T

Tacit Knowledge

Sidney G. Winter
University of Pennsylvania, Wharton School of Business, Philadelphia, PA, USA

Abstract

As emphasized by Michael Polanyi, there are types of knowledge that cannot be conveyed effectively by means of language. The validity of this observation is particularly apparent in the case of the knowledge required to execute psychomotor skills, such as riding a bicycle. In strategic management, tacitness is generally seen as a barrier to the imitation of valuable knowledge, hence as a relevant factor in accounting for firm heterogeneity and potentially for the sustainability of competitive advantage. However, knowledge that is tacit may be transferrable nevertheless, though not by means of words exclusively.

Definition Tacit knowledge is knowledge that cannot be conveyed effectively from one person to another by means of language.

The concept of 'tacit knowledge' is best evoked by contrast with its opposite, fully articulable knowledge. Articulable knowledge can be conveyed from the knower to a recipient by means of language. The 'language' may be everyday language of whatsoever nationality, or it may be the language of some technical specialty, or perhaps a notation system like that of mathematics or chemistry. The conveyance may be in oral or written form, or perhaps Morse code or just a bit string – provided the recipient's *understanding* of the message is ultimately linguistic. In sum, 'language' is the thing that conveys meaning in symbolic form. The purpose of the conveyance may simply be explanation, or it may be to enable a certain performance by the recipient. Lock combinations and phone numbers are simple examples of symbol strings that accomplish such conveyance for a practical purpose; they approach the limit of full articulation – though only in contexts where certain background conditions are (tacitly) taken for granted.

For a variety of reasons, knowledge is often not fully articulable. A prominent reason is that there is continuous intermediation by the brain and neurophysiology of a human being between the potentially articulable world of meaning and the world of physical effects. For example, sound waves come in, impinging on ear drums, and muscles may ultimately move in response – but in between there is a translation from sound waves to phonemes to meanings, and then another from meanings to intentions to muscular actions. While the ability to learn these translations is a universal feature of human intelligence, the learning is not at the level of the meanings of symbols – how could it be, since meaning itself is what these translations determine and reflect?

Thus, we humans are highly capable of acquiring knowledge that cannot be articulated – which

© Macmillan Publishers Ltd., part of Springer Nature 2018
M. Augier, D. J. Teece (eds.), *The Palgrave Encyclopedia of Strategic Management*,
https://doi.org/10.1057/978-1-137-00772-8

is hardly surprising, considering the impressive learning capabilities displayed by our non-human relatives and ancestors. To say, for example, that 'driving involves tacit knowledge' is to imply that there would be no driver training schools if a little book on how to drive could fully convey driving skill. Learning how to drive involves creating new neuronal linkage at the unconscious level. This basic point seems quite clear, and it is indeed clear and important across a very wide domain of examples.

The insight that knowledge can be tacit has a long history, but the most influential development of the idea was by Michael Polanyi in his 1958 book *Personal Knowledge*: *Towards a Post-Critical Philosophy*. Polanyi's ambitions for that volume were large indeed: the questions illuminated ranged across philosophy, mathematics, science, technology, politics and religion.

Although the book does not mention strategy per se, its discussion of tacit knowing has nevertheless had substantial impact on the strategic management field. The explanation for this derives first from the fact that Polanyi discussed skillful performance as a leading illustration of the reality of tacit knowledge (Polanyi 1962: ch. 4). He analysed specific examples, including the oft-cited case of the knowledge required to ride a bicycle. The realities of skill have obvious relevance to any assessment of how the world's work actually gets done. Polanyi also commented more specifically, though tangentially, on the role of tacit knowledge in traditional manufacturing technology and in ► technology transfer, and more extensively on the relationships of scientific and technical knowledge (Polanyi 1962: 56, 174–183).

From his general appreciation of tacit skill comes the insight that the knowledge requirements of productive activity do not consist of articulable knowledge alone, and a very different way of thinking is needed as a complement to the understanding derived from a focus on the latter. Such an alternative view has strongly shaped the development of theories of routines and capabilities. Polanyi's more specific remarks on technology point directly to central concerns of the strategic management field, particularly the question of whether an economically successful performance can be replicated or imitated. This question has direct bearing on the analysis of persistent profitability in the '► resource-based view'. If a firm has mastered tacit knowledge that enables an 'abnormally profitable' productive performance, does that imply that it has a 'sustainable competitive advantage'?

As usual, careful reflection can reveal hazy limits to the range of supposedly clear examples. In the digital age, we confront the fact that a bit string can convey a picture or a video. Does a video of a gymnastic stunt become an 'articulate' account of it just because it can be transmitted as a bit string? No, because the understanding of it is not 'ultimately linguistic'. Yet the video is clearly not in the same category as a string of words. Very likely it is more promising as a transfer mechanism in practice, though hardly adequate by itself. Thus, we are driven to the somewhat awkward conclusion that a bit string can be a medium for the conveyance of tacit knowledge.

Tacitness in some degree is actually the typical situation as regards the character of productive knowledge. The contrasting domain of *fully articulable knowledge* is narrowly circumscribed by multiple constraints and necessary conditions. But, because articulation is an extremely powerful influence on the progress and diffusion of knowledge *within* its limited domain, its importance is not open to serious question, and it often pre-empts attention in discussions of knowledge.

If language, in any form, does not suffice for knowledge transfer, what does? More specifically, how are we to account for the fact that some highly tacit skills – such as driving, touch typing, or effective use of a mobile phone or a saxophone – have become so widely diffused in the modern world? The answer is that in many cases there are relatively costly, non-linguistic transfer mechanisms available – and the costs are incurred when there are incentives to do so. The unifying element of these costly mechanisms is a requirement for serious engagement by the individual or organization that is the recipient of the knowledge – engagement that is most characteristically reflected in time spent in practice or in tentative, experimental use. A second general point complements and interprets the first: there is an alternative path to tacit knowledge that does not really require its transfer at all: reinvention. Available articulated knowledge (the 'how to' book) plus substantial effort devoted to trial- and-error learning plus a reasonably clear success criterion or feedback mechanism will often suffice to enable the learner to acquire the needed tacit knowledge at a satisfactory level. When deliberate transfer mechanisms are employed, such as authoritative demonstration by a master, critique by a coach, or the efforts of a corporate tech-transfer team, the required engagement of the learner generally involves an element of such effort at reinvention.

There are general implications of strategic significance: the more tacit the knowledge involved, the greater the diversity in performance one should expect to see. It is not like the highly uniform execution you would expect among the diverse holders of your phone number. When the requisite knowledge is largely tacit, the dependence on the engagement of the learner implies that variance contributed by the diversity of the learners – involving their motivations, efforts, coaches, physical attributes, local circumstances and random choices – can feed forward into the quality and details of the performances. Also, while tacitness by itself is not a major barrier to imitation (see the driving skill example again), a strong barrier can be established when there is tacitness plus idiosyncrasy (or 'rareness') plus a degree of complexity and a semblance of secrecy (there is low 'observability in use'). There is then little chance that somebody can get out the door of a company with its crucial competitive secret on a Post-it in his pocket.

The following suggestions are offered for the reader who might wish to pursue issues that are only touched upon in this brief article. Polanyi (2009) provides a concise treatment of the concept and its place in his philosophy, which of course is covered much more extensively in the classic source, Polanyi (1962). The sources of tacitness are discussed also by Nelson and Winter (1982: 80–82) and Nightingale (2003). Relationships and contrasts between tacit and articulate or codified knowledge are featured in Nonaka and Takeuchi (1995), Cowan et al. (2000), Nightingale (2003), and Zollo and Winter (2002). Implications of tacitness in the realms of technology and business competition are discussed in these works and also in Bohn (2005) and Kogut (2008).

See Also

- ▶ Competitive Heterogeneity
- ▶ Knowledge Articulation
- ▶ Technology Transfer

References

Bohn, R.E. 2005. From art to science in manufacturing: The evolution of technological knowledge. *Foundations and Trends in Technology, Information and Operations Management* 1: 1–82.

Cowan, R., P. David, and D. Foray. 2000. The explicit economics of knowledge codification and tacitness. *Industrial and Corporate Change* 9: 211–253.

Kogut, B. 2008. *Knowledge, options and institutions*. Oxford: Oxford University Press.

Nelson, R.R., and S.G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.

Nightingale, P. 2003. If Nelson and Winter are only half right about tacit knowledge, which half? A Searlean critique of 'codification'. *Industrial and Corporate Change* 12: 149–183.

Nonaka, I., and H. Takeuchi. 1995. *The knowledge creating company*. New York: Oxford University Press.

Polanyi, M. [1958] 1962. *Personal knowledge: Towards a post-critical philosophy*. Chicago: University of Chicago Press.

Polanyi, M. [1966] 2009. *The tacit dimension*. Chicago: University of Chicago Press.

Zollo, M., and S.G. Winter. 2002. Deliberate learning and the evolution of dynamic capabilities. *Organization Science* 13: 339–351.

T

Taylor, Frederick Winslow (1856–1915)

Hindy Lauer Schachter
New Jersey Institute of Technology, School of Management, Newark, NJ, USA

Frederick Winslow Taylor was born into a wealthy Pennsylvania family in 1856. His supervision of factory workers at Midvale Steel

Company and the Manufacturing Investment Company (a paper products firm) convinced him that managers needed more information on how to define a fair day's standard for tasks under their jurisdiction and that such information would emerge from time study, that is, time and motion-based experiments on the individual components of factory tasks. Earlier attempts to increase production efficiency relied on traditional job times. Taylor did experiments to show how long workers took to do a given task if they used different tools and materials. These experiments allowed him to estimate how long a trained worker should take to complete tasks using the best approach.

With time study Taylor put new responsibilities on managers' shoulders: they would have to conduct experiments, learn how to make changes to increase productivity and train workers in the best methods. Taylor wanted worker–manager communication to improve through clear performance feedback. In addition, introduction of time study required an organizational shift through the creation of a centralized planning department. This unit served as a repository of information garnered through experimentation. The planning department archives and forms would supplant the line supervisor's memory as the source of preferred work methods.

Taylor (1947: 36) believed that his approach supplied management with 'well recognized, clearly defined and fixed principles'. Eventually he thought it would spark manager–worker cooperation as both entities gained from production increases (Taylor 1947). To hasten that end, after 1893 Taylor used time study as a consultant at Bethlehem Steel Company, Simmonds Rolling Machine Company, Cramps Ship and Engine Building Company, and other organizations. He applied experimentation to tasks such as shovelling coal or loading pig iron that many people probably thought were too simple to require any study at all. In the over 200 American businesses that accepted some of Taylor's ideas between 1901 and 1915, employee wages tended to increase while fewer decisions depended on the personal biases of supervisors (Nelson 1992).

Taylor published during the Progressive Era, a time running approximately from 1895 to 1915, when middle-class reformers were trying to increase efficiency and honesty in government through such devices as using civil service exams to appoint administrators. Taylor's idea that successful management required a kind of scientific expertise aligned him with the early twentieth-century American Progressive-Era search to use knowledge to improve efficiency in the private and public sectors (Haber 1973). For that reason Progressive lawyer Louis Brandeis (who would subsequently become a United States Supreme Court Justice) labelled the theory 'scientific' management rather than using Taylor's original 'shop' management terminology – a change that Taylor liked and that has come to define his ideas (Savino 2009). In fact, a key Taylor work is called *The Principles of Scientific Management* (Taylor 1947).

Taylor's ideas always excited controversy. In the late twentieth century, scholars still debated the role time study actually played in organizations. Braverman (1974) argued that scientific management stayed the development of democratic workplaces by transferring the hard-won knowledge of line workers to planning departments. This point of view echoes early twentieth-century International Association of Machinist arguments that time study degrades skilled work.

Other contemporary analysts have responded to Braverman's argument by noting that scientific management actually enhanced workplace humanization because it made expertise rather than managerial whim the foundation of decision-making (Nyland 1989; Schachter 1989). Historical evidence that many early twentieth-century Progressives took this view includes the pattern of turn-of-the-century social reformers such as Josephine Goldmark and Mary van Kleeck publicizing shop management insights. When Dartmouth University's Harlow Pearson set up the Taylor Society after Taylor's death in 1915, a number of reformers joined the organization, seeing it as a natural home for Progressive change.

It is generally agreed that Taylor is an important figure in management history. ▶ Peter Drucker (1954) even went so far as to label scientific management the most important American contribution to Western thought since the late eighteenth-century Federalist papers.

See Also

- ▶ Business History
- ▶ Drucker, Peter: The Drucker Strategic Management System

References

Braverman, H. 1974. *Labor and monopoly capital: The degradation of work in the twentieth century*. New York: Monthly Review Press.

Drucker, P. 1954. *The practice of management*. New York: Harper & Row.

Haber, S. 1973. *Efficiency and uplift: Scientific management in the progressive era, 1890–1920*. Chicago: University of Chicago Press.

Nelson, D. 1992. Scientific management in retrospect. In *A mental revolution: Scientific management since Taylor*, ed. D. Nelson. Columbus: Ohio State University Press.

Nyland, C. 1989. *Reduced worktime and the management of production*. Cambridge: Cambridge University Press.

Savino, D., and D. Louis. 2009. Brandeis and his role promoting scientific management as a progressive movement. *Journal of Management History* 15: 38–49.

Schachter, H.L. 1989. *Frederick Taylor and the public administration community: A reevaluation*. Albany: State University of New York Press.

Selected Works

[1903] 1947. *Shop management*. New York: Harper & Brothers.

[1911] 1947. *The principles of scientific management*. New York: Harper & Brothers.

[1912] 1947. *Testimony before the Special House Committee*. New York: Harper & Brothers.

Technical Fitness

Constance E. Helfat
Dartmouth College, Dartmouth, NH, USA

Definition Technical fitness denotes how effectively a capability performs its intended function when normalized by its cost. Thus, technical fitness has two independent dimensions: quality and cost. Technical fitness applies to both dynamic and non-dynamic capabilities, and is an intermediate measure of performance that, in turn, may affect the environmental fitness of capabilities.

Technical fitness denotes how effectively a capability performs its intended function when normalized by its cost (Helfat et al. 2007). Technical fitness applies to both dynamic and non-dynamic (ordinary or operational) capabilities, and measures the quality of a capability per unit of cost.

The quality of a capability refers to how well a capability performs its intended function. For example, the quality of a dynamic capability for new product development can be measured on a number of dimensions, including how well a new product performs its intended function. A newly developed car may perform better or worse on a number of technical dimensions such as speed and reliability, regardless of whether the product succeeds in the marketplace.

The cost of a capability includes that of creating or acquiring the capability, as well as the cost of utilizing and maintaining it. Capabilities have a strong ‘use-it-or-lose-it’ character, because organizations tend to remember by doing. Thus, firms maintain capabilities by using them.

The cost of a capability and its quality do not have a one-to-one relationship. Some capabilities, such as for just-in-time logistics, can have both high quality (delivery of the correct item when and where needed) and low costs of utilization. Although the costs of creating such a capability may be substantial, when amortized over a large number of units the cost per unit is likely to be low.

Technical fitness can take on only positive values, because both dimensions of technical fitness exceed zero. If a capability does not perform its intended function at all, it is not a capability (Winter 2003; Helfat and Winter 2011) and the concept of technical fitness does not apply. Thus, any measure of how well a capability performs its intended function, or its quality, is strictly positive. In addition, capabilities are rarely costless. Firms incur positive amounts of cost to build (or acquire) and maintain capabilities. Thus, when the quality of a capability, which has a positive value, is normalized by a positive amount of costs, the resulting value is also positive.

Technical fitness also varies across firms. Firms differ both in the cost of a capability and in how well a capability performs its intended function. For example, relative to its competitors, Wal-Mart has a superior logistics capability with respect to the costs of the capability and intended outcomes such as on-time and accurate delivery of products to its stores.

Technical fitness may or may not improve ► environmental fitness, in terms of how well a capability enables a firm to make a living. The latter depends on several factors, including whether competitors have equal or better technical fitness of a capability. In addition, if the technical fitness of a capability does not match the needs of a market, technical fitness will not have a positive effect on environmental fitness. For example, high technical fitness of a capability to utilize obsolete technology is unlikely to enhance environment fitness.

Empirical studies have used the quality dimension of technical fitness as a dependent variable. For example, Henderson and Cockburn (1994) used patent counts to measure the productivity of drug discovery, and Stadler et al. (2013) used the number of successful wells drilled to measure the productivity of upstream oil exploration and development. These studies showed that differences between firms in the attributes of their capabilities predicted differences in technical fitness. In addition, technical fitness can be used as an intermediate measure of capability performance in empirical work seeking to assess the impact of capabilities on measures of firm performance.

See Also

- ► Competitive Advantage
- ► Dynamic Capabilities
- ► Environmental Fitness
- ► Measuring Competence

References

Helfat, C.E., and S.G. Winter. 2011. Untangling dynamic and operational capabilities: strategy for the [n]ever-changing world. *Strategic Management Journal* 32: 1243–1250.

