological arguments are made, (4) theory that focuses exclusively on the conditions for establishing equilibrium, and (5) theory that omits time as an independent variable.

- Farjoun (1998): This article examines empirically the joint effect of skill-based and physical-based related diversification on accounting and financial measures of performance. For a sample of 158 large diversified manufacturing firms, the joint effort of skill-based and physical-based related diversification had a strong

(Continued)

(Continued)

positive effect on most indicators of performance. This finding corroborates resource-based theory that related diversification that builds on both skill-based and physical-based resources allows firms to create economic value by sharing and transferring these resources and to use activities and routines in which these resources interact.

- Lieberman and Montgomery (1998): Building on Lieberman (1987) and Lieberman and Montgomery (1988), the authors of this article argue that resource-based theory and first-mover (dis)advantage are related conceptual frameworks that can benefit from closer linkages.

- Argote (1999): This book presents evidence that organizations vary tremendously in the rate at which they learn. Argote argues that differences in patterns of knowledge creation, retention, and transfer contribute to differences in the rates at which organizations learn.

- Brush and Artz (1999): Using a sample of 193 veterinary practices, this article investigates contingencies among resources, capabilities, and performance in veterinary medicine. Empirical evidence supports the view that the economic value of resources and capabilities depends on the information asymmetry characteristics of the product market.

- Silverman (1999): This article considers how a firm's resource base affects the choice of industries into which the firm diversifies and offers two main extensions of prior resource-based research. First, the paper operationalizes technological resources at a more fine-grained level than in prior empirical studies, thereby enabling a more detailed analysis concerning the direction of diversification. This analysis indicates that the predictive power of resource-based theory is greatly improved when resources are measured at a more fine-grained level. Second, the article integrates transaction costs theory and resource-based theory to provide more detailed predictions concerning diversification. Empirical evidence suggests circumstances where resources (that have high asset specificity) can be and are used through contracting rather than through becoming a diversified firm.

- Williamson (1999): This article suggests that one way of looking at research opportunities in strategic management is to view transaction costs theory as feeding into the organizational capabilities perspective. Both transaction costs theory and resource-based theory are viewed as needed in our efforts to understand complex business phenomena as we build a science of organization.
- Yeoh and Roth (1999): This article empirically examines the impact of firm resources and capabilities using a sample of 20 pharmaceutical firms that operated as separate entrepreneurs between 1971 and 1989. The empirical results indicate that R&D and sales force expenditures have direct and indirect effects on sustainable competitive advantage.
- Ahuja and Katila (2001): Using a sample of acquisition and patent activities of 72 leading firms from the global chemicals industry from 1980 to 1991, the relatedness of acquired and acquiring knowledge-based resources has a nonlinear impact on innovation output. In particular, acquisition of firms with high levels of both relatedness and unrelatedness prove inferior to acquiring firms with moderate levels of knowledge-based relatedness.
- Bowman and Helfat (2001): This article examines the resource-based theory that there is a significant role for corporate strategy based on the use of common resources by related businesses within a firm (Peteraf, 1993; Teece, 1982). Based on an analysis of the variance decomposition research literature, Bowman and Helfat conclude that corporate strategy (Andrews, 1980; Ansoff, 1965), in fact, does matter for economic performance.
- Makadok (2001): This article provides a mathematical model synthesizing resource-based and dynamic capabilities views of economic value creation. Resource picking (emphasized by resource-based theory) and capability building (emphasized by the dynamic capabilities approach) for the purpose of achieving economic rent creation are shown to be complementary in some business circumstances but are shown to be substitutes in other business circumstances.

(Continued)

(Continued)

- Mahoney (2001): In this article, I argue that resource-based theory is primarily a theory of economic rents, whereas transaction costs theory is primarily a theory of the existence of the firm. These two theories are complementary and are connected in the following way: Resource-based theory seeks to delineate the set of market frictions that would lead to firm growth and sustainable economic rents (via isolating mechanisms), whereas transaction costs theory seeks to delineate the set of market frictions that explain the existence of the firm. The article submits that the set of market frictions that explain sustainable firm rents (in resource-based theory) will be sufficient market frictions to explain the existence of the firm (in transaction costs theory). I also argue that the resource-based theory of the strategic (rent-generating and rent-sustaining) firm cannot assume away opportunism.

- Afuah (2002): This article provides a model for mapping firm capabilities into competitive advantage. Using a sample of 78 observations for cholesterol drugs in the market from 1988 to 1994, the author illustrates how the model can be used to estimate competitive advantage from technological capabilities.

- Coff (2002): Empirical results from a sample of 324 acquisitions that closed or failed to close in the years 1988 and 1989 offer evidence in support of the hypothesis that related human capital expertise between the acquirer and acquired enterprise can mitigate opportunism hazards associated with human capital asset specificity (Becker, 1964). In this business setting, related knowledge-based resources, in the form of related human expertise, increases the probability that a given transaction will close.

- Madhok (2002): This article maintains that a strategic theory of the firm should not only address the decision with respect to hierarchical governance or market governance but should also take into account how a firm's resources and capabilities can best be developed and deployed in the search for competitive advantage. Or, put differently, transaction costs theory should be coupled with resource-based theory.

- Thomke and Kuemmerle (2002): Using a combination of field research, discovery data from nine pharmaceutical firms, and

data on 218 alliances involving new technologies for experimentation and testing, several causes affecting resource accumulation are identified and described. The article provides empirical support that the difficulty of imitating a particular resource is affected by the interdependencies with other resources.

- Adner and Helfat (2003): This article adds to the study of competitive heterogeneity by measuring the economic effect of specific corporate-level managerial decisions, driven by dynamic managerial capabilities, on the variance of economic performance among U.S. energy companies. The empirical results also strongly suggest that corporate managers matter.
- Helfat and Peteraf (2003): This article introduces the capability life cycle, which identifies general patterns and paths in the evolution of organizational capabilities over time. The framework is intended to provide a theoretical structure for a more comprehensive approach to dynamic resource-based theory.
- Hoopes, Madsen, and Walker (2003): This article maintains that the resource-based view's accomplishments are clearer when seen as part of a larger theory of competitive heterogeneity. Combining economics, organization theory, and traditional business policy, the resource-based view suggests how, in a competitive environment, firms maintain unique and sustainable positions.
- Knott (2003a): The author of this article finds that franchising routines are both valuable and can lead to sustainable competitive advantage. The upshot of this empirical research is that tacit knowledge is not necessary for having an isolating mechanism.
- Knott (2003b): This article outlines a theory of sustainable innovation fueled by persistent heterogeneity. Knott shows that there exist conditions that generate persistent heterogeneity and sustainable innovation with each firm behaving optimally, taking other firms' behaviors into account.
- Lippman and Rumelt (2003): This article critiques the microfoundations of neoclassical theory and develops further the

(Continued)

(Continued)

concept of rent. The article also provides insights on rent sensitivity analysis and a payments perspective of strategic management.

- Makadok (2003): This article models mathematically the joint impact of two determinants of profitable resource advantages: the accuracy of managers' expectations about the future economic value of a resource and the severity of agency problems that cause managers' interests to diverge from those of shareholders. The conclusion is that future research on the origins of competitive advantage should examine agency and governance issues along with, not apart from, resource-based issues.

- Szulanski (2003): This research book on sticky knowledge addresses an important question for managers: Why don't best practices spread within organizations? Szulanski explores the effect of motivational and knowledge barriers on knowledge transfer and presents the empirical results of statistical analyses that stem from data collected through a two-step questionnaire survey. The research relies on 271 surveys studying the transfer of 38 (technical and administrative) practices in eight companies. Szulanski finds that knowledge barriers to transfer have a larger effect on the stickiness of knowledge than motivational barriers, and the two barriers jointly explain nearly 75% of the variance in stickiness.

To conclude this chapter, I focus on a particularly important set of dynamic capabilities that are embedded in real options in strategic decision making. Trigeorgis (1996) provides both rigor and relevance concerning strategic (real) options.

Real Options: Managerial Flexibility and Strategy in Resource Allocation (*Trigeorgis, 1996*)

Financial theory, properly applied, is critical to managing in an increasingly complex and risky business climate. . . . Option analysis provides a more flexible approach to valuing our research investments. . . . To me all kinds of business decisions are options. (Judy Lewent, as cited in Nichols, 1994, and reprinted in Trigeorgis 1996, p. xv)

Trigeorgis (1996) deals with the classical subject of resource allocation or project appraisal under uncertainty, particularly with the economic valuation of managerial operating flexibility and strategic actions as corporate real options. Similar to options on financial securities, real options involve property rights (with no obligations) to acquire or exchange an asset for a specified alternative price. The ability to value options (e.g., to defer, abandon, and grow) has brought a revolution to modern corporate finance theory on resource allocation.