Helfat, C.E., S. Finkelstein, W. Mitchell, M.A. Peteraf, H. Singh, D.J. Teece, and S.G. Winter. 2007. *Dynamic Capabilities: Understanding Strategic Change in Organizations*. Malden: Blackwell.

Henderson, R., and I. Cockburn. 1994. Measuring competence: exploring firm effects in pharmaceutical research. *Strategic Management Journal* 15: 63–84.

Stadler, C., C.E. Helfat, and G. Verona. 2013. The impact of dynamic capabilities on resource access and development. *Organization Science*. https://doi.org/10.1287/orsc.1120.0810.

Winter, S.G. 2003. Understanding dynamic capabilities. *Strategic Management Journal* 24: 991–996.

Technological Change

Jens Frøslev Christensen
Copenhagen Business School, Copenhagen, Denmark

Abstract

Technological change is both a central source of increased productivity and welfare and a source of unwanted side effects on health and environment. Technological change is a key force affecting firms' competitive positions, and, over the past decades, a comprehensive literature has demonstrated that firms face non-trivial strategic challenges when responding to technological change conducted by rivals, and when trying to set the agenda themselves through technological innovation.

Definition Technological change is here defined as change in science and engineering knowledge as applied in products, services and operations processes – encompassing both the material knowledge and the underlying immaterial theoretical and practical knowledge.

Technological change is here defined as change in science and engineering knowledge applied in products, services and operations processes (widely interpreted), both change in material knowledge (e.g., products and machines) and change in the underlying theoretical knowledge, practical skills and routines. Technological ► innovation is a subset of technological change. While technological innovation refers to the endeavours by business enterprises to create and commercialize products, services and processes with new technical features, technological change also refers to change at the level of whole societies, industries or technological fields (e.g., biotechnology), representing aggregate endeavours and outcomes of multiple firms, research institutions and other agents.

In market-based economies business enterprises are key generators and adopters of technological change. Hence technological change represents an important parameter in the strategic management of most business firms. The research and practice of ► management of technology and innovation centrally address the managerial and strategic aspects of how firms generate and adopt new technology.

Technological change is a central source of increased productivity and social welfare. It is also a source of unwanted side effects in the form of environmental degradation or negative health consequences. Economists have demonstrated that the historic rate of economic growth cannot entirely be accounted for by growth in labour and capital, and Solow (1957) proposed that this residual growth indicated technological change. As an economic term, technological change is reflected in both greater volume of output associated with enhanced production processes, and in qualitatively superior output in the form of products or services (Rosenberg 1982). In most indicators of technological change, in particular R&D and patent statistics, product-orientated activities feature much more prominently than process-orientated activities. Nevertheless, most economists tend to treat the phenomenon as if it were solely volume- and productivity-enhancing. Technological progress is measured in terms of increased *market* value of productive output, and negative social and environmental impacts may be due to problems of *market failure*. Thus, in his economic analysis of climate changes, Stern (2007) concludes that man-induced climate changes constitute the widest-ranging market failure ever seen.

Schumpeter, a pioneer in the theory of technological innovation, situated technological change, and in particular radical product innovation conducted by entrepreneurs or large companies, as *the* central driver of economic development. He made a distinction between two stages of technological change: invention and innovation (Schumpeter 1939), to which was later added a third, diffusion. The invention process encompasses the generation of new ideas, the innovation process reflects the development of inventions into marketable products and processes, and in the diffusion stage the new products and processes spread across the potential market. In this phase, technological change tends to be characterized by incremental innovation, whose cumulative productivity impacts are much greater than those incurred by the early radical innovations (Stoneman 1995).

In the aftermath of the Second World War, the invention–innovation–diffusion conception was extended into a framework for understanding the role of science in technological change. In a report presented to President Truman in 1945, Bush (1990) asserted that the principal source of technological development is basic research transferred into applied research and subsequently development and production. This view became known as the *linear model* that for several decades became the prevailing paradigm in both science and technology policy, and in the management of the technology and innovation of companies. The idea of a divide between basic research and R&D was reinforced by the economic argument that basic research is a public good that governments should fund, whereas R&D is close to the market and should be funded by private firms (Stokes 1997, pp. 95–96).

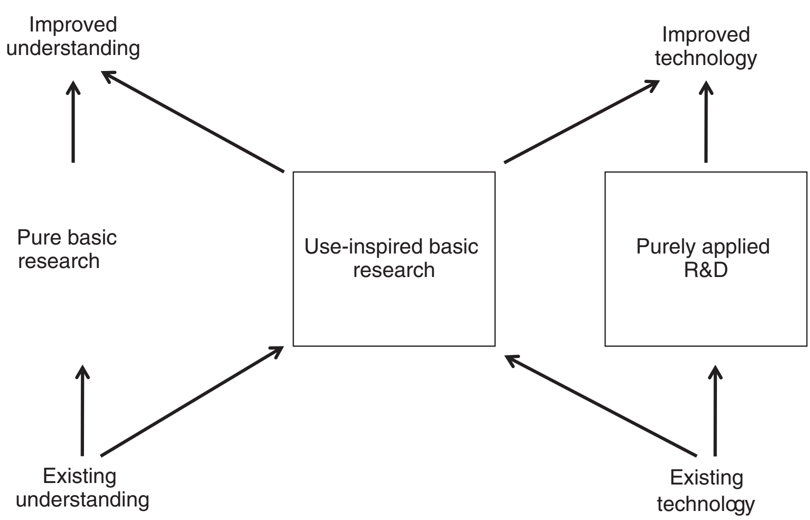
Since the 1980s it has become evident that the linear model gives an incomplete account of technological change in most industries. Thus, for example, the formidable competitive achievements of Japanese automobile and electronics companies since the late 1950s were *not* based on a strong foundation in basic science, but rather on technological/engineering competencies in product and process development. Kline and Rosenberg (1986) offer an alternative and more realistic 'chain-linked' model that allows for technology to be developed without necessarily being based on science inputs, and for technology also to potentially influence science. Christensen (1995) proposes a more differentiated taxonomy of innovative assets than the linear R&D conception and argues that technological innovation may result from different constellations of such assets. Stokes (1997) suggests a distinction between two trajectories, one sustaining basic scientific understanding, another sustaining technological knowhow. These trajectories are considered to be loosely coupled. 'But each of these trajectories is at times strongly influenced by the other, and this influence can move in either direction, with use-inspired basic research often cast in the linking role' (p. 87) (see also Fig. 1).

The nature of technological change has been further explored in the concepts of technological paradigms and trajectories. Dosi (1984) defines a *technological paradigm* as 'a model and a "pattern" of solution of selected technological problems, based on selected principles derived from natural sciences and selected material technologies' (p. 14). Technological paradigms align technological and economic types of problems with scientific or engineering principles of search behaviour (Dosi 1988, pp. 1127). *Technological trajectories* are both constrained and enabled by their corresponding technological paradigms, and '. . . technical progress is largely endogenously driven by a competitive process whereby firms continuously try to improve on their basic technologies and artefacts' (Dosi 1988, pp. 1142). Pavitt (1984) has further developed a taxonomy of firm-specific innovation and technological trajectories.

This perspective emphasizes the idea that technological change is a key force affecting firms' competitive positions. Over the past three decades a comprehensive literature has addressed the particular strategic and managerial challenges that firms face both when responding to technological change conducted by rivals, and when trying to set the agenda themselves through technological innovation. Burgelman et al. (2004) and Storey (2004) offer authoritative selections of articles on strategic management of technology and innovation. Some important insights from the literature on management of technology and innovation shall be briefly mentioned.

User-Centric Innovation

While the views on technological change in the Schumpeterian tradition tend to focus on science- or technology-based dynamics, the



Technological Change, Fig. 1 Stoke's dynamic model of science-technology interplay (Source: Stokes 1997, p. 88) central role of users has also increasingly become acknowledged. Von Hippel (2005) and other scholars have studied various ways in which users – especially so-called lead-users – contribute to technological innovation. However, Christensen and Raynor (2003) point to the dilemma that firms relying primarily on innovative inputs from mainstream customers may miss out on opportunities for disruptive innovations.

Appropriability

Appropriability refers to the factors that govern the innovator's ability to capture profits from technological innovation, most importantly the efficiacy of intellectual property rights protection and the nature of the technology (e.g., its complexity). Teece (1986) pioneered the strategic analysis of appropriability, and one of his key propositions was that access to complementary assets, for example, marketing and operations, is particularly critical for technological innovators operating in weak 'appropriability regimes'. In conventional economic and strategic thinking, firms' incentives to invest in technological innovation diminish if appropriability conditions allow easy imitation. However, this position has been modified by recent research on more open innovation strategies (Pisano 2006; West 2006).

Technology Cycles

In the 'classical' strategy theory of technological change, firms and industries are subject to long periods of cumulative technological trajectories punctuated by periods of radical breakthroughs as new technologies evolve and eventually come to replace the old (Utterback 1994). Anderson and Tushman (1990) distinguish between *competence-enhancing* and *competence-destroying* new technology and find that incumbents embedded in old technology face great difficulties in responding to the latter. This proposition has later been challenged by, among others, Henderson and Clark (1990), Pavitt (1998), and Christensen and Raynor (2003).

Markets for Technology

Although markets for technology have generally been considered to be subject to substantial imperfections, there is mounting evidence that such markets have become increasingly common. This has important implications for the strategic management of firms. Thus, firms can choose to license in the technology instead of developing it in-house or they can license out their technology. Well-functioning markets for technology promote the diffusion of existing technology and encourage specialized technology-based entrepreneurship (Arora et al. 2001).

The Systemic and Open Nature of Technological Change

In recent years, there has been increasing emphasis on understanding the multi-technology and systemic nature of technological change in many industries (e.g., automobiles and software). This issue has been investigated under the headings of modular technological innovation (Sanchez and Mahoney 1996), systems integration (Prencipe et al. 2003) and platform strategy (Gawer and Cusumano 2002). Moreover, since the sources of technological change have become increasingly distributed across nations, firms and knowledge institutions, technological change must increasingly be studied as interorganizational rather than primarily firm-specific processes – and the strategic ramifications hereof are currently being explored under the theme of open innovation in both scholarly work (Chesbrough et al. 2006, 2013) and in practice. The themes of user-centric innovation and markets for technology also reflect more open forms of conducting technological innovation.

See Also

- ▶ Innovation
- ▶ Management of Technology

T

References

Anderson, P., and M.L. Tushman. 1990. Technological discontinuities and dominant designs: A cyclical model of technological change. *Administrative Science Quarterly* 35: 604–633.

Arora, A., A. Fosfuri, and A. Gambardella. 2001. Markets for technology and their implications for corporate strategy. *Industrial and Corporate Change* 10: 419–451.

Burgelman, R.A., C.M. Christensen, and S.C. Wheelwright. 2004. *Strategic management of technology and innovation*, 4th ed. Boston: McGraw-Hill Irwin.

Bush, V. 1945/1990. *Science: The endless frontier*. Washington, DC: National Science Foundation.

Chesbrough, H., W. Vanhaverbeke, and J. West (eds.). 2006. *Open innovation: Researching a new paradigm*. Oxford: Oxford University Press.

Chesbrough, H., W. Vanhaverbeke, and J. West (eds.). 2013. *Exploring next wave of open innovation research*. Oxford: Oxford University Press.

Christensen, J.F. 1995. Asset profiles for technological innovation. *Research Policy* 24: 727–745.

Christensen, C.M., and M. Raynor. 2003. *The innovator's solution*. Boston: Harvard Business School Press.

Dosi, G. 1984. *Technical change and industrial transformation*. London: Macmillan.

Dosi, G. 1988. Sources, procedures, and microeconomic effects of innovation. *Journal of Economic Literature* 26: 1120–1171.

Gawer, A., and M.A. Cusumano. 2002. *Platform leadership*. Boston: Harvard Business School Press.

Henderson, R., and K.B. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35: 9–30.

Kline, S.J., and N. Rosenberg. 1986. An overview of innovation. In *The positive sum strategy: Harnessing technology for economic growth*, ed. R. Landau and N. Rosenberg. Washington, DC: National Academy Press.

Pavitt, K. 1984. Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy* 13: 343–373.

Pavitt, K. 1998. Technologies, products and organization in the innovating firm: What Adam Smith tells us and Joseph Schumpeter doesn't. *Industrial and Corporate Change* 7: 433–452.

Pisano, G. 2006. Profiting from innovation and the intellectual property revolution. *Research Policy* 35: 1122–1130.

Prencipe, A., A. Davies, and M. Hobday (eds.). 2003. *The business of systems integration*. Oxford: Oxford University Press.

Rosenberg, N. 1982. *Inside the black box: Technology and economics*. Cambridge: Cambridge University Press.

Sanchez, R., and J.T. Mahoney. 1996. Modularity, flexibility and knowledge management in organizational design. *Strategic Management Journal* 17: 63–76.

Schumpeter, J. 1939. *Business cycles: A theoretical, historical, and statistical analysis of the capitalist process*. New York: McGraw-Hill.

Solow, R.M. 1957. Technical change and the aggregate production function. *Review of Economics and Statistics* 39: 312–320.

Stern, N. 2007. *The economics of climate change: The stern review*. Cambridge: Cambridge University Press.

Stokes, D.E. 1997. *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.

Stoneman, P. 1995. *The handbook of economics of innovation and technological change*. Cambridge, MA: Blackwell.

Storey, J. (ed.). 2004. *The management of innovation*, vol. 1 & 2. Cheltenham: Edward Elgar.

Teece, D.J. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* 15: 285–305.

Utterback, J.M. 1994. *Mastering the dynamics of Innovation: How companies seize opportunities in the face of technological change*. Boston: Harvard Business School Press.

Von Hippel, E. 2005. *Democratizing innovation*. Cambridge, MA: The MIT Press.

West, J. 2006. Does appropriability enable or retard open innovation? In *Open innovation: Researching a new paradigm*, ed. H. Chesbrough, W. Vanhaverbeke, and J. West. Oxford: Oxford University Press.

Technological Inertia

Mary Tripsas
Boston College, Carroll School of Management, Chestnut Hill, MA, USA

Definition Technological inertia is the propensity of incumbent firms with historical expertise in one generation of technology to continue development of that generation and not effectively develop and commercialize products based on a new generation of technology.

Technological inertia is the propensity of incumbent firms with expertise in one generation of technology to continue development of that generation and not effectively develop and commercialize products based on a new generation of technology. The phenomenon has been well documented empirically in a range of industries, including the movement from X-ray to CT scanners, electromechanical to electronic calculators, and mechanical to electronic watches among others (Cooper and Smith 1992).

Research on the underlying drivers of technological inertia has been at two levels of analysis. First are longitudinal industry studies focused on developing typologies of ▶ technological change that help explain the conditions under which incumbents fail. Second are studies that look inside the firm, attempting to discern what organizational factors contribute to technological inertia.

Technological inertia can be thought of as a special case of organizational inertia – one that is exacerbated by the somewhat deterministic nature of technological progress in an industry, which, within a particular technological regime, follows a 'natural trajectory,' in which 'advances seem to follow advances in a way that appears somewhat "inevitable"' (Nelson and Winter 1977: 56). These periods of incremental innovation within a particular regime are punctuated by technological discontinuities that eventually culminate in a new technological regime.

With this pattern as a backdrop, scholars have focused on understanding what type of technological discontinuity results in incumbent failure. In one of the most influential studies in this tradition, Tushman and Anderson (1986) found that it did not matter how radical a new technology was in terms of performance improvement. Instead, 'competence-destroying' technology, which required fundamentally different scientific and engineering expertise, was most difficult for incumbents. In an important extension of this work, Henderson and Clark (1990) found that technological inertia was strong even for seemingly minor innovations that were architectural, since they required new architectural knowledge.

Moving beyond the role of a firm's technological capabilities, Christensen and Bower (1996) categorized new technology by its effect on customers. Disruptive technology, which appealed to a different customer segment with different purchase criteria, caused incumbents to fail, not because they couldn't develop new technological expertise, but because they were focused on existing customers and overlooked the new segment. In a similar vein, Tripsas (1997) found that, if specialized complementary assets for commercialization retained their value, incumbents were less likely to fail, even if a new technology was competence-destroying.

Looking inside the firm, technological inertia has been examined through both behavioural and cognitive lenses. Behavioural explanations emphasize the stickiness of organizational routines and procedures associated with a particular technological regime (Nelson and Winter 1982). Organizations are more likely to search locally for technological solutions, and, as a result, they fail to notice and develop distant new technologies. The nature of learning also leads to ▶ competency traps, which, along with myopia, result in behaviours that promote incremental change within an existing technological regime (Levinthal and March 1993).