Corporate value creation and competitive positioning are critically determined by corporate resource allocation and by the proper evaluation of investment alternatives. Trigeorgis (1996) argues that traditional quantitative techniques such as discounted cash flow (DCF) analysis (that consider the size, timing, and uncertainty of cash flows) have failed in business practice because these techniques traditionally have not properly captured managerial flexibility to adapt and revise later decisions in response to unexpected market developments. Moreover, these techniques traditionally neither capture the strategic value resulting from proving a technology viable nor capture the economic impact of project interdependencies and competitive interactions. In the Nelson and Winter (1982) sense, organizational capabilities that enhance adaptability and strategic positioning provide the infrastructure for the creation, preservation, and exercise of corporate real options that can have significant economic value.

Trigeorgis (1996) notes that, in practice, managers have often been willing to overrule traditional investment criteria to accommodate operating flexibility and other strategic decisions that managers consider just as valuable as direct cash flows. It is now widely recognized, for example, that traditional DCF approaches to the appraisal of capital investment projects, such as the standard net-present-value (NPV) rule of accepting positive NPV projects, do not properly capture management's flexibility to adapt and revise later decisions in response to unexpected market developments. Or, put differently, a theoretically accurate NPV analysis would include real options values.

Trigeorgis (1996) argues that in the business marketplace, which is characterized by change, uncertainty, and competitive interactions, the realization of cash flows will probably differ from what management expected at the outset. As new information arrives and uncertainties about market conditions and future cash flows are gradually resolved, management may have valuable flexibility to alter its initial operating plan to capitalize on favorable future opportunities or to react so as to mitigate economic losses. For example, management may be able to

defer, expand, contract, abandon, or otherwise alter a project at various stages of the project's useful operating life.

This managerial operating flexibility is likened to financial options. An American call option of an asset (with current value V) gives the right, with no obligation, to acquire the underlying asset by paying a prespecified price (the exercise price, I) on or before a given maturity. Similarly, an American put option gives the right to sell (or exchange) the underlying asset and receive the exercise price. *The asymmetry derived from having the right but not the obligation to exercise an option is at the heart of the option's value.*

Trigeorgis (1996) notes that as with options on financial securities, management's flexibility to adapt its future actions in response to altered future market conditions and competitive reactions expands a capital investment opportunity's value by improving its upside potential while limiting its downside economic losses relative to the initial expectations of a passive management. The resulting asymmetry calls for a strategic investment criterion, reflecting both value components: the traditional static NPV of direct cash flows and the real option value of operating flexibility and strategic interactions.

Trigeorgis (1996) argues that a real options approach to capital budgeting has the potential to conceptualize and quantify the value of options from active management and strategic interactions. This economic value is typically manifest as a collection of real options embedded in capital investment opportunities, having as the underlying asset the gross project value of discounted expected operating cash inflows. Many of these real options (e.g., to defer, contract, shut down, or abandon a capital investment) occur naturally; other real options may be planned and built in at some extra cost from the outset (e.g., to expand capacity, to build growth options, to default when investment is staged sequentially, or to switch between alternative inputs or outputs). Let us now consider various real options.

1. *Option to Defer Investment.* The real option to defer an investment decision is analogous to an American call option on the gross present value of the completed project's expected operating cash flows, V , with an exercise price equal to the required outlay, I . Management holds a lease on (or an option to buy) valuable land or resources. Management can wait x years to see if output prices justify constructing a building or a plant or developing a field. The option to wait is particularly valuable in natural-resource extraction industries, farming, paper products, and real estate development.

Consider the following example from Dixit and Pindyck (1994): How should we decide whether or not to enter into a business? If we refer to the literature on finance, the traditional approach is to use cash flow analysis using a net present value criterion. For example, let us imagine a situation in which we are considering entering the business of making widgets. Assume that it costs \$1,600 to build a widget factory and that our current cost of capital is 10%. In addition, we sell only one widget per year, and the current price of a widget is \$200. Although we know the current price for widgets, we are somewhat uncertain about the future prices. Forecasts indicate that there is a 50% chance that prices will go up to \$300 next period (and remain there forever); however, there is also a 50% chance that prices will go down to \$100. This forecast implies that the expected price of widgets in the future is $\$200 = (0.5 \times \$300 + 0.5 \times \$100)$.

Using these numbers, we can evaluate this project with a standard cash flow analysis. The expected cash flow from entering the widget business appears in the first column of Table 5.1. In period 0, we build the plant ($-\$1,600$) and begin production, receiving \$200 in revenues ($-\$1,600 + \$200 = -\$1,400$). From that period on, we have positive expected cash flow of \$200. We can use this cash flow series to arrive at the NPV for the project, which is \$600. (Because the value at T_0 of a perpetuity cash flow [CF] beginning at T_1 with a discount rate r equals CF/r then here it is: $\$200/.1 = \$2,000$; then we take \$2,000 and subtract \$1,400 to arrive at \$600.) We would then proceed with the project because the NPV of \$600 is greater than zero.

However, what if we wait a period to find out whether the price goes up or down? That is, what if we choose to keep our options open and remain flexible in our decision? Two different scenarios could occur. The first possibility is that the price goes up to \$300, in which case we would experience the cash flow under Scenario 1 in Table 5.1. The second possibility is that the price goes down to \$100, in which case we obtain the cash flows under Scenario 2. Now, one will notice that under Scenario 1, the NPV (in period 0) is positive (i.e., $NPV = \$1,545$); however, under Scenario 2, the NPV is negative (i.e., $NPV = -\$455$). (The present value of the perpetuity is $\$300/.1 = \$3,000$, from which we subtract $\$1,600/.1 = -\$1,455$ to arrive at $\$1,545$, and for the low demand scenario we have $\$100/.1 = \$1,000$ and subtract $\$1,600/.1 = -\455 .) Thus, if we waited a period and the price went up to \$300, we would proceed with the project, whereas if the price went down, we would not proceed with the project. Thus, under the second scenario, the actual NPV would not be $-\$455$ but would be \$0; that is, we would not invest

Table 5.1 Calculations of Discounted Cash Flows

<i>Time</i>	<i>Expected Cash Flow (Traditional NPV)</i>	<i>Expected Cash Flow (Scenario 1)</i>	<i>Expected Cash Flow (Scenario 2)</i>
0	\$ (1,400)	\$ --	\$ --
1	\$ 200	\$ (1,300)	\$ (1,500)
2	\$ 200	\$ 300	\$ 100
3	\$ 200	\$ 300	\$ 100
4	\$ 200	\$ 300	\$ 100
5	\$ 200	\$ 300	\$ 100
6	\$ 200	\$ 300	\$ 100
7	\$ 200	\$ 300	\$ 100
8	\$ 200	\$ 300	\$ 100
Inf.	\$ 200	\$ 300	\$ 00
NPV	\$ 600	\$ 1,545	\$ (455)

in a negative NPV project. What does this tell us about the value of waiting and remaining strategically flexible?

One way of answering this question is to reframe our cash flow analysis. Instead of taking the NPV of the expected cash flows, let us calculate the expected NPV of the two scenarios combined. That is, we have a 50% chance of the price going up and getting an NPV of \$1,545 and a 50% chance of the price going down and getting \$0. The expected combined NPV is therefore approximately \$773 ($= 0.5 \times \$1,545 + 0.5 \times 0$). The NPV where we wait, find out the true price, and then make the decision is larger (by \$173) than going ahead right now. There is (an option) value to waiting of \$173. Thus, we can increase our expected returns by waiting a year and then deciding whether to undertake the sunk-cost investments in a new plant.

Summary. The previous example illustrates that even when the static (positive) NPV calculation suggests a go, when the real options value of flexibility is taken into account, the top-level manager should wait. The option to wait is equivalent to a call option on the investment project. The call is exercised when the firm commits to the project. But often it is better to defer a positive-NPV project to keep the call option alive. Deferral is most attractive when uncertainty is great and immediate project cash flows—which are lost or postponed by waiting—are small.

2. *Option to Default During Staged Construction (Time-to-Build Option).* Each stage of an investment can be viewed as an option on the economic value of subsequent stages by incurring the installment-cost

outlay (e.g., I_1) required to proceed to the next stage and can therefore be valued similar to options on options (or compound options). Staging the investment as a series of outlays creates the real option to abandon the project in midstream if new information is unfavorable. This real option is valuable in R&D-intensive industries (especially pharmaceuticals); in highly uncertain long-development, capital-intensive industries (such as energy-generating plants or large-scale construction); and in venture capital.