Cognitive explanations for technological inertia have emphasized the rigidity of historical mindsets and how they constrain interpretations of new technology. Top managers do not always notice new technology, limiting their ability to respond (Kaplan 2008). Even when they do notice and invest, prior cognitions are difficult to change. For instance, Tripsas and Gavetti (2000) show how the mindset of Polaroid's top management team resulted in a focus on digital imaging technologies that used an outdated instant photography business model. Promising avenues for future research on technological inertia involve the integration of behavioural and cognitive perspectives as well as multiple levels of analysis.

See Also

▶ Cognition and Technical Change
▶ Competency Trap
▶ Disruptive Technology
▶ Dominant Design
▶ Innovation
▶ Management of Technology
▶ Radical and Incremental Technical Change
▶ Technological Change
▶ Technological Paradigms and Technological Trajectories
▶ Technology Cycles

References

Christensen, C.M., and J.L. Bower. 1996. Customer power, strategic investment, and the failure of leading firms. *Strategic Management Journal* 17: 197–218.

Cooper, A.C., and C.G. Smith. 1992. How established firms respond to threatening technologies. *Academy of Management Executive* 6: 55–70.

Henderson, R., and K.B. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35: 9–30.

Kaplan, S. 2008. Cognition, capabilities, and incentives: Assessing firm response to the fiber-optic revolution. *Academy of Management Journal* 51: 672–695.

Levinthal, D.A., and J.G. March. 1993. The myopia of learning. *Strategic Management Journal* 14 (special issue): 95–112.

Nelson, R.R., and S. Winter. 1977. In search of a useful theory of innovation. *Research Policy* 6: 36–76.

Nelson, R.R., and S. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.

Tripsas, M. 1997. Unraveling the process of creative destruction: Complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal* 18: 119–142.

Tripsas, M., and G. Gavetti. 2000. Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal* 21: 1147–1161.

Tushman, M.L., and P.C. Anderson. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* 31: 439–465.

Technological Paradigms and Technological Trajectories

Giovanni Dosi[1] and Richard R. Nelson[2]

[1]Scuola Superiore Sant'Anna, Pisa, Italy

[2]Columbia University, New York, NY, USA

Abstract

The notions of technological paradigms and technological trajectories are central to the interpretation of innovation as an evolutionary process and to the understanding of invariances in the knowledge structure and in the ways technological knowledge accumulates and, together, what distinguishes different fields and different periods of technological advance.

Definition Technological paradigms comprise specific knowledge bases building on selected chemical or physical principles, problem-solving procedures, search heuristics and often also some 'dominant design' of the artefacts produced on grounds of the paradigm itself. Trajectories map the relatively ordered patterns of advance in the techno-economic characteristics of products and in the efficiencies in inputs use. Industries evolve as heterogeneous firms, explore with different degrees of success the innovative opportunities entailed by each paradigm, and compete with each on the markets.

As discussed at greater length in Dosi and Nelson (2010), each technology needs to be understood as comprising (a) a specific body of practice – in the form of processes for achieving particular ends – together, of course, with an ensemble of required artefacts on the 'input side'; (b) quite often some distinct notion of a design of a desired 'output' artefacts; and, (c) a specific body of understanding, some relatively private, but much of it shared among professionals in a field. These elements, together, can be usefully considered as constituent parts of a *technological paradigm* (Dosi 1982, 1988), somewhat analogous with Thomas Kuhn's scientific paradigm (Kuhn 1962).

A paradigm embodies an *outlook*, a definition of the relevant problems to be addressed and the patterns of enquiry in order to address them. It entails a view of the purported needs of the users and the attributes of the products or services they value. It encompasses the scientific and technical principles relevant to meeting those tasks, and the specific technologies employed. A paradigm entails *specific patterns of solution to selected techno-economic problems* – that is, specific families of recipes and routines – based on highly selected principles derived from natural sciences, jointly with specific rules aimed at acquiring related new knowledge. Together, the paradigm includes a (generally imperfect) understanding about just how and (to some extent) why prevailing practice works.

An important part of paradigmatic knowledge takes the form of *design concepts*, which characterize in general the configuration of the particular artefacts or processes that are operative at any time. Shared general design concepts are an important reason why there is often strong similarity among the range of particular products manufactured at any time – for example, the large passenger aircraft produced by different aircraft companies or the different television sets available at the electronics stores. Indeed, the establishment of a given technological paradigm is quite often linked with the emergence of some *dominant design* (Abernathy and Utterback 1978; Rosenbloom and Cusumano 1987; Henderson and Clark 1990; Utterback and Suarez 1993; Suarez and Utterback 1995; and the critical review of the whole literature in Murmann and Frenken 2006). A dominant design is defined in the space of artefacts and is characterized both by a set of core design concepts embodied in components that correspond to the major functions performed by the product and by a product architecture that defines the ways in which these components are integrated (Murmann and Frenken 2006; drawing upon Henderson and Clark 1990). However, sometimes the establishment of a dominant paradigm is *not* associated with a dominant design. A revealing case in point is pharmaceutical technologies which *do* involve specific knowledge basis, specific search heuristics and so on – that is, the strong mark of paradigms – without, however, any hint at any dominant design. Molecules, even when aimed at the same pathology, might have quite different structures: in that space, one is unlikely to find similarities akin those linking even a Volkswagen Beetle 1937 and a Ferrari 2000. Nevertheless, the notion of 'paradigm' holds in terms of the underlying features of knowledge bases and search processes.

Whether or not the establishment of a dominant paradigm entails also the establishment of a dominant design is of importance in terms of the dynamics of industry structure along the lifecycle of the industries to which a particular paradigm is associated.

Technological paradigms identify the operative constraints on prevailing best practice and the *problem-solving heuristics* deemed promising for pushing back those constraints. More generally, they are the cognitive frames shared by technological professionals in a field, that orient what such professionals think they can do to advance a technology (Constant 1980). Technological paradigms also encompass normative aspects, such as criteria for assessing performance, and thus provide ways of judging what is better as well as identifying goals for the improvement of practice. Each paradigm involves a specific 'technology of technical change', that is, specific *heuristics of search*. So, in sectors such as organic chemicals for example, these heuristics relate to the ability of coupling basic scientific knowledge with the development of molecules that present the required characteristics, while in pharmaceuticals the additional requirement is the ability to match molecular knowledge with receptors and pathologies. In microelectronics, search concerns methods for further miniaturization of electrical circuits, the development of the appropriate hardware capable of 'writing' semiconductor chips at such a required level of miniaturization and advances in the programming logic to be built into the chip. There are numerous examples: some are discussed in Dosi (1988). Here notice in particular that distinct (paradigm-specific) search and learning procedures first imply diverse modes of creating and accessing novel technological opportunities, and, second, entail different organizational forms suited to such research procedures. As we shall see, both properties will turn out to be central when trying to characterize distinct 'regimes' of technological evolution.

Together, the foregoing features of technological paradigms provide a focus for efforts to advance a technology and channel them along distinct *technological trajectories*, with advances (made by many different agents) proceeding over significant periods of time in certain relatively invariant directions, in the space of techno-economic characteristics of artefacts and production processes. As paradigms embody the identification of the needs and technical requirements of users, trajectories may be understood in terms of the progressive refinement and improvement in the supply responses to such potential demand requirements.

A growing number of examples of technological trajectories include aircraft, helicopters, various kinds of agricultural equipment, cars, semiconductors (Gordon and Munson 1981; Sahal 1981, 1985; Dosi 1984; Grupp 1992; Saviotti and Trickett 1992; Saviotti 1996). So, for example, technological advances in aircraft technologies have followed two distinct trajectories (one civilian and one military) characterized by log-linear improvements in the trade-offs between horsepower, gross take-off weight, cruise speed, wing load and cruise range (Sahal 1985; Frenken et al. 1999; Frenken and Leydesdorff 2000; Giuri et al. 2007; and, more specifically on aircraft engines, Bonaccorsi et al. 2005). Analogously, in microelectronics, technical advances are accurately represented by an exponential trajectory of improvement in the relationship between density of electronic chips, speed of computation and cost per bit of information (see Dosi 1984, but the trajectory has persisted since then). In fact, it is fair to say that trajectory-like patterns of technological advance have generally been found whenever the analyst bothered to plot over time the fundamental techno-economic features of discrete artefacts or processes, say from the DC3 to the Airbus 380, among aircrafts; or from crucible to Bessemer to basic oxygen reduction among steel-making processes. (Admittedly, trajectories in the space of processes and related input intensities have been studied much less than trajectories in the output characteristic space, and this is indeed a challenging research area ahead.)

The emergence of relatively ordered trajectories, as already suggested, is not always associated with the emergence of dominant designs. When it is, the trajectories appear to be driven by 'hierarchically nested technological cycles' entailing both relatively invariant core components improving over time and a series of bottlenecks and 'technological imbalances' (Rosenberg 1976) regarding the consistency among all the components of the systems (cf. Murmann and Frenken 2006). Some properties of trajectories are important to note.

First, trajectories *order* and *confine* but do not at all eliminate the persistent *generation* of *variety*, in the product and process spaces that innovative search always produces. The paradigm defines proximate boundaries of feasibility and together shapes the heuristics of search. However, there continues to be plenty of possible trade-offs between output characteristics, which different producers explore (Saviotti 1996) and which will be eventually the object of (imperfect and time-consuming) market selection.

Second, by the same token, trajectories 'extrapolated forward' – in so far as their knowledge is shared by the community of firms, practitioners, engineers – are powerful *uncertainty-reducing representations* of what the future is likely to yield in technological terms. However, this remains a far cry from any unbiased expectation on the time and costs involved in 'getting there' – wherever 'there' means – and, even more so, of the probability distributions of individual actors over both technological and economic success. That is, trajectories are not means to reduce Knightian uncertainty into probabilizable risk. Indeed, notwithstanding roughly predictable trajectories of advance, both *substantive uncertainty* – concerning future states of the world – and *procedural uncertainty* – regarding future problem-solving procedures – continue to be ubiquitous.

Note that there is no a priori economic reason why one should observe limited clusters of technological characteristics at any one time and ordered trajectories over time. On the contrary, as we have already argued in Dosi (1988) – given consumers with different preferences and equipment users with different technical requirements and different relative prices over different countries, if technologies were perfectly 'plastic' and malleable – as standard economic representations are implicitly suggesting – one would tend to observe sorts of 'isoquants' in the space of techniques and technoeconomic characteristics and products with the familiar shape. And, over time, if technological recipes – both in the procedural aspects and their input contents – could be freely added, divided, recombined, substituted and so on, one would also tend to observe an increasingly disperse variety of technical and performance combinations in products, production inputs and available techniques (even if not necessarily in their use, given relative prices). The ubiquitous evidence on trajectories, on the contrary, suggests that technological advances are circumscribed within a quite limited subset of the techno-economic characteristics space. We could say that the paradigmatic, cumulative nature of technological knowledge provides *innovation avenues* (Sahal 1985) which channel technological evolution, while major discontinuities tend to be associated with changes in paradigms. Indeed, here and throughout what we shall call 'normal' technical progress those advances occur along a given trajectory – irrespective of how 'big' they are and how fast they occur – while we reserve the name of 'radical innovations' for those innovations linked with paradigm changes.

A change in the paradigm generally implies a change in the trajectories. Together with different knowledge bases and different prototypes of artefacts, the techno-economic dimensions of innovation also vary. Some characteristics may become easier to achieve, new desirable characteristics may emerge, some others may lose importance. Relatedly, the engineers' vision of future technological advances will change, together with a changing emphasis on the various trade-offs that characterize the new artefacts. So, for example, the technological trajectory in active electrical components based on thermionic valves had as fundamental dimensions heat-loss vacuum-parameters, miniaturization and reliability over time. With the appearance of solid state components (the fundamental building block of the microelectronic revolution) heat loss became relatively less relevant, while miniaturization increased enormously in importance. Similar examples of change in the dimensions of the design space can be found in most transitions from one paradigm to another.

Are there some features which most technological trajectories share? A common feature which characterizes trajectories in process technologies and in the related equipment-embodied technologies is a powerful trend towards mechanization and/or automation of production activities. Recent pieces of evidence are in Klevorick and colleagues (1995), but the phenomenon has been noticed in the classical literature and plays an important role in the analyses of the dynamics of capitalist economies by Adam Smith and Karl Marx. Note that such a tendency holds across sectors and across countries characterized by different capital intensities, and broadly occurs irrespective of variations in relative prices. Due to its generality, in another work (Nelson and Winter 1977) it was called a 'natural trajectory': of course, there is nothing 'natural', strictly speaking, but it is indeed a general reflection of a long term trend towards the substitution of inanimate energy to human and animal efforts, and, more recently, also of inanimate information processing to human cognition and control.

There is another relatively common feature of trajectories of innovation (even if we still do not know how common – a task for future empirical research), namely *learning curves* (see Thompson 2010). This original statement of the 'law' comes from Wright (1936), in reference to aircraft manufacturing (see also Alchian 1963). Similar regularities appear in various energy-producing technologies, in computers, light bulbs, and many other artefacts and processes: for technology-specific evidence and surveys see Conley (1970), Baloff (1971), Dutton and Thomas (1984), Gritsevskyi and Nakicenovic (2000), MacDonald and Schrattenholzer (2001), Neij (1997), Yelle (1979), Argote and Epple (1990) and Thompson (2010). Semiconductors offer an archetypical example of a trajectory driven by miniaturization efforts yielding the so-called Moore's Law involving the doubling of the density of elementary transistor-per-chip and later microprocessors every 2–3 years (more details in Gordon and Munson 1981; Dosi 1984; Jovanovic and Rousseau 2002; Nordhaus 2007).

Interestingly, a steady fall in unit labour inputs seems, at least in some circumstances, to appear even when holding the equipment constant. This is the so-called Horndahl effect, named after a Swedish steel mill (Lundberg 1961), an effect which contributed to Arrow's (1962) inspiration for the concept of learning by doing. Notice that learning effects appear at the levels of industry, firms and plants, even if distinct rates and intertemporal variabilities with micro learning displaying higher irregularities over time than industry-level rates of progress (for some discussion of the evidence see Auerswald et al. 2000). The interpretation of learning mechanisms underlying the observed performance trajectories of their differences across different paradigms are important research tasks for future evolutionary analyses of innovation.

Together with differences across paradigms in the rates of technological advance, one observes major differences in the processes through which such advances occur. In fact, significant progress has been made in the conceptualization of what different technological paradigms have in common and how they differ in terms of the sources of knowledge upon which they draw – the *technological opportunities* which they tap – the mechanisms through which such opportunities are seized, and the possibilities they entail for innovators to extract economic benefit from their technological advances – that is, the *appropriability conditions*.

Technological Opportunities, The Processes of Knowledge Accumulation and Their Cumulativeness

Prevailing technological paradigms differ over time and across fields regarding the nature of the knowledge underlying the opportunities for technical advances. Relatedly, they differ in the extent to which such knowledge has been gained through operating experience, as opposed to scientific research.

While in most fields there is a mix, in the fields generally thought of as 'high tech' a more significant contribution is nowadays grounded in the specialized fields of science or engineering.

Where operating experience and learning by doing and using are the primary bases for professional understanding, the learning trajectory will advance, paced by experience with actual new designs (and nowadays with the advances incorporated into new vintages of capital equipment and the ability of using these). On the other hand, understanding can advance rapidly when there are fields of science dedicated to that objective. Several studies (see, e.g., Klevorick et al. 1995; Nelson and Wolff 1997) have shown that the fields of technology which, by a variety of measures, have advanced most rapidly are associated with strong fields of applied science or engineering. Moreover, firms operating in these fields also tend to have levels of R&D intensity that are higher than average. In fact, in a secular perspective, the evidence is in tune with Mokyr's general conjecture that the 'epistemic' elements of technological knowledge – that is, those elements associated with an explicitly casual knowledge of natural phenomena – are of crucial (and increasing) importance in modern technological advances (Nelson and Wolff 1997; Mokyr 2002, 2010; Nelson and Nelson 2002; Nelson 2003).

Since the Industrial Revolution, the contribution of science to technology has been increasing, and, in turn, such a science base has been largely the product of publicly funded research, while the knowledge produced by that research has been generally open and available for potential innovation to use (more in David 2001a, b, 2004; Pavitt 2001; Nelson 2004).

This, however, is not sufficient to corroborate any simple 'linear model' from pure to applied science to technological applications.

First, the point made elsewhere by Rosenberg (1982), Kline and Rosenberg (1986), Pavitt (1999) and Nelson (1981) continues to apply: scientific principles are helpful but are rarely enough. An enlightening case in a 'science-based' area – medical innovation – is discussed in Rosenberg (2009). Semiconductors technology is another good example. For many decades, efforts to advance products and process technology – crucially involving the ability to progressively make circuits smaller and smaller – have taken advantage of the understandings in material science and the underlying solid state physics. However, much more pragmatic and tacit elements of technological know-how have been persistently crucial.