3. *Option to Expand, Contract, Shut Down, or Restart Operations.* If market conditions are more favorable than expected, the firm can expand the scale of production or accelerate resource use. Conversely, if conditions are less favorable than expected, the firm can reduce the scale of operations. In extreme cases, production may be halted and restarted. Applications can be found in natural-resource industries (e.g., mining), facilities planning and construction in cyclical industries, fashion apparel, consumer goods, and commercial real estate.

4. *Option to Abandon for Salvage Value.* Management may have a valuable real option to abandon a project in exchange for its salvage value. Naturally, more general-purpose capital assets would have a higher salvage value and abandonment option value than special-purpose assets. Valuable abandonment options are generally found in capital-intensive industries (such as airlines and railroads), in financial services, and in new-product introductions in uncertain markets. Abandonment should not be exercised lightly if it might lead to eventual erosion of valuable expertise and other crucial organizational capabilities that could be applied elsewhere in the business or that could prevent the firm from participating in future technological developments. Moreover, abandonment may lead to the loss of goodwill from customers.

5. *Option to Switch Use (e.g., Inputs or Outputs).* Generally, process flexibility can be achieved not only via technology (e.g., by building a flexible facility that can switch among alternative energy inputs) but also by maintaining relationships with a variety of suppliers and switching among them as their relative prices change. Process flexibility is valuable in feedstock-dependent facilities, such as oil, electric power, chemicals, and crop switching. Product flexibility—enabling the firm to switch among alternative outputs—is more valuable in industries such as machine parts, automobiles, consumer electronics, toys, specialty paper, and pharmaceuticals, where product differentiation and diversity

are important or product demand is volatile. In such business cases it might be worthwhile to install a more costly flexible capacity to acquire the dynamic capability to alter product mix or production scale in response to changing market conditions.

6. *Corporate Growth Options.* Corporate growth options that set the path of future opportunities are of considerable strategic importance. Although in isolation a proposed facility may appear unattractive, such a facility may be only the first in a series of similar facilities if the process is successfully developed and commercialized, and it may even lead to entirely new by-products. Many early investments (e.g., in R&D) can be seen as prerequisites or links in a chain of interrelated projects. The value of the early projects derives not so much from their expected directly measurable cash flows as from the future growth opportunities they may unlock (e.g., access to a new market or strengthening of the firm's core capabilities and its strategic positioning). An opportunity to invest in a first-generation high-tech product, for example, is analogous to an option on options (an interproject compound option). Despite a negative static NPV, the infrastructure, experience, and potential by-products generated during the development of the first-generation product may serve as springboards for developing lower cost or higher quality future generations of that product, or even for generating entirely new applications. But unless the firm makes the initial investment, subsequent generations or other business applications will not even be feasible. The infrastructure and experience gained, if maintained as proprietary knowledge, can place the firm at a competitive advantage, which may even reinforce itself when learning-cost-curve effects are present. Growth options are found in all infrastructure-based or strategic industries—especially in high technology, R&D, industries with multiple product generations or applications (e.g., semiconductors, computers, pharmaceuticals), multinational operations, and strategic acquisitions.

7. *Multiple Interacting Options.* Real-life projects often involve a collection of various options. Upward-potential-enhancing and downward-protection options are present in combination. Their combined economic value may differ from the sum of their separate values (i.e., they interact). They may also interact with financial flexibility options. Applications include most industries listed previously.

Trigeorgis (1996) argues that real options have the potential to make a significant difference in strategic management. Sustainable competitive advantages resulting from proprietary technologies, scale, ownership of

valuable natural resources, managerial capital, reputation, brand name, or patents (Andersen, 2001; Arora, Fosfuri, & Gambardella, 2001) empower companies with valuable real options to grow through future profitable investments and to more effectively respond to unexpected adversities or opportunities in a changing technological, competitive, or general business environment. Students studying the economics of organization have ample opportunity to supplement real options analysis (i.e., often decision-theoretic) with game-theoretic tools capable of incorporating strategic competitive responses, and this research area promises to be an important and challenging direction for strategic management and corporate finance research.