Second, it is quite common that scientific advances have been made possible by technological ones, especially in the fields of instruments: think of the example of the electronic microscope with respect to scientific advances in life sciences (more in Rosenberg 1982, 1994).

Third, it is not unusual that technologies are made to work before one understands why they do so: the practical (steam) engine was developed some years before science modelled the theoretical Carnot engine; even more strikingly, the aeroplane was empirically proved to work decades before applied sciences 'proved' that it was theoretically possible. In fact, the specificities of the links between technological advances and advances in applied sciences are a major discriminating factor among different technological paradigms and different sectors (see below on sectoral taxonomies).

Generally speaking, while it usually holds that technological advance tends to proceed rapidly where scientific understanding is strong and slowly where it is weak, the key has often been the ability to design controllable and replicable practices that are broadly effective around what is understood scientifically.

Given potential opportunities for innovation, what are the properties of the processes through which they are tapped? An important feature distinguishing different paradigms relates to the *cumulativeness* of innovative successes. Intuitively, the property captures the degrees to which 'success breeds success', or, in another fashionable expression, the measure to which innovative advances are made by dwarves standing on the shoulders of past giants (as such, possibly, the integral of many dwarves). Cumulativeness captures the incremental nature of technological search, and, crucially, varies considerably across different innovative activities (Malerba and Orsenigo 1996; Breschi et al. 2000; see also below). More formally, a way to capture cumulativeness is in terms of *future probabilities* of success conditional on *past realizations* of the stochastic process. In that respect, it is a widespread instance of *knowledge-based dynamic increasing returns*.

A number of technological paradigms embodying knowledge that is to a large extent generated endogenously tend to display dynamics of knowledge accumulation that are more cumulative than trajectories of advance, which are, so to speak, fuelled 'from outside' (e.g., via the acquisition of new pieces of equipment generated in other industrial sectors). A further distinction concerns the *domain* at which cumulative learning tends to occur: is it at the level of individual firms or is it at the level of the overall community of firms, would-be entrepreneurs, technical communities associated with each paradigms, for example? In Teece and colleagues (1994), examples such as Intel are given, where cumulativeness applies at both paradigm and firm level. At the opposite extreme, many instances point at patterns of technological change which are anti-cumulative in that they imply *competence-destruction* at the level of individual incumbents (cf. Tushman and Anderson 1986). Yet other historical examples highlight discontinuities engendered by firms' specific *organizational diseconomies of scope* even under largely cumulative industry-level patterns of accumulation of technological knowledge: Bresnahan et al. (2008) offer a vivid illustration concerning the introduction of the PC and the browser in the case of IBM and Microsoft, respectively.

Means of Appropriation

Most researchers at universities and public laboratories do their work, which on occasion may result in a significant technological advance, without expectation of benefiting directly from it financially. Some inventors invent because of the challenge and the sense of fulfilment that comes with solving a difficult problem. And, more importantly, in contemporary societies most scientific knowledge – of both the 'pure' and 'applied' nature – is generated within a regime of *open science*. The fundamental vision underlying and supporting such a view of publicly supported open science throughout much of the twentieth century entailed (i) a sociology of the scientific community largely relying on self-governance and peer evaluation, (ii) a shared culture of scientists emphasizing the importance of motivational factors other than economic ones and (iii) an ethos of disclosure of search results driven by 'winner takes all' precedence rules. In

Nelson (2006), David and Hall (2006), and Dosi et al. (2006a) the dangers from the erosion of open science institutions are discussed. We have already mentioned above the importance of (free-flowing) advances in pure and applied sciences as a fundamental fuel for technological advance – albeit with significant variation across technologies, sectors and stages of development for each technological paradigm. However, the major share of inventive activities finalized to economically exploitable technologies that goes on in contemporary capitalist societies is made in profit-seeking organizations with the hope and expectation of being economically rewarded, if that work is successful. In turn, the very existence of a relation between economically expensive search efforts by private agents and (uncertain) economic rewards from successful innovations, entails the fundamental incompatibility – originally pointed out by Marx and Schumpeter – between any sort of zero-profit general equilibrium and any incentive to *endogenous* innovation (that is, endogenous to the private, 'capitalist', sector of the economy).

Granted this is the case, however, two major sets of questions arise.

First, how profound is such a trade-off, if any, between monopolistic departures from competitive (zero profit) conditions and incentives to innovate? More precisely, what is the evidence, if any, on the monotonic relation between (actual and expected) returns from innovation, on the one hand, and innovative efforts, on the other?

Such a monotonic relation is, in fact, built-in as one of a core assumption within most 'neo-Schumpeterian' models of growth, while the limited ability to appropriate returns to invention and innovation is often offered as the reason why the rate of technological progress is very slow in some industries. The aforementioned studies on the nature and sources of technological opportunities suggest that this is unlikely to be the primary reason. The far more likely reason lies in differences in the strength and richness of technological opportunities. More generally, let us suggest that the widespread view that the key to increasing technological progress is in strengthening appropriability conditions, mainly through making patents stronger and wider, is deeply misconceived. Obviously, inventors and innovators must have a reasonable expectation of being able to profit from their work, where it is technologically successful and happens to meet market demands. However, in most industries this is already the case. And there is no evidence that stronger patents will significantly increase the rate of technological progress. (See, further, Mazzoleni and Nelson 1998; Granstrand 1999; Jaffe 2000; Dosi et al. 2006b; and the growing literature cited therein.) In fact, in many instances the opposite may well be the case. We have noted that, in most fields of technology, progress is cumulative, with yesterday's efforts, both failures and successes, setting the stage for today's efforts and achievements. If those who do R&D today are cut off from being able to draw from and build on what was achieved yesterday, progress may be hindered significantly. Historical examples, such as those presented in Merges and Nelson (1994) on the Selden patent around the use of a light fuel in an internal combustion engine to power a car or the Wright brothers' patent on an efficient stabilizing and steering system for flying machines, are good cases in point, showing how the intellectual property rights (IPR) regime probably slowed down considerably the subsequent development of cars and aircrafts, due to the time and resources consumed by lawsuits against the patents themselves. The current debate on property rights in biotechnology suggests similar problems, whereby granting very broad claims on patents might have a detrimental effect on the rate of technical change, insofar as they preclude the exploration of alternative applications of the patented inventions.

This is particularly the case when inventions concerning fundamental techniques or knowledge are concerned, as with genes or the Leder and Stewart patent on the achievement of a genetically engineered mouse that develops cancer. This is clearly a fundamental research tool. To the extent that such techniques and knowledge are critical for further research that proceeds cumulatively on the basis of the original invention, the attribution of broad property rights might severely hamper further developments. Even more so if the patent protects not only the product the inventors have achieved (the 'onco-mouse') but all the class of products that could be produced through that principle, that is, 'all transgenic non-human mammals', or all the possible uses of a patented invention (say, a gene sequence), even though they are not named in the application. In this respect, Murray and colleagues (2009) offer a striking illustration of how 'opening up upstream' (again, in the case of the mouse) – in such an instance, a discrete change in the IPR regime in the US – yielded more search and more diverse rates of exploration of 'downstream' research paths.

In general, today's efforts to advance a technology often need to draw from a number of earlier discoveries and advances which painstakingly build upon each other. Under these circumstances, IPRs are more likely to be a hindrance than an incentive to innovate (see Merges and Nelson 1994; Heller and Eisenberg 1998). If past and present components of technological systems are patented by different parties, there can be an *anti-commons* problem (the term was coined by Heller and Eisenberg). While in the standard commons problem (such as an open pasture) the lack of proprietary rights is argued to lead to overutilization and depletion of common goods, in instances such as biotechnology the risk may be that excessive fragmentation of IPRs among too many owners may well slow down research activities because each owner can block the other. Further empirical evidence on the negative effects of strong patent protection on technological progress is in Mazzoleni and Nelson (1998); and, at a more theoretical level, see the insightful discussion in Winter (1993) showing how tight appropriability regimes in evolutionary environments can deter technical progress (cf. also the formal explorations in Marengo et al. 2009). Conversely, one can document, well before the contemporary movement of 'open source' software, cases in which groups of competing firms or private investors, possibly because of some awareness of the anti-commons problem, have preferred to avoid claiming patents and to deliberately operate in a weak IPR regime somewhat similar to that of open science, involving the free disclosure of inventions to one another (see Allen 1983, and Nuvolari 2004 on blast furnaces and the Cornish pumping engine, respectively). Interestingly, these cases of 'collective invention' have been able to yield rapid rates of technical change. Similar phenomena of free revelation of innovation appear also in the communities of users innovators (see von Hippel 2005).

The *second* set of questions concerns the characteristics of the regimes stimulating and guiding technological advance in a field of activity – that is, *how* inventors appropriate returns. The conventional wisdom has long been that patent protection is the key to being able to appropriate returns. But this is the case only in some fields of technology. Pharmaceuticals is an important example. However, a series of studies (Mansfield et al. 1981; Levin et al. 1985; Cohen et al. 2002, among others) has shown that in many industries patents are not the most important mechanism enabling inventors to appropriate returns. Thus Levin et al. (1985) find that, for most industries, 'lead time and learning curve advantages, combined with complementary marketing efforts, appear to be the principal mechanisms of appropriating returns to product innovations' (p. 33).

Patenting often appears to be a complementary mechanism for appropriating returns to product innovation, but not the principal one in most industries. For process innovations (used by the innovator itself) secrecy is often important, patents seldom so. These findings were largely confirmed by a follow-on study by Cohen et al. (2002). David Teece (1986) and a rich subsequent literature (cf. the special issue of *Research Policy*, 2006; taking stock of the advancements since his original insights) have analysed in some detail the differences between inventions for which strong patents can be obtained and enforced, and inventions where patents cannot be obtained or are weak, in the firm strategies needed for reaping returns to innovation. A basic and rather general finding is that, in many cases, building the organizational capabilities to implement and complement new technology allows high returns to R&D, even when patents are weak. Thus, despite the fact that patents were effective in only a small share of the industries considered in the study by Levin et al. (1985), some three-quarters of the industries surveyed reported the existence of at least one effective means of protecting process innovation, and more than 90% of the industries reported the same regarding product innovations (Levin et al. 1985). These results have been confirmed by a series of other subsequent studies conducted for other countries (see, for example, the PACE study for the European Union; cf. Arundel et al. 1995).

If there are some bottom lines so far to this broad area of investigation, they are that, *first*, there is no evidence on any monotonic relation between degrees of appropriability and propensity to undertake innovative search, above some (minimal) appropriability threshold; *second*, appropriability mechanisms currently in place are sufficient (in fact, possibly over-abundant); *third*, the different rates of innovation across sectors and technological paradigms can hardly be explained by variations in the effectiveness of appropriability mechanisms, and, *fourth*, they can be explained even less by differences in the effectiveness of IPR protection.

See Also

- ▶ Innovation
- ▶ Learning and Adaptation
- ▶ Path Dependence in Technologies and Organizations
- ▶ Radical and Incremental Technical Change
- ▶ Science and Innovation
- ▶ Tacit Knowledge

References

Abernathy, W.J., and J. Utterback. 1978. Patterns of innovation in industry. *Technology Review* 80: 40–47.

Alchian, A. 1963. Reliability of progress curves in airframe production. *Econometrica* 31: 679–693.

Allen, R.C. 1983. Collective invention. *Journal of Economic Behavior & Organization* 4: 1–24.

Argote, L., and D. Epple. 1990. Learning curves in manufacturing. *Science* 247: 920–924.

Arora, A., A. Fosfuri, and A. Gambardella. 2002. *Markets for technology*. Cambridge, MA: The MIT Press.

Arrow, K. 1962. The economic implications of learning by doing. *Review of Economic Studies* 29: 155–173.

Arundel, A., G. van de Paal, and L. Soete. 1995. *Innovation strategies of Europe's largest firms. Results of the PACE survey*, European innovation monitoring system, Report No. 23.. Brussels: European Commission.

Auerswald, P., S. Kaufmann, J. Lobo, and K. Shell. 2000. The production recipe approach to modelling technological innovation: An application to learning by doing. *Journal of Economic Dynamics and Control* 24: 389–450.

Baloff, N. 1971. Extension of the learning curve: Some empirical results. *Operation Research Quarterly* 22: 329–340.

Bonaccorsi, A., P. Giuri, and F. Pierotti. 2005. Technological frontiers and competition in multi-technology sectors: Micro evidence from the aero-engine industries. *Economics of Innovation and New Technology* 14: 23–42.

Breschi, S., F. Malerba, and L. Orsenigo. 2000. Technological regimes and Schumpeterian patterns of innovation. *Economic Journal* 110: 388–410.

Bresnahan, T., S. Greenstein, and R. Henderson. 2008. Schumpeterian competition within computing markets and organizational diseconomies of scope. Working paper, Kellogg School of Management, Northwestern University.

Cohen, W., R.R. Nelson, and J.P. Walsh. 2002. Links and impacts: The influence of public research on industrial R&D. *Management Science* 48: 1–23.

Conley, P. 1970. Experience curves as a planning tool. *IEEE Spectrum* 7: 63–68.

Constant, E. 1980. *The origins of the turbojet revolution*. Baltimore: Johns Hopkins University Press.

David, P.A. 2001a. From keeping nature's secrets to the institutionalization of open science. Discussion Papers in Economic and Social History, University of Oxford.

David, P.A. 2001b. Path dependence, its critics and the quest for 'historical economics'. In *Evolution and path dependence in economic ideas: Past and present*, ed. P. Garrouste and S. Ioannides. Cheltenham: Edward Elgar.

David, P.A. 2004. Understanding the emergence of 'open science' institutions: Functionalist economics in historical context. *Industrial and Corporate Change* 13: 571–589.

David, P.A., and B. Hall. 2006. Property and the pursuit of knowledge: IPR issues affecting scientific research. *Research Policy* 35: 767–771.

Dosi, G. 1982. Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy* 11: 147–162.

Dosi, G. 1984. *Technical change and industrial transformation*. London: Macmillan.

Dosi, G. 1988. Sources, procedures and microeconomic effects of innovation. *Journal of Economic Literature* 26: 1120–1171.

Dosi, G., and R.R. Nelson. 2010. Technical change and industrial dynamics as evolutionary processes. In *Handbook of the economics of innovation*, vol. 1, ed. B.H. Hall and N. Rosenberg. Burlington: Academic.

Dosi, G., P. Llerena, and M. Sylos Labini. 2006a. Science-technology-industry links and the 'European Paradox': Some notes on the dynamics of scientific and technological research in Europe. *Research Policy* 35: 1450–1464.

Dosi, G., L. Marengo, and C. Pasquali. 2006b. How much should society fuel the greed of innovators? On the relations between appropriability, opportunities and rates of innovation. *Research Policy* 35: 1110–1121.

Dutton, J.M., and A. Thomas. 1984. Treating progress functions as a managerial opportunity. *Academy of Management Review* 9: 235–247.

Frenken, K., and L. Leydesdorff. 2000. Scaling trajectories in civil aircraft (1913–1997). *Research Policy* 29: 331–338.

Frenken, K., P.P. Saviotti, and M. Trommetter. 1999. Variety and niche creation in aircraft, helicopters, motorcycles and microcomputers. *Research Policy* 28: 469–488.

Giuri, P., C. Tomasi, and G. Dosi. 2007. *L'industria aerospaziale. Innovazione, tecnologia e strategia economica*. Milan: Il Sole 24 Ore e Fondazione Cotec.

Gordon, T.J., and T.R. Munson. 1981. *Research into technology output measures*. Glastonbury: The Future Group.

Granstrand, O. 1999. *The economics and management of intellectual property*. Cheltenham: Edward Elgar Publishing.

Gritsevskyi, A., and N. Nakicenovic. 2000. Modeling uncertainty of induced technological change. *Energy Policy* 28: 907–921.

Grupp, H. 1992. *Dynamics of science-based innovation*. Berlin: Springer.

Heller, M., and R. Eisenberg. 1998. Can patents deter innovation? The anti-commons in biomedical research. *Science* 280: 698–701.

Henderson, R.M., and K.B. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35: 9–30.

Jaffe, A.B.. 2000. The U.S. patent system in transition: Policy innovation and the innovation process. *Research Policy* 29: 531–577.

Jovanovic, B., and P.L. Rousseau. 2002. Moore's law and learning by doing. *Review of Economic Dynamics* 5: 346–375.

Klevorick, A.K., R.C. Levin, R.R. Nelson, and S.G. Winter. 1995. On the sources and significance of interindustry differences in technological opportunities. *Research Policy* 24: 185–205.

Kline, S.J., N. Rosenberg, S.J. Kline, and N. Rosenberg. 1986. An overview of innovation. In *The positive sum strategy: Harnessing technology for economic growth*. Washington, DC: National Academy Press.