Applications of the Real Options Perspective

- In 1984 the W. R. Grace Corporation made an investment in a new technology for automotive catalytic converters. Although the technology proved uncompetitive on price in the automotive market, new applications arose in cogeneration plants and in utility emission control as a result of the U.S. Clean Air Act.
- In research and development, many high-technology companies invest heavily in technologies that may result in a wide range of possible outcomes and new potential markets but with a high probability of technical or market failure. In the pharmaceutical industry, for example, on average, it costs \$360 million and takes a decade to bring a new drug to the market. Only 1 explored chemical in 10,000 becomes a prescription drug, and once a drug reaches the market it faces a 70% chance of failing to earn the cost of invested capital. Such investments are hard to sell to top management on financial grounds; their benefits are remote and hard to quantify, even though intuitively their growth potential seems promising. Instead of ignoring these technologies, a company can make a capital commitment in stages, effectively taking a call option on the underlying technology (or future applications). The initial outlay is not made so much for its own cash flows as for its growth-option value.
- Merck and Company embarked on extensive automation, starting with a drug packaging and distribution plant, even though

(Continued)

(Continued)

technical success was uncertain and projected labor savings did not seem to justify the investment. Operations valuation allowed engineers to articulate the whole range of outcomes and their benefits. Indeed, building and using real options-based planning methods were viewed as having created a valuable new capability for Merck (Nichols, 1994). In fact, the more uncertain the technology or the future market demand, the higher the value of the real option.

- The case of the adoption of the thin-slab caster by Nucor involved consideration of sunk-cost commitments and real options (Ghemawat, 1997):

- *The Option to Wait (and Learn) Before Investing.* In the Nucor case, it was very unlikely that another firm would be willing to be the pioneering firm to deploy this new technology. Thus, if Nucor were to wait, the reduction in uncertainty would have been small.

- *The Option to Make Follow-On Investments if the Immediate Investment Project Succeeds.* In the Nucor case, the first plant appeared to have a slightly negative NPV on a stand-alone basis across a majority of likely scenarios. However, the experience gained in building the first plant would substantially improve the economics of subsequent plants. Thus, the first plant could merely be the price of admission representing a necessary learning curve. Thus, even though the narrow (negative) NPV calculation for Nucor suggested a no-go, the growth options tipped the scale to go. Nucor took into account the strategic value of taking on this negative-NPV project. A close look at Nucor's payoffs reveals a call option on follow-on projects in addition to the immediate project's cash flows. Today's investments can generate tomorrow's opportunities.

- *The Option to Abandon the Project.* Even if the compact strip production (CSP) fails, that component of the mill could be potentially replaced by another technology; the bulk of the mill, such as electric arc furnace and rolling mills, may be useable even with another thin-slab technology. Thus, when the narrow (negative) NPV calculation suggests a no-go, a high options value of abandonment (i.e., low switching

costs) may tip the scale to go. The option to abandon a project provides partial insurance against failure. This is a put option; the put option's exercise price is the value of the project's assets if sold or shifted to a more valuable use.

The flexibility provided by flexible manufacturing systems, flexible production technology, or other machinery having multiple uses can be analyzed from the real options perspective. Recently, the flexibility created by modular design that connects components of a larger system through standardized interfaces (and its impact on organization design) has captured attention in strategic management (Baldwin & Clark, 2000; Bowman & Kogut, 1995; Garud & Kumaraswamy, 1995; Garud, Kumaraswamy, & Langlois, 2003; Langlois, 2002; Sanchez & Mahoney, 1996, 2001; Schilling, 2000; Worren, Moore, & Cardona, 2002). Students studying the economics of organization have an opportunity to evaluate such flexibility using the real options framework.

In conclusion, the current academic research literature in corporate finance has largely framed real-options problems as decision theoretic. However, we now need to move on to considerations where the timing of investments also depends on how other players will respond. Thus, strategic management must take into account both decision-theoretic problems and game-theoretic problems in the next generation of real options research.⁴

Concluding Comments. The resource-based, dynamic capabilities, and real options literatures are potentially highly synergistic for theory development, empirical testing, and business applications. Students with research interests in business history (e.g., Chandler, 1990), evolutionary theory and organizational capabilities (e.g., Nelson & Winter, 1982), corporate finance (e.g., Trigeorgis, 1996), strategic human resource management (Baron & Kreps, 1999), and entrepreneurship (Penrose, 1959) are anticipated to contribute to the evolving science of organization.

⁴For strategic management contributions to the real options perspective, see Bowman and Hurry (1993), Chi (2000), Folta (1998), Folta and Miller (2002), Kogut (1991), McGrath (1997, 1999), Miller (2002), Miller and Folta (2002), and Sanchez (1993, 2003). Schwartz and Trigeorgis (2001) provide a number of classical readings and recent contributions on real options and investment under uncertainty.