Kuhn, T. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Levin, R.C., W.M. Cohen, and D.C. Mowery. 1985. R&D appropriability, opportunity and market structure: New evidence on some Schumpeterian hypotheses. *American Economic Review, Papers and Proceedings* 75: 20–24.

Lundberg, E. 1961. *Produktivitet och RûÊntabilitet*. Stockholm: Studieförbundet Näringsliv och samhälle.

MacDonald, A., and L. Schrattenholzer. 2001. Learning rates for energy technologies. *Energy Policy* 29: 255–261.

Malerba, F., and L. Orsenigo. 1996. The dynamics and evolution of industries. *Industrial and Corporate Change* 5: 51–87.

Mansfield, E., M. Schwartz, and S. Wagner. 1981. Imitation costs and patents: An empirical study. *The Economic Journal* 91: 907–918.

Marengo, L., C. Pasquali, M. Valente, and G. Dosi. 2009. *Appropriability, patents, and rates of innovation in complex products industries*, LEM working paper series, 2009/05. Pisa: Scuola Superiore Sant'Anna.

Mazzoleni, R., and R.R. Nelson. 1998. The benefits and costs of strong patent protection: A contribution to the current debate. *Research Policy* 27: 273–284.

Merges, R.P., and R.R. Nelson. 1994. On limiting or encouraging rivalry in technical progress: The effect of patent scope decisions. *Journal of Economic Behavior & Organization* 25: 1–24.

Mokyr, J. 2002. *The gifts of Athena: Historical origins of the knowledge economy*. Princeton: Princeton University Press.

Mokyr, J. 2010. The contribution of economic history to the study of innovation and technical change. In *Handbook of the economics of innovation*, vol. 1, ed. B.H. Hall and N. Rosenberg. Burlington: Academic.

Murmann, J.P., and K. Frenken. 2006. Toward a systematic framework for research on dominant designs, technological innovations, and industrial change. *Research Policy* 35: 925–952.

Murray, F.E., P. Aghion, M. Dewatripont, J. Kolev, and S. Stern. 2009. *Of mice and academics: Examining the effect of openness on innovation*. NBER Working Paper Series, 14819. Cambridge, MA.

Neij, L. 1997. Use of experience curves to analyse the prospects for diffusion and adoption of renewable energy technology. *Energy Policy* 25: 1099–1107.

Nelson, R.R. 1981. Research on productivity growth and productivity differences: Dead ends and new departures. *Journal of Economic Literature, American Economic Association* 19: 1029–1064.

Nelson, R.R. 2003. On the uneven evolution of human know-how. *Research Policy* 32: 909–922.

Nelson, R.R. 2004. The market economy, and the scientific commons. *Research Policy* 33: 455–471.

Nelson, R.R. 2006. Reflections on 'The Simple Economics of Basic Scientific Research': Looking back and looking forward. *Industrial and Corporate Change* 15: 145–149.

Nelson, R.R., and K. Nelson. 2002. On the nature and evolution of human know-how. *Research Policy* 31: 719–733.

Nelson, R.R., and S.G. Winter. 1977. In search of a useful theory of innovation. *Research Policy* 6: 36–76.

Nelson, R.R., and E.N. Wolff. 1997. Factors behind cross-industry differences in technical progress. *Structural Change and Economic Dynamics* 8: 205–220.

Nordhaus, W.D. 2007. Two centuries of productivity growth in computing. *Journal of Economic History* 67: 128–159.

Nuvolari, A. 2004. Collective invention during the British Industrial Revolution: The case of the Cornish pumping engine. *Cambridge Journal Economics* 28: 347–363.

Pavitt, K. 1999. *Technology, management and systems of innovation*. Cheltenham and Lyme: Edward Elgar.

Pavitt, K. 2001. Public policies to support basic research: What can the rest of the world learn from US theory and practice? (and what they should not learn). *Industrial and Corporate Change* 10: 761–779.

Rosenberg, N. 1976. *Perspectives on technology*. Cambridge: Cambridge University Press.

Rosenberg, N. 1982. *Inside the black box: Technology and economics*. Cambridge: Cambridge University Press.

Rosenberg, N. 1994. *Exploring the black box: Technology, economics, and history*. Cambridge: Cambridge University Press.

Rosenberg, N. 2009. Some critical episodes in the progress of medical innovation: An Anglo-American perspective. *Research Policy* 3: 234–242.

Rosenbloom, R.S., and M.A. Cusumano. 1987. Technological pioneering and competitive advantage: The birth of the VCR industry. *California Management Review* 29: 51–76.

Sahal, D. 1981. *Patterns of technological innovation*. New York: Addison-Wesley.

Sahal, D. 1985. Technological guideposts and innovation avenues. *Research Policy* 14: 61–82.

Saviotti, P.P. 1996. *Technological evolution, variety and the economy*. Cheltenham: Edward Elgar Publishing.

Saviotti, P.P., and A. Trickett. 1992. The evolution of helicopter technology, 1940–1986. *Economics of Innovation and New Technologies* 2: 111–130.

Suarez, F.F., and J.M. Utterback. 1995. Dominant designs and the survival of firms. *Strategic Management Journal* 16: 415–430.

Teece, D.J. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* 15: 285–305.

Teece, D.J., R. Rumelt, G. Dosi, and S.G. Winter. 1994. Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization* 23: 1–30.

Thompson, P. 2010. Learning by doing. In *Handbook of the economics of innovation*, vol. 1, ed. B.H. Hall and N. Rosenberg. Burlington: Academic.

Tushman, M.L., and P. Anderson. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* 31: 439–465.

Utterback, J.M., and F.F. Suarez. 1993. Innovation, competition, and industry structure. *Research Policy* 22: 1–21.

Von Hippel, E. 2005. Democratizing innovation: The evolving phenomenon of user innovation. *Journal für Betriebswirtschaft* 55: 63–78.

Winter, S.G. 1993. Patents and welfare in an evolutionary model. *Industrial and Corporate Change* 2: 211–231.

Wright, T.P. 1936. Factors affecting the costs of airplanes. *Journal of Aeronautical Sciences* 10: 302–328.

Yelle, L.E. 1979. The learning curve: Historical review and comprehensive survey. *Decision Sciences* 10: 302–308.

Further Reading

Beardsley, G., and E. Mansfield. 1978. A note on the accuracy of industrial forecasts of the profitability of new products and processes. *Journal of Business* 51: 127–135.

Bresnahan, T.F., and M. Trajtenberg. 1995. General purpose technologies: Engines of growth? *Journal of Econometrics* 65: 83–108.

Bush, V. 1945. *Science: The endless frontier*. Washington, DC: GPO.

Castaldi, C., R. Fontana, and A. Nuvolari. 2009. 'Chariots of Fire': The evolution of tank technology, 1915–1945. *Journal of Evolutionary Economies* 19: 545–566.

Chataway, J., J. Tait, and D. Wield. 2004. Understanding company R&D strategies in agro-biotechnology: Trajectories and blind spots. *Research Policy* 33: 1041–1057.

Cohen, W., and R. Levin. 1989. Empirical studies of innovation and market structure. In *Handbook of industrial organization*, vol. 2, ed. R. Schmalensee and R. Willig. Amsterdam: Elsevier.

Consoli, D. 2005. The dynamics of technological change in UK retail banking services: An evolutionary perspective. *Research Policy* 34: 461–480.

Dasgupta, P., and P.A. David. 1994. Towards a new economics of science. *Research Policy* 23: 487–521.

Dawid, H. 2006. Agent-based models of innovation and technological change. In *Handbook of computational economics, Vol. 2: Agent-based computational economics*, ed. L. Tesfatsion and K.L. Judd. Amsterdam: Edward Elgar.

Dew, N. 2006. Incommensurate technological paradigms? Quarreling in the RFID industry. *Industrial and Corporate Change* 15: 785–810.

Dosi, G., and M. Egidi. 1991. Substantive and procedural uncertainty: An exploration of economic behaviours in changing environments. *Journal of Evolutionary Economics* 1: 145–168.

Dosi, G., K. Pavitt, and L. Soete. 1990. *The economics of technical change and international trade*. Brighton/Wheatsheaf/New York: New York University Press.

Freeman, C., and C. Perez. 1988. Structural crises of adjustment: Business cycles and investment behavior. In *Technical change and economic theory*, ed. C. Freeman, R.R. Nelson, G. Silverberg, and L. Soete. London: Pinter Publishers.

Freeman, C., and L. Soete. 1997. *The economics of industrial innovation*, 3rd ed. London/Washington, DC: Pinter.

Gary, M.S., G. Dosi, and D. Lovallo. 2008. Boom and bust behavior: On the persistence of strategic decision bias. In *The Oxford handbook of organizational decision making*, ed. G.P. Hodgkinson and W.H. Starbuck. Oxford/New York: Oxford University Press.

Geuna, A., A. Salter, and W.E. Steinmuller (eds.). 2003. *Science and innovation: Rethinking the rationale for funding and governance*. Cheltenham: Edward Elgar.

Hall, B.H., and N. Rosenberg (eds.). 2010. *Handbook of the economics of innovation*. Burlington: Academic.

Kerker, M. 1961. Science and the steam engine. *Technology and Culture* 2: 381–390.

Merton, R.K. 1973. *The sociology of science: Theoretical and empirical investigations*. Chicago: University of Chicago Press.
Mina, A., R. Ramlogan, G. Tampubolon, and J.S. Metcalfe. 2007. Mapping evolutionary trajectories: Applications to the growth and transformation of medical knowledge. *Research Policy* 36: 789–806.
Needham, J. 1962–3. The pre-natal history of the steam engine. *Transactions of the Newcomen Society* 35: 3–58.
Nelson, R.R., and S.G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.
Nightingale, P. 1998. A cognitive model of innovation. *Research Policy* 27: 689–709.
Perez, C. 1985. Microelectronics, long waves and world structural change: New perspectives for developing countries. *World Development* 13: 441–463.
Perez, C. 2010. Technological revolutions and techno-economic paradigms. *Cambridge Journal of Economics* 34: 185–202.
Polanyi, M. 1962. *Personal Knowledge: Towards a Post-Critical Philosophy.* Chicago: University of Chicago Press.
Possas, M.L., S. Salles-Filho, and J.M. Silveira. 1996. An evolutionary approach to technological innovation in agriculture: Some preliminary remarks. *Research Policy* 25: 933–945.
Soete, L. 1979. Firm size and inventive activity: The evidence reconsidered. *European Economic Review* 12: 319–340.
Starbuck, W., and J.M. Mezias. 1996. 'Opening Pandora's Box': Studying the accuracy of managers' perceptions. *Journal of Organizational Behavior* 17: 99–117.
Stoneman, P. 1995. *Handbook on the economics of innovation and technical change*. Oxford: Blackwell.

Technology Adoption

Chris Forman[1], Avi Goldfarb[2] and Shane Greenstein[3]
[1]Georgia Institute of Technology, Atlanta, GA, USA
[2]University of Toronto, Rotman School of Management, Toronto, OH, Canada
[3]Northwestern University, Evanston, IL, USA

Abstract

Using examples from information technology adoption, we emphasize the role of costs, benefits, communications channels and dynamic considerations in the decision to adopt new technology. We discuss differences between adoption by consumers and adoption by firms. We emphasize the adoption of business process innovations, which alter organizational practices and often involve the post-adoption invention of complementary business processes and adaptations. Within the context of business adoption, we discuss the inherent challenges in identifying the decision maker and the role of competition in influencing the benefits to adoption.

Definition Technology adoption occurs when an individual, firm or other agent first makes use of a new technology. In this setting, technology can refer to a new product, service or management innovation.

There is a vast economics literature studying the use of newly available technologies. When an individual, firm or other agent (e.g., government or non-profit) first makes use of the new service or product, this act is often labelled 'adoption'. An important facet of this literature analyses the economic determinants of heterogeneity in adoption behaviour, principally the timing of adoption and the willingness to pay more for new products and services. In this entry we discuss several key issues for understanding the economics of the adoption of new technologies, using information technology as an example.

A General Discussion of Costs and Benefits of Adopting a New Technology

In weighing the decision to adopt a new technology, the basic economic model compares the costs of adoption with the benefits. The costs include the physical set-up costs, the costs of learning to use the new technology and the costs of purchasing any services that are complementary to the technology being adopted. The benefits include the initial increase in utility or productivity through the use of the new technology as well as the longer-term benefits.

The most common method for comparing costs and benefits is the 'Probit model of adoption', developed by David (1969). In this model, agents adopt the new technology when the benefits exceed the costs. Empirical studies of adoption regress the adoption decision on factors that might affect costs and benefits. For example, Forman et al. (2005) examine internet adoption by businesses. They use a comprehensive cross-section of US business establishments and regress the establishments' adoption decision on features of the establishments and features of their locations.

The probit model implicitly measures the adoption decision as the reduced form of a dynamic process in which agents weigh the present value of all future costs and benefits of adopting a new technology. Agents might wait to adopt in order to take advantage of lower future prices, better future complementary technologies or better future quality; or they might adopt early in order to benefit from the technology over a longer period. This suggests that past investments in technology can reduce the marginal benefit of adopting the next generation (Forman 2005).

The probit model of adoption is missing one important aspect of the economics of technology diffusion: that diffusion happens slowly as agents learn about the existence of, and benefits of, the new technology. Thus, learning and communication channels matter to the diffusion of new technology. Goolsbee and Klenow (2002) emphasize the importance of learning by seeing what other people are doing as a key aspect of the diffusion of personal computers. Their work highlights the econometric challenges of identifying the effects of learning and communication from other factors that might drive correlation in user behaviour, such as network externalities or other types of user spillovers.

Below, we provide examples of the costs and benefits of adoption, the dynamic considerations and the communication channels through which people learn about new technology. Because the technology adoption process differs depending on whether the adopters are households or firms, we discuss households and firms separately.

Technology Adoption by Households

Prince's (2008) study of personal computer (PC) ownership illustrates the tenor of many studies of household adoption of new technology. Three main factors help explain both the timing of adoption and the willingness to pay: (1) differences across consumers in their valuation of the product (benefits), (2) the barriers to the initial purchase (costs) and (3) dynamic aspects of the decision. His results indicate that the marginal utility of PC quality is strongly increasing in income and education, and strongly decreasing in age. Further, as prices fall and quality rises over time, the decision about whether to buy a new PC is complicated by the dynamic decision of when to buy a new PC. Furthermore, the households that become first-time purchasers are more price sensitive than repeat purchasers, and perceive large costs in the initial adoption.

Goldfarb and Prince (2008) examine a probit model of adoption and emphasize the current period trade-off between costs and benefits. While the relative costs of adoption vary according to consumer budget constraints, the benefits depend in large part on the opportunity cost of time. Specifically, they find that people with high levels of income and education are more likely to adopt, but spend less time online when they do adopt. Their results suggest that this is driven by the large array of other leisure activities available to people with high levels of income and education.

Household adoption decisions can often be shaped by the behaviour of other users or through institutions. For example, Goolsbee and Klenow (2002) add spillovers to the standard probit, studying how users' adoption decisions are interrelated either through learning effects or through network externalities. Goldfarb (2006) documents the role of universities as a communications channel that aided the diffusion of internet technology. Internet technology diffused from university researchers to undergraduate students in the mid-1990s. Those students then brought the technology home with them and it diffused to other household members.

Overall, technology adoption by households involves a (potentially dynamic) trade-off between costs and benefits, and a communication process through which individuals learn about the existence of, and benefits to, the new technology.

Technology Adoption by Firms

The study of adoption in business tends to raise additional questions beyond those faced by households. Unlike household adoption studies, in a business setting it can be difficult to identify the decision maker or ‘adopter’. This confusion arises for many reasons, but principally because there may be sharing of non-capital investments across a wide array of processes. Though the unit costs of sharing are lower for large organizations, the sharing usually does not occur instantaneously or without high coordination costs across many parts of an organization (Astebro 2002, 2004). Many decision makers and considerations can shape the coordination of adoptions decisions. As a result, organizations may adopt new technologies that take time to diffuse across users.

Competitive pressure is another key factor in technology adoption by businesses. That is, there first may be a minimal level of investment necessary just to be in business. Second, adoption of some technologies may confer competitive advantage vis-à-vis rivals. As an illustration, computing frequently enables the invention of entirely new services and products that may or may not provide permanent or temporary competitive advantage. When new services are reasonably permanent, a private firm may see returns to the investment in the form of increases in final revenue or other strategic advantages. If a new product or service is quickly imitated by all firms, it rapidly becomes a standard feature of doing business in a downstream market. The benefits from the new technology are quickly passed on to consumers in the form of lower prices and better products. In this case, the benefits to a firm do not appear as an increase in revenues but they exist, nonetheless, in the form of losses avoided by the businesses in question.

McElheran (2012) uses such reasoning to investigate the order in which existing manufacturing firms adopted electronic commerce during its first wave of diffusion in the late 1990s. She shows that many of the most productive firms adopted electronic commerce in their procurement processes – to secure inputs – which led to a reinforcement of cost leadership. On the other hand, many of the most productive firms were followers in adopting customer-facing electronic commerce, which had much riskier payoffs and did not integrate as well with existing business processes. That led many leading firms to be more cautious, while a second tier of firms took the lead in experimenting.

One particularly important type of technology adoption by businesses is the adoption of *business process innovations*. Such innovations alter organizational practices, generally with the intent of improving services, reducing operational costs and taking advantage of new opportunities to match new services to new operational practices. Typically, this type of innovation involves changes in the discretion given to employees, changes to the knowledge and information that employees are expected to retain and employ, and changes to the patterns of communications between employees and administrators within an organization. Such innovations involve the retraining of employees and the redesign of organizational architecture, such as its hierarchy, lines of control, compensation patterns and oversight norms.

Prior studies stress the importance of *co-invention*, the post-adoption invention of complementary business processes and adaptations aimed at making adoption useful (Bresnahan and Greenstein 1996). For example, an initial investment in information technology (IT) is not sufficient for ensuring that productivity gains arise. Those gains depend on whether the employees of the adopting organization find new uses or services to take advantage of the new capabilities, and/or invent new processes for unanticipated problems.

The adoption costs of business process innovations may depend on the availability of third-party services, such as third-party consulting, which may not be present at early moments in a new technology's diffusion (Bresnahan et al. 2002). In the absence of market solutions, adopters may have to divert internal resources to solve idiosyncratic issues. The incentives around utilization and investment can also change considerably over time, owing to changes in the restructuring of the organization's hierarchy and operational practices (Bloom et al. 2011).

Co-invention costs are an important factor in explaining business adoption of the Internet. By the late 1990s, implementation of first-generation internet applications such as email was straightforward. It involved a PC, a modem, a contract with an ISP and some appropriate software. In contrast, investment in the use of the Internet for an application module in a suite of Enterprise Resource Planning software was anything but routine. Such an implementation included technical challenges beyond the Internet's core technologies, such as security, privacy and dynamic communication between browsers and servers. Usually organizational procedures also changed. In particular, Forman et al. (2003a, b, 2005, 2008, 2012) examine the causes and consequences of business adoption of internet technology, and emphasize the importance of coinvention as a driver of technology adoption by firms. For example, Forman et al. (2008) showed that firms with easy access to skilled IT workers (whether locally or within the firm) were much more likely to adopt advanced internet applications.

Directly connecting adoption of business process innovation with performance is challenging. Building on their work on internet adoption, Forman et al. (2012) show that the aggregate benefits (in terms of wages) of such adoption are much higher in locations with a ready supply of IT expertise. Hubbard (2000, 2003) and Baker and Hubbard (2003, 2004) examine the productivity benefits of a business process innovation at a more micro level. In particular, they examine the use of computing technologies to monitor the performance of trucks. They document the fact that such technologies improve the ability of trucking firms and private fleets to coordinate assets and better match trucks to tasks, and monitoring of trucker actions. Both lead to improved output.

Overall, the study of business adoption raises additional issues owing to the inherent challenges in identifying the decision maker, the role of competition and, especially, the role of co-invention in the adoption of business process innovations.

Summary

Using examples from information technology adoption, we have emphasized the role of costs, benefits, communications channels and dynamic considerations in the decision to adopt new technology. We have also discussed differences between adoption by consumers and adoption by firms. Of course, such a short article cannot comprehensively cover all issues related to technology adoption and diffusion. Literature reviews, with various perspectives, include Rogers' (1995) review of the communications literature, Stoneman's (2002) review of the economics literature, and Forman and Goldfarb's (2006) review of the drivers of information technology adoption by businesses.

See Also

- ▶ Business Process Re-engineering
- ▶ Complementarities
- ▶ General-Purpose Technology
- ▶ Geography of Innovation
- ▶ Information Technology and Strategy
- ▶ Innovation Diffusion
- ▶ Learning and Adaptation
- ▶ Network Effects

References

Astebro, T. 2002. Noncapital investment costs and the adoption of CAD and CNC in U.S. metalworking industries. *RAND Journal of Economics* 33: 672–688.

Astebro, T. 2004. Sunk costs and the depth and probability of technology adoption. *Journal of Industrial Economics* 52: 381–399.

Baker, G., and T.N. Hubbard. 2003. Make versus buy in trucking: Asset ownership, job design, and information. *American Economic Review* 93: 551–572.

Baker, G., and T.N. Hubbard. 2004. Contractibility and asset ownership: On-board computers and governance in US trucking. *Quarterly Journal of Economics* 119: 1443–1480.
Bloom N, Garicano L, Sadun R, Van Reenen J. 2011. The distinct effects of information technology and communication technology on firm organization. Working paper, Stanford University.
Bresnahan, T. and Greenstein, S. 1996. Technical progress and co-invention in computing and in the use of computers. *Brookings Papers on Economic Activity: Microeconomics* 1–78.
Bresnahan, T.F., E. Brynjolfsson, and L.M. Hitt. 2002. Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. *Quarterly Journal of Economics* 117: 339–376.
David, P. A. 1969. A contribution to the theory of diffusion. Memorandum No. 71, Stanford Center for Research in Economic Growth, Stanford University.
Forman, C. 2005. The corporate digital divide: Determinants of internet adoption. *Management Science* 51: 641–654.
Forman, C., and A. Goldfarb. 2006. Diffusion of information and communications technology to business. In *Handbooks in information systems*, Economics and Information Systems, vol. 1, ed. T. Hendershott. Amsterdam: Elsevier.
Forman, C., A. Goldfarb, and S. Greenstein. 2003a. Which industries use the internet? In *Organizing the new industrial economy*, ed. M. Baye. Amsterdam: Elsevier.
Forman, C., A. Goldfarb, and S. Greenstein. 2003b. The geographic dispersion of commercial internet use. In *Rethinking rights and regulations: Institutional responses to new communications technologies*, ed. L. Cranor and S. Wildman. Cambridge, MA: The MIT Press.
Forman, C., A. Goldfarb, and S. Greenstein. 2005. How did location affect adoption of the internet by commercial establishments? Urban density versus global village. *Journal of Urban Economics* 58: 389–420.
Forman, C., A. Goldfarb, and S. Greenstein. 2008. Understanding inputs into innovation: Do cities substitute for internal firm resources? *Journal of Economics and Management Strategy* 17: 295–316.
Forman, C., A. Goldfarb, and S. Greenstein. 2012. The internet and local wages: A puzzle. *American Economic Review* 102: 556–575.
Goldfarb, A. 2006. The (teaching) role of universities in the diffusion of the internet. *International Journal of Industrial Organization* 24: 203–225.
Goldfarb, A., and J. Prince. 2008. Internet adoption and usage patterns are different: Implications for the digital divide. *Information Economics and Policy* 20: 2–15.
Goolsbee, A., and P. Klenow. 2002. Evidence on learning and network externalities in the diffusion of home computers. *Journal of Law and Economics* 45: 317–344.
Hubbard, T. 2000. The demand for monitoring technologies: The case of trucking. *Quarterly Journal of Economics* 115: 533–560.
Hubbard, T. 2003. Information, decisions, and productivity: On-board computers and capacity utilization in trucking. *American Economic Review* 93: 1328–1353.
McElheran, K. 2012. Do market leaders lead in business process adoption? Customers and co-invention in e-business adoption. Working paper, Harvard Business School.
Prince, J. 2008. Repeat purchase amid rapid quality improvement: Structural estimation of demand for personal computers. *Journal of Economics and Management Strategy* 17: 1–33.
Rogers, E. 1995. *The diffusion of innovations*, 4th ed. New York: Free Press.
Stoneman, P. 2002. *The economics of technological diffusion*. Oxford: Blackwell.

Technology Cycles

James M. Utterback
Massachusetts Institute of Technology,
Cambridge, MA, USA

Abstract

New technologies can create whole new industries. New technologies can and often do successfully disrupt and eventually overwhelm prominent firms, which have built their positions based on prior product concepts or process techniques. Technology cycles are often described as following a pre-determined or predictable trajectory. Progress is seen to involve a succession of cycles, each ending in a discontinuity as a new trajectory is established and a new cycle begins (Sahal, *Patterns of technological innovation*. Reading: Addison Wesley; Dosi. 1982. *Research Policy* 11: 147–162, 1981). Understanding technology cycles may lead to better defining opportunities, threats and potential competitive outcomes of a firm's strategic choices (Gavetti and Levinthal, *Management Science* 50: 1309–1318, 2004). The most popular ideas about mapping and predicting technology

T

cycles and performance though are proving to be seriously oversimplified and misleading when subjected to searching examination. It is more vital than ever for strategists to understand the changing texture of technology, but in a richer and more nuanced way.

Definition Technological change is seen to be a force of increasing salience for value creation and competitive success of firms. Cycles of emergence of new technology, growth in performance and effectiveness and eventual stagnation or stasis are thought to be recurring phenomena in industries past and present.

Performance as a Predictable Trajectory

Technological performance has been thought of as a monotonically increasing function of effort expended on development and is assumed to follow a rising double exponential or 'S'-shaped curve. The assumption of slow, rapid and then again slow improvements forms the conceptual cornerstone of many works on ▶ technology strategy and ▶ innovation.

The shape of performance curves is usually explained by suggesting that in the early days of a technology there is a plethora of possible approaches and directions for development and great uncertainty about which to pursue. Progress will be slow at first due to many failed experiments and dead ends. As early failures and successes build expertise, uncertainties will be resolved and progress accelerated along a more and more clearly defined trajectory. Soon, however, many of the most promising avenues will have been exploited, and physical limits and binding constraints may appear. These may attenuate or even tightly limit further improvements in performance. Much early work attempted to predict technological progress using various techniques for trend extrapolation and for defining limits (Ayres 1969: ch. 6). The problem of multiple performance measures, especially for sophisticated products and systems, was addressed by the creation of various indices and weighting schemes (for example, Saviotti et al. 1982). The reality that performance is rarely depicted by a clear set of points, but rather by a broad smear of data including leading and lagging examples was noted. In an extreme case the differences between leading performance and technology in general use spanned four orders of magnitude (Hilbrink 1989)!

Hirooka (2006: 336–340) usefully relates research trajectories to development trajectories and diffusion trajectories for a number of technologies including solar cells, fuel cells and superconductors. Porter and colleagues (1991: 170–172) summarize the popular idea that an envelope of performance over time can be constructed by stacking S-curves of successive individual technologies, such as the vacuum tube, transistor and integrated circuit, and suggests the resulting envelope will also have a double exponential shape. An early test of this hypothesis led to disturbing results. Far from being predictable the envelope for generations of typesetting technologies was relatively stagnant from 1500 to 1960, but progress in typesetting speeds has been nearly exponential since then with no sign of diminishing (Mohn 1972). Moreover, Mohn (1972: 227) provides data for hot metal linotype setting speeds that negates all of the ideas just summarized. From 1886 speeds increase rapidly until they reach a long plateau, continuing until about 1960, of roughly 10,000 characters per hour. Then, another surge of improvement is observed, just the opposite of expectations regarding limits. This is but the earliest of many contradictions of popular opinion to be covered in more detail below.

As mentioned above, technological performance has been thought of as a monotonically increasing function of effort expended on development. Cumulative effort is difficult to define and to measure, especially in sectors with numerous and often changing competitors. As a consequence, much of the research on technology cycles resorts to plotting performance as a function of cumulative production volume or even simply over time assuming that production quantities and development efforts increase in constant proportions with time (Fusfeld 1970: 308; Sahal 1981: 186–187).

The Importance of Technology Cycles for Strategy

The critical importance for strategy of understanding technology cycles is that most firms invest far too heavily in development efforts long past the time that rewarding improvements might be expected. Further, once threatened, firms typically redouble their investment in familiar concepts and fail to make a transition to an emerging concept. Most research on technology cycles presumes that a technology with superior performance will be preferred by the market. But the potency of a newly introduced technology may be that it offers different performance, enabling a great expansion of the market and openings for new rivals (Cooper and Schendel 1976; Foster 1986).

In his influential book *Innovation*: *the Attacker's Advantage*, Richard Foster gives two examples, both homogeneous products from the chemical and synthetic fibre industries, of performance related to R&D effort. Foster (1986: 157) observes that firms almost inevitably wait too long to attempt a transition from a dominant established technology to an emerging challenger, and that they overestimate their ability to anticipate discontinuities, to identify potential new competitors and to time the onset of a new technology cycle. Building on the widely accepted ideas about the dynamics of performance improvement summarized above he argues firms should switch their investments to new concepts much earlier than they usually do. In essence he believes that the ► myopia of long-established firms gives attackers a significant advantage during transitions between generations of technologies. Foster cites a long list of products in the container and packaging business whose market positions have been overturned by innovative competitors: glass bottles by steel cans; steel cans by aluminum cans; glass bottles by plastic bottles; plastic-coated milk cartons by plastic jugs; and so forth. In each case, he notes that market leadership passed from one set of firms to another. Today's leaders, in these cases, were never leaders in the next product generation. 'I don't know of any comprehensive statistics that would stand up to academic scrutiny, but my feeling is that leadership changes hands in about seven out of ten cases when discontinuities strike. A change in technology may not be the number-one corporate killer, but it certainly is among the leading causes of corporate ill health' (Foster 1986: 116). He argues that an optimal pattern of development spending would tend to follow a normal distribution, with the majority of spending concentrated in the centre of the cycle, rather than constantly rising and being concentrated toward the end of product life.

Cooper and Schendel (1976), examining 22 cases of technological discontinuities, observe that threatened firms typically redouble their investment in familiar concepts and fail to make a transition. Often they ride out the failure of the old while making only defensive investments in the new. Cooper and Schendel conclude that such a dual strategy is simply not viable. Their explanation for the observed syndrome is that, 'decisions about allocating resources to old and new technologies within the organization are loaded with implications for the decision makers; not only are old product lines threatened, but also old skills and positions of influence' (Cooper and Schendel 1976: 68–69). It is not uncommon for a threatened firm to develop a hybrid or intermediate form of product combining aspects of both new and old. An example was Lockheed's Electra aircraft, a hybrid of jet engine and propeller. The Electra was not a successful competitor to rapidly evolving turbojets (Girifalco 1991: 112–113). Hybrid strategies seem to have a general history of failure. Examples of failed hybrids are legion from the steaming sailing ship to Thomas Edison's GEM lamp and to Good-year's bias-ply belted tire (Utterback 1994).

Irvine and Martin (1984) extend expectations regarding predictable trajectories to the arena of public policy, developing the concept of macro-strategic research. Their work has led to the further development of research milestones or strategic technology road mapping at a national level. If ideas about the predictability of technology cycles and trajectories are incorrect, then such rigorous attempts at planning will lead to errors and misallocations of resources both for corporations, and more importantly for industries and whole economies. The weight of research reviewed below suggests that the most popular ideas about mapping and predicting technology cycles and performance are seriously over simplified and misleading. It is more vital than ever for strategists to understand the changing texture of technology, but in a richer and more nuanced way.

A Critique of Widely Accepted Theory and Applications

Hints of trouble ahead are evident in many of the sources reviewed. These include technologies with linearly increasing performance as well as ones with continuing exponential improvement despite repeated predictions of limits, Moore's Law being a famous example (Mollick 2006). Moreover, many technologies, which seemingly have reached limits, are dramatically reborn when faced with competition from something new. The gas lighting industry when faced with Edison's carbon filament incandescent lamp responded by developing an incandescent 'mantle' of ceramic filaments improving the efficiency of gas lamps by nearly threefold. 'For a number of years the potential superiority of the incandescent lamp remained in doubt, and even its survival was sometimes questioned' (Bright 1949: 127). Edison himself faced with competition from European innovations in metal filaments created a new form of carbon, the 'GEM' lamp, which performed nearly as well. Bright (1949: 181) reports that, 'the filament which resulted in 1904 from Whitney's work was the greatest improvement made in the carbon lamp since 1884'. These and myriad other examples (in Utterback and Kim 1986; Utterback 1994) are stark evidence not only of firms' tendency toward retrenchment and tradition when threatened, but also of complacent incremental improvement at other times.

In a searching review of the technology cycle literature including a further detailed study of 14 technologies, Sood and Tellis (2005: 152) conclude that, 'the results contradict the prediction of a single S-curve. Instead, technological evolution seems to follow a step function, with sharp improvements in performance following long periods of no improvement. Moreover, paths of rival technologies may cross more than once or not at all.'

Koh and Magee (2006) take a more broadly functional approach to technology cycles by focusing on three broad categories; matter, energy and information, and then classify technologies according to the manner in which they operate in different domains such as: transform, transport, store, exchange and control. Their approach provides a way to deal with the varied multidimensionality of individual technologies. In applying their technique to information storage, calculation and communication using a 100-year data series Koh and Magee find generally continuous progress for each functional category independent of the specific underlying technologies dominating at different times. In essence they suggest that the envelope of performance for a succession of technologies in each category is more stable and continuous than is that for any individual cycle. In a later study (2008) of storage, transportation and transformation of energy, Koh and Magee report substantial variability of progress rates is found within given functional categories for energy compared to relatively small variation within any one category for information technology.

In a similar general argument, McNerney and colleagues (2012) make the prediction that the rate of improvement of a technology depends on its design complexity, that is the number of components incorporated and the number of connections among them. The possibility that design complexity may be reduced through time will, they suggest, be correlated with rates of performance improvement. Thus, products having many interconnected parts may advance more rapidly than simpler, more homogeneous, and more integrated examples, and as designs become simpler over time the rate of advance would be expected to slow. Their model and hypotheses are resonant with histories presented in Utterback (1994) in which patterns of innovation for complex products such as electric lighting, typewriters, calculators and computers are contrasted with homogeneous products such as sheet glass and rayon, and products and processes do indeed become simpler and more integrated over time.

Transitions Between Cycles

Much research on technology cycles implies that growth of a new technology comes at the expense of displacing prior art. A key point for this review though is that an innovation often both enables growth into a broader market and prospers through it, rather than displacing prior art in an existing market. As we have seen, a presumption is that technology with superior performance on a traditional figure of merit will be preferred by the market. But the salient feature of a newly introduced technology may be that it offers dimensions of performance not possessed by the traditional offering, allowing the market to be re-framed (Christensen 1992a, b; Levinthal 1998; Kaplan and Tripsas 2008). Thus, radio was first defined as ‘wireless telegraphy’ and for years occupied just a niche market, communication over water, where wired connections were absent. Only when its ability to reach a mass audience with news and music was recognized did its transforming potential become realized. There were no extant competitors with broad reach, and the growth of ‘radio broadcasting’ was explosive.

A widely held belief is that at the time an invading technology first appears an established technology generally offers better performance or cost than does the challenger. The new technology may be viewed objectively as crude, leading to the belief that is will find only limited application. The performance superiority of an established technology may prevail for quite some time, but if the new has real merit, it typically enters a period of rapid improvement – just as the established technology enters a stage of slow improvement. Eventually, the newcomer improves its performance characteristics to the point where they match those of the established technology and rockets past it, still in the midst of rapid improvement.

Utterback (1994: ch. 9) in a meta-analysis of 46 transitions in technology cycles finds that a discontinuous change may drastically increase the aggregate demand for the products of an industry. The replacement of the vacuum tube by the transistor, and later the integrated circuit, has increased the sales of the electronics industry from several billions of dollars to hundreds of billions. The replacement of piston aircraft engines by turbojets has correspondingly dramatically reduced the costs and increased the seat miles flown by commercial aviation. Innovations that broaden the market may create room for new firms to start. Innovation-inspired substitutions conversely may cause established firms to hang on all the more tenaciously, making it extremely difficult for an outsiders to gain a foothold and the cash flow needed to expand and become a player in the industry. Innovations that substitute for established products and processes thus may arise more often inside an industry. Some innovations create a wholly new market niche, encouraging the entry of many new entrants. Here, established firms are unlikely to enter successfully and new firms have greater survival odds.

Christensen (1997), in a comprehensive study of the Winchester disc drive industry, finds that each transition in the leadership of the industry was led by new firms addressing an un-served market segment with simpler and less expensive architectural innovations, rather than with wholly new technology. He observes that performance curves of product variants need not intersect for change in leadership to follow. Firms having different products address a spread of performance levels and demands. While older firms focus on demanding major customers, demand is more elastic and rapidly growing in the un-served market for simpler variants. Demand for performance is better described by a broad rising band rather than a line, just as is performance itself.

In a study of the evolution of laser printers de Figueiredo and Kyle (2006: 242) observe two innovation frontiers; a top frontier that is the traditional ‘make it better, faster’ one; and a bottom frontier of ‘make it cheaper and accessible’. Advances evidently occur at both boundaries, allowing penetration of a broad range of new market segments. They suggest this to be true of many other technologies including integrated circuits, microprocessors and digital cameras.

Technology Cycles as a Function of Experiment and Synthesis of Diverse Inputs

Abernathy and colleagues (1982) suggest that performance, rather than being a monotonically increasing function of effort expended on development, is a function of frequency and diversity of experimentation in the market, and of conditions that encourage experimentation. Within the rich mixture of experimentation and competition at the start of a cycle some centre of gravity eventually forms, usually in the shape of product standards or production practices, that is, a 'dominant design'. The bases of competition change radically, and firms are put to tests that very few pass. Before too long, the ecology of competing firms changes from many competitors to few. One might consider each firm's investments and product introductions as experiments, which provide corrective and stimulating feedback to that firm and to the industry about product and market requirements. Thus, the earliest period in the development of a product line or industry, in which few firms participate, would necessarily be a period of relatively slow technical progress and productivity advance. As larger numbers of firms enter the arena, thus broadening the range of experimentation and the definition of the product technology, greater innovation with correspondingly greater technology progress and productivity advance should be expected. Finally, as a few firms come to dominate the industry with superior product technology and productivity, both experimentation and progress would be expected to slow (Utterback 1994).

This is not to say that the narrowing of search never reflects a genuine lack of technological opportunities. But it is to say that, more often than not, firms become so structured that they only search very narrowly. Indeed, it is this structuring that critically limits progress: the more specialized firms are, the more narrowly they tend to search for new opportunities (Abernathy et al. 1982: 7).

Radical innovation, and the growth of new industries, is probably more likely when a firm occupies the confluence or convergence of distinct streams of emerging technology. Research progress at the intersection of fields is more likely to occur when cross-disciplinary new product development teams are designed by the organization and when routines and processes are designed to support cross-disciplinary learning. A confluence of technologies is characterized both by the bringing together of formerly disparate fields of knowledge and by the creation of new product markets. When a confluence of technology streams occurs, rich opportunities for experiment and progress may result (Utterback 1994). These may ultimately lead to an emerging industry and newly dominant firms (Maine et al. 2012).

In summary, amid rapidly changing markets, extended supply chains, and widening sources of competition, dislocations and opportunities frequently arise from surprising sources. We might more creatively think of innovation and firm formation as a process of experimentation in the market. Rather than seeking to reduce uncertainty and to optimize, perhaps we should seek to increase possibilities for experimentation and for broader search and synthesis. Technology cycles then are anything but consistent and predicable trajectories.

See Also

▶ Innovation
▶ Myopia
▶ Technological Change
▶ Technology Strategy

References

Abernathy, W.J., B.H. Klein, J. Dopico, and J.M. Utterback. 1982. *A proposal for policy oriented research on the automobile industry*, Working paper no. 81-14. Cambridge, MA: MIT Center for Policy Alternatives.

Ayres, R.U. 1969. *Technological forecasting and long-range planning*. New York: McGraw-Hill.

Bright Jr., A.A. 1949. *The electric lamp industry: Technological change and economic development from 1800 to 1947*. New York: Macmillan.

Christensen, C.M. 1992a. Exploring the limits of the technology S-curve part I: Component technologies. *Production and Operations Management* 1: 334–357.

Christensen, C.M. 1992b. Exploring the limits of the technology S-curve part II: Architectural technologies. *Production and Operations Management* 1: 358–366.
Christensen, C.M. 1997. *The innovator's dilemma: When new technologies cause great firms to fail*. Boston: Harvard Business School Press.
Cooper, A., and D. Schendel. 1976. Strategic responses to technological threats. *Business Horizons* 19: 61–69.
De Figueiredo, J., and M. Kyle. 2006. Surviving the gales of creative destruction: The determinants of product turnover. *Strategic Management Journal* 27: 241–264.
Dosi, G. 1982. Technological paradigms and technological trajectories. *Research Policy* 11: 147–162.
Foster, R. 1986. *Innovation: The attacker's advantage*. New York: Summit Books.
Fusfeld, A.R. 1970. The technological progress function: A new technique for forecasting. *Technological Forecasting* 1: 301–312.
Gavetti, G., and D. Levinthal. 2004. The strategy field from the perspective of *Management Science*: Divergent strands and possible integration. *Management Science* 50: 1309–1318.
Girifalco, L.A. 1991. *Dynamics of technological change*. New York: Van Nostrand Reinhold.
Hilbrink, J.O. 1989. Economic impact and technological change. *IEEE Transactions on Engineering Management* 36: 37–46.
Hirooka, M. 2006. *Innovation dynamism and economic growth: A nonlinear perspective*. Cheltenham: Edward Elgar.
Irvine, J., and B.R. Martin. 1984. *Foresight in science: Picking the winners*. London: Frances Pinter.
Kaplan, S.L., and M. Tripsas. 2008. Thinking about technology: Applying a cognitive lens to technological change. *Research Policy* 37: 790–805.
Koh, H., and C.L. Magee. 2006. A functional approach for studying technological progress: Application to information technology. *Technological Forecasting & Social Change* 73: 1061–1083.
Koh, H., and C.L. Magee. 2008. A functional approach for studying technological progress: Application to energy technology. *Technological Forecasting & Social Change* 75: 735–758.
Levinthal, D.A. 1998. The slow pace of rapid technological change: Gradualism and punctuation in technological change. *Industrial and Corporate Change* 7: 217–247.
Maine, E., M.J. Bliemel, A. Murira, and J. Utterback. 2012. *Knowledge diversity in the emerging global bio-nano sector*, Working paper no. ESD 2012-20. Cambridge, MA: MIT Engineering Systems Division.
McNerney, J., J.D. Farmer, S. Redner, and J.E. Trancik. 2012. Role of design complexity in technology improvement. *Proceedings of the National Academy of Sciences*.
Mohn, N.C. 1972. Application of trend concepts in forecasting: Typesetting technology. *Technological Forecasting & Social Change* 3: 225–253.
Mollick, E. 2006. Establishing Moore's law. *IEEE Annals of the History of Computing* 28: 62–75.
Porter, A., A.T. Roper, T.W. Mason, F.A. Rossini, and J. Banks. 1991. *Technological forecasting*. New York: Wiley.
Sahal, D. 1981. *Patterns of technological innovation*. Reading: Addison Wesley.
Saviotti, P.P., et al. 1982. An approach to the construction of indexes of technological change and technological sophistication: The case of agricultural tractors. *Technological Forecasting & Social Change* 21: 133–147.
Sood, A., and G. Tellis. 2005. Technological evolution and radical innovation. *Journal of Marketing* 69: 152–168.
Utterback, J.M. 1994. *Mastering the dynamics of innovation*. Boston: Harvard Business School Press.
Utterback, J.M., and L. Kim. 1986. Invasion of a stable business by radical innovation. In *The management of productivity and technology in manufacturing*, ed. P. Kleindorfer. New York: Plenum Press.

Technology Strategy

Robert A. Burgelman
Stanford University, Graduate School of Business, Stanford, CA, USA

Abstract

This contribution identifies and discusses four substantive dimensions – ► competitive strategy stance, value chain stance, resource commitment stance and management stance – in the formulation of an organization's technology strategy, and three key tasks – internal and external technology sourcing, deploying technology in product and process development, and using technology in technical support activities – in the enactment of its technology strategy. Based on this discussion it articulates two normative conjectures. It also identifies and discusses several key factors that drive the dynamics of an organization's technology strategy and thereby help it survive and thrive.

Definition Technology strategy concerns the generation and deployment of technological resources; that is, theoretical and practical knowledge, skills and artefacts that can be used to develop products and services as well as their production and delivery systems, for competitive advantage in business strategy.

Technology is a resource that, like financial and human resources, is pervasively important in all, not only high-technology, organizations, and needs to be managed strategically. Technology strategy provides a basis for making decisions about the sourcing, development and deployment of technological resources. It encompasses but extends beyond research and development (R&D) strategy. Following Burgelman and Rosenbloom (1989), in this article I discuss technology strategy in terms of its substance and enactment, and briefly highlight some of the determinants of its evolution.

Substance of Technology Strategy

Technology strategy can be conceptualized in terms of four substantive dimensions: (1) ► competitive strategy stance, (2) value chain stance, (3) resource commitment stance and (4) management stance.

The *competitive strategy stance* involves technological choices such as those related to product components and architectures (Henderson and Clark 1990), ► dominant design (Abernathy and Utterback 1978) and sustaining or disruptive technologies (Christensen 1997). It also involves decisions with respect to (i) technology leadership (e.g., relative advantage in the command of a body of technological competencies and capabilities, (ii) pioneering the development of new technology and so on, (iii) technology entry timing (which may depend on the appropriability regime (Teece 1986)) and (iv) technology licensing (e.g., to pre-empt competitors with alternative technologies).

The *value chain stance* recognizes that the strategic deployment of technology goes beyond its use in products and services and may be important in all the value chain activities (Porter 1985). Considering technology strategy in relation to the value chain helps distinguish between core and peripheral technologies (at a particular moment in time), with the former being key to the firm's competitive advantage. The value chain stance defines the scope of the firm's technology strategy, and may be determined to a significant extent by its scale and business focus.

The *resource commitment stance* concerns the intensity of its resource commitments to technology, which may vary widely between high-technology firms and those in which technology plays a minor role. The resource commitment stance is manifest in the depth of a firm's technology strategy, which can be expressed in terms of the number of technical options the firm has available at any moment in time.

The *management stance* involves the choices of a management approach and organization design that are consistent with the stances taken on the other three substantive dimensions. A key criterion of the management stance is organizational fit; that is, the extent to which the management approach and organization design meet the requirement flowing from the competitive strategy, value chain and resource commitment stances. For instance, a firm committed to being a technology leader and early entrant will need to create and effectively manage a central R&D function.

Enactment of Technology Strategy

Experience derived from enacting technology strategy in practice provides feedback concerning its effectiveness. Technology strategy is enacted in practice through several key tasks: (1) internal and external technology sourcing, (2) deploying technology in product and process development, and (3) using technology in technical support activities.

Technology sourcing can be internal and/or external. Relatively few companies in R&D-intensive industries can afford to do fundamental research in-house; most firms emphasize applied research in support of existing and emerging businesses. In order to be able to maintain their 'absorptive capacity' of new externally developed technologies, however, firms need to maintain some internal R&D capability (Cohen and Levinthal 1990). External sourcing of technology often takes the form of ► strategic alliances, and these require careful strategic management of interdependencies (e.g., Doz et al. 1989).

Product and process development reflect the explicit or implicit stances taken on the four substantive dimensions of technology strategy, and the relative success of product and process development produces information about its effectiveness. Wheelwright and Clark (1992) point out how firms can create a product and process development strategy framework to help them consistently integrate technology strategy with product market strategy. Iansiti (1997) takes the idea of technology integration further by focusing on technology evaluation and selection processes that precede actual product development processes and which affect the speed and productivity of these processes at the project level.

Technical support capabilities may also provide important feedback about technology in use to enhance the firm's technology strategy going forward (e.g., Rosenberg 1982). In some industries, important innovations sometimes originate with users (von Hippel 1977). Hence, two-way flows of information are relevant: expert knowledge from product developers can enhance the effectiveness of field operations, and feedback from the field informs further technology and product development.

Two normative conjectures about technology strategy seem to flow from the preceding discussion. First, a firm's technology strategy should be *comprehensive*; that is, each of the areas of enactment should be informed by the four substantive stances. Second, technology strategy should be *integrated*; that is, each of the four substantive stances taken across the various areas of enactment should be consistent.

Dynamics of Technology Strategy

To survive and thrive, a firm's technology strategy must evolve. The dynamics of its technology strategy are driven by several forces, such as the trajectories of the broader areas of technology of which the firm's capabilities are part (e.g., Dosi 1982); the interplay between product and process technology development within design configurations (e.g., Abernathy 1978); the emergence of new competing technologies (e.g., Foster 1986), sometimes with increasing returns to adoption (e.g., Arthur 1988), which may be competence-enhancing or competence-destroying (e.g., Tushman and Anderson 1986); dematurity, or renewed technological ▶ innovation in the context of established markets (e.g., Abernathy et al. 1983); speciation, or the application of existing technology to a new domain of application (e.g., Levinthal 1998); exaptation, or the success of a technology today that depends on the prior selection of a trait that was non-adaptive at the time (e.g., Dew et al. 2004; Mokyr 2000); as well as organizational determinants of ▶ technological change (e.g., Tushman and Rosenkopf 1992).

See Also

▶ Competitive Strategy
▶ Core Competence
▶ Corporate Strategy
▶ Dominant Design
▶ Dynamic Capabilities
▶ Innovation
▶ Strategic Groups
▶ Technological Change

References

Abernathy, W.J. 1978. *The productivity dilemma: Roadblock to innovation in the automobile industry*. Baltimore: Johns Hopkins University Press.

Abernathy, W.J., and J. Utterback. 1978. Patterns of industrial innovation. *Technology Review* 80: 41–47.

Abernathy, W.J., K.B. Clark, and A.M. Kantrow. 1983. *Industrial renaissance*. New York: Basic Books.

Arthur, W.B. 1988. Competing technologies: An overview. In *Technical change and economic theory*, ed. G. Dosi. New York: Columbia University Press.

Burgelman, R.A., and R.S. Rosenbloom. 1989. Technology strategy: An evolutionary process perspective. In *Research on technological innovation, management, and policy*, vol. 4, ed. R.A. Burgelman and R.S. Rosenbloom. Oxford: Elsevier Science.

Christensen, C.M. 1997. *The innovator's dilemma*. Cambridge, MA: Harvard Business School Press.

Cohen, W.M., and D.A. Levinthal. 1990. Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly* 35: 128–152.

Dew, N., S.D. Sarasvathy, and S. Venkataraman. 2000. The economic implications of exaptation. *Journal of Evolutionary Economics* 14: 69–84.

Dosi, G. 1982. Technological paradigms and technological trajectories: A suggested interpretation of the

determinants and directions of technical change. *Research Policy* 11: 147–162.
Doz, Y.L., Hamel, G., and Prahalad, C.K. 1989. Collaborate with your competitors – and win. *Harvard Business Review*, Jan–Feb, 133–139.
Foster, R.N. 1986. *Innovation: The attacker's advantage*. New York: Summit.
Henderson, R.M., and K.B. Clark. 1990. Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35: 9–30.
Iansiti, M. 1997. *Technology integration*. Cambridge, MA: Harvard Business School Press.
Levinthal, D.A. 1998. The slow pace of rapid technological change: Gradualism and punctuation in technological change. *Industrial and Corporate Change* 7: 217–247.
Mokyr, J. 2000. Evolutionary phenomena in technological change. In *Technological innovation as an evolutionary process*, ed. J. Ziman. Cambridge: Cambridge University Press.
Porter, M.E. 1985. *Competitive advantage*. New York: Free Press.
Rosenberg, N. 1982. *Inside the black box*. Cambridge: Cambridge University Press.
Teece, D.I. 1986. Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy* 15: 285–305.
Tushman, M.L., and P. Anderson. 1986. Technological and organizational environments. *Administrative Science Quarterly* 31: 439–465.
Tushman, M.L., and L. Rosenkopf. 1992. Organizational determinants of technological change. In *Research in organizational behavior*, vol. 14, ed. B. Staw and L. Cummings. Greenwich: JAI Press.
Von Hippel, E. 1978. Has a customer already developed your next product? *Sloan Management Review* 18: 63–74.
Wheelwright, S.C., and K.B. Clark. 1992. *Revolutionizing product development: Quantum leaps in speed, efficiency, and quality*. New York: Free Press.

Technology Transfer

David J. Teece
Berkeley Research Group, LLC, Emeryville, CA, USA
Haas School of Business, University of California, Berkeley, Berkeley, CA, USA

Abstract

Technology transfer is the act of conveying product- or process-related 'industrial' knowledge from one organization or subunit to another. It requires a series of activities and leads to learning by the recipient organization. Transfer is often costly in terms of financial and other resources. The cost is likely to be higher when the knowledge involved is more tacit than codified, where the know-how is less (rather than well) understood by both the transferor and the transferee, and when the transfer occurs between firms rather than between units of a single firm. Between firms, transfer of proprietary know-how is complicated because monitoring is necessary to avoid leakage of technology beyond what is required by the agreements governing the transfer.

Definition Technology transfer is the act of conveying product- or process-related 'industrial' knowledge (technical and organizational) from one organization to another, either between separate entities or within a single firm.

Technology transfer is the act of conveying knowledge from one organization or subunit to another. The knowledge will relate to a product or process, but it can be anything from a precise formula for a molecule to an unwritten set of heuristics useful for keeping a complex process within required tolerances.

Technology transfer is itself generally a process rather than an event. Teece (1977a) identified four stages: (1) pre-engineering technological exchanges; (2) transfer of process design and associated process engineering for process innovation, product design and production engineering for product innovation; (3) use of R&D to adapt and troubleshoot the technology; (4) pre-start-up training and process debugging costs. Szulanski (1996) similarly characterized it as a four-stage sequence: initiation, implementation, ramp-up and integration.

An implicit precursor to technology transfer is the identification of valuable know-how. In large firms, the identification of exchange or transfer opportunities can be difficult (Teece 2000).

Because the technology is ultimately absorbed and adapted into a new organizational context, the subject of technology transfer is increasingly studied as an aspect of the broader topic of ▶ organizational learning (Argote 1999). A special case is international technology transfer, which is often studied in terms of its impact on a foreign subsidiary of a multinational enterprise (Teece 1977a) or on the recipient country's economic development (e.g., Reddy and Zhou 1990).

Transfers entail costs that vary with the nature of the technology. The transfer will be least costly, in terms of both financial and other resources, when the technology is fully codified (e.g., a chemical formula) or fully embodied in machinery (Teece 1977a, 1981b). Other factors that affect the cost of transfer are the number of times the technology has been applied in the past, and how well the technology is understood (Teece 1977a). Put differently, there is a steep learning curve associated with technology transfer itself because each transfer is another observation of the factors that impact the performance of the technology, at least for process technology. Not surprisingly, the total cost has also been shown to increase with the desired speed of transfer, especially when the technology is relatively new (Teece 1977b).

Many elements of a technology may start out tacit but are later codified. For example, a company may determine that there are sufficient benefits to warrant the time and expense of making tacit know-how explicit. Codification, however, can increase the risk of unintentional leakage to rivals because codified knowledge is relatively easy to copy. Nevertheless, codification may be economically worthwhile because of the greater ease of replication, the greater ease of scaling, and the reduced dependence on specific knowledgeable individuals or improved communication with customers.

Even well-documented technologies will usually involve some knowledge that is tacit, accessible only by observing or interviewing those who have actually used it. These tacit aspects of a technology can only be transferred to other organizations (whether internal or external to the firm) by sending knowledgeable people (a scarce, valuable resource) for direct instruction and demonstration (Teece 1981a). Identifying the right people to convey the knowledge and oversee the transfer can itself be a demanding task (Teece 2000: 39). Furthermore, the ability of employees to transfer knowledge may be limited by contextual or social factors that differentiate the sending and receiving organizations (Argote and Ingram 2000). And when the transfer occurs between firms, any contract governing the loan or 'secondment' of employees is likely to be incomplete, so maximum flexibility is needed (Teece 1986: 30).

In general, the cost of technology transfer is less between business units within a firm than between separate firms (Teece 1976). The enterprise form of organization is very efficient for the development, internal transfer, and orchestration of differentiated organizational and technological capabilities relative to an equivalent group of separate companies linked only by market-based contracts. When transferred across separate companies, proprietary technology would need to be assigned a price, restrictions on certain uses by the recipients would need to be specified, transfer efforts would need to be spelled out in as much detail as possible and contract (licence) performance would need to be carefully monitored.

The complexity of such arrangements is one reason the ▶ multinational corporation remains a common channel for technology to be transferred across national borders (Teece 1981a). A comparison by Bloom and colleagues (Bloom et al. 2012: 12) of the diffusion of best management practices across 20 countries found that multinational affiliates as a group were always better managed than domestic, non-multinational firms. They also found that information about even the existence of best practices wasn't known by managers in local firms (Bloom et al. 2013). The management practice gap was widest in countries where competitive forces were weakest, which suggests that multinationals are generally very successful at transferring their know-how regardless of local conditions.

There are other reasons for the persistence of multinational corporations even in an era when many of the frictions affecting cross-border alliances, such as communication and travel, have been significantly reduced. One such is that careful oversight of transfers between firms is required to ensure that no sensitive information is divulged beyond what is needed to complete the task at hand, yet the monitoring procedures must not unduly impede the desired transfer. Between units of a single firm, disclosure problems are greatly reduced.

However, even within a single firm, technology transfer is a complex undertaking. Szulanski (1996) found that the greatest barriers to the successful transfer of best practices between units of a firm are weak understanding about the practice itself, poor ► absorptive capacity of the recipient unit and communication difficulties between units.

The ► licensing or sale of inventions is a type of interfirm technology transfer that involves still another difficulty. These transactions include the added problem of the paradox of markets for information: the buyer or licensor cannot fully evaluate the technology until after it has been transferred (Arrow 1962: 615). As a result, some types of information cannot be easily traded. Quite simply, where non-disclosure agreements are difficult to enforce, the market for know-how may not function very well, and technology transfer between unaffected entities might simply fail to take place.

A well-studied example of such transfers is the effort by US universities to license technologies developed by faculty to the private sector for further development. In this situation, there is a clearly demarcated gatekeeper, the university's technology transfer office (TTO), a role that would be played by a legal or similar department in a private firm. Siegel et al. (2003) found that the effectiveness of TTOs was impeded by various institutional problems, including weak incentives for inventors to participate in the licensing process, a poor understanding of technology by those in charge of negotiating with potential licensing partners, and cumbersome procedures that make it difficult to complete a licensing agreement. Each of these could also impede technology transfer by a private firm to licensees.

See Also

- ► Absorptive Capacity
- ► Foreign Direct Investment (FDI) and Economic Development
- ► Information and Knowledge
- ► Licensing
- ► Multinational Corporations
- ► Organizational Learning
- ► Outsourcing
- ► Strategic Business Unit (SBU)
- ► Tacit Knowledge

References

Argote, L. 1999. *Organizational learning: Creating, retaining, and transferring knowledge*. Boston: Kluwer.

Argote, L., and P. Ingram. 2000. Knowledge transfer: A basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes* 82: 150–169.

Arrow, K.J. 1962. Economic welfare and the allocation of resources of invention. In *The rate and direction of inventive activity: Economic and social factors*, ed. National Bureau of Economic Research. Princeton: Princeton University Press.

Bloom, N., C. Genakos, R. Sadun, and J. Van Reenen. 2012. Management practices across firms and countries. *Academy of Management Perspectives* 20: 12–33.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts. 2013. Does management matter? Evidence from India. *Quarterly Journal of Economics* 128: 1–51.

Reddy, N.M., and L. Zhao. 1990. International technology transfer: A review. *Research Policy* 19: 285–307.

Siegel, D.S., D. Waldman, and A. Link. 2003. Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: An exploratory study. *Research Policy* 32: 27–48.

Szulanski, G. 1996. Exploring internal stickiness: Impediments to the transfer of best practice within the firm. Winter special issue. *Strategic Management Journal* 17: 27–43.

Teece, D.J. 1976. *The multinational corporation and the resource cost of international technology transfer*. Cambridge, MA: Ballinger.

Teece, D.J. 1977a. Technology transfer by multinational firms: The resource cost of transferring technological know-how. *The Economic Journal* 87: 242–261.

Teece, D.J. 1977b. Time–cost tradeoffs: Elasticity and determinants for international technology transfer projects. *Management Science* 23: 830–837.

Teece, D.J. 1981a. The multinational enterprise: Market failure and market power considerations. *Sloan Management Review* 22: 3–17.

Teece, D.J. 1981b. The market for know-how and the efficient international transfer of technology. *Annals of the Academy of Political and Social Science* 458: 81–96.

Teece, D.J. 1986. Transaction cost economics and the multinational enterprise: An assessment. *Journal of Economic Behavior & Organization* 7: 21–45.
Teece, D.J. 2000. Strategies for managing knowledge assets: The role of firm structure and industrial context. *Long Range Planning* 33: 35–54.

Teece, David J. (Born 1948)

Neil M. Kay[1] and Christos N. Pitelis[2]
[1]Department of Economics, University of Strathclyde, Argyll, UK
[2]Brunel University London, Uxbridge, UK

Abstract

David Teece has made important contributions to strategy, innovation and ▶ international business (IB) and public policy, including reasons for the existence and organization of firms; how appropriability/value capture and transactional issues can influence innovative activity and value creation and the boundaries of the firm; and the role of dynamic capabilities in contributing to the sustainable competitive advantage (SCA) of firms. Recent work on dynamic capabilities integrates a considerable body of social science and management research and fashions a novel framework to help deepen understanding of the nature, essence, strategy and performance of the business enterprise.

Background

Born in New Zealand, Teece attended the University of Canterbury before going to the Wharton School of the University of Pennsylvania (M.A. 1973 and Ph.D. 1975) where he studied economics, with specializations in industrial economics/organization, international trade and technological innovation. He taught at the Stanford Graduate School of Business from 1975 to 1982 before moving to the University of California (Berkeley) where he became Professor of Business Administration (in 1982); and for over 20 years he held various academic and research positions at UC Berkeley including Director of the Institute of Business Innovation and its predecessor the Institute for Management, Innovation, and Organization, and (since 2013), the Tusher Center for Intellectual Capital.

Major Contributions to Scholarship in the Field of Strategy

Teece's work has integrated streams of transaction cost theory, evolutionary economics, and behavioural organization theory, among others (part of the intellectual history of Teece's work also detailed in work with co-authors, see for example Augier and Teece 2005). In this entry, we emphasize some major themes.

Boundaries of the Firm

Teece's early work on firm boundaries was heavily influenced by his PhD adviser Edwin Mansfield and by Oliver Williamson. Williamson (1975) had set out the basic framework for what was to become known as ▶ transaction cost economics (TCE); at its heart lay a core hypothesis that was as deceptively simple as it was potentially powerful; the more transaction-specific the assets underlying a transaction between buyer and seller, the more likely the buyer would internalize the transaction to avoid the danger of opportunistic appropriation of quasi-rents once the buyer had become committed to the transaction.

Teece's work has both extended and contributed to the TCE literature. For instance, one major lacuna in early TCE was lack of explicit recognition of internalization in the context of internationalization/ cross-border expansion, as represented by the growth of the multinational enterprise (MNE) and the choice of foreign direct investment (FDI) as opposed to licensing. This was not something that Oliver Williamson (1975) had been explicitly concerned with. This lacuna had been partially filled by scholars in ▶ international business (IB) such as ▶ Stephen Hymer (1976) and Buckley and Casson (1976). However, interpretation of what was meant by transaction costs in the IB field (described as the

T

"Internalization School" in Teece 1986a) could be eclectic and was generally developed to deal with the special case of the MNE. Teece (1981a, 1986a) approached the problem by starting with the governance perspective associated with TCE and extending it to deal with the characteristics of the nature of technology, assets and capabilities, regimes of appropriability and the characteristics of the markets in which the firm operates. Along with Teece (1980, 1982) this helped apply TCE in novel areas and demonstrated its potential as a framework that can help explain the boundaries of the firm in general and not just the special case of vertical integration and/or the MNE.

Knowledge Assets and Competitive Advantage

Another interest that evolved out of Teece's dissertation resulted in contributions to the study of IB and technology transfer and reflected the influence of Mansfield. Teece's empirical study of international technology transfer in the chemicals, petroleum and machinery industries challenged the conventional treatment of knowledge in the economics textbooks as a public good that could be transferred between uses and users at low to zero marginal cost. Contrarily, Teece found:

> The resources required to transfer technology internationally are considerable. Accordingly, it is quite inappropriate to regard existing technology as something that can be made available to all at zero social cost. Furthermore, transfer costs vary considerably, especially according to the number of previous applications of the innovation, and how well the innovation is understood by the parties involved. (1977a: 259)

The frame of reference Teece derived from these early insights (of knowledge transfer as a potentially costly activity) served him well two decades later in a series of papers on intellectual property (IP) strategy. These papers could be regarded as representing a confluence and integration of two recurring themes in Teece's work over the years; knowledge assets and capabilities as the foundation of the theory and SCA of firms (see, for instance, Grindley and Teece 1997; Teece 1998, 2000).

Complementary Assets and Technological Innovation

Teece's previous work on knowledge assets and the boundaries of the firm were to serve as part foundation for what was to become one of his most influential and cited pieces of work to date. In 1986 he published "Profiting from technological innovation: implications for integration, collaboration, licensing and public policy" in *Research Policy* (Teece 1986b).

In the paper, Teece set about creating a framework that would help innovators and innovating firms assess possibilities for the exploitation of commercially relevant inventions and/or innovation, including solely in-house strategic options versus contractual ones. This was an early exploration into what we now think of as business models for technology commercialization (Teece 2010). He recognized that there could be a gap between the conventional wisdom (that there was a direct link between firm innovativeness and commercial success) and the reality (the progenitor for what could become a successful product often failed to reap the rewards for its pioneering efforts). In principle the gap was visible, but with the information diffuse and scattered in the public domain it awaited collation and interpretation. Teece collected, collated, and noted numerous examples that apparently breached the notion of first mover advantage with the original innovator often losing out to later followers. The more difficult question was why this should happen, and to answer that question Teece weaved together separate strands comprising complementary assets, the appropriability regime, timing, and the evolution of standards, all of which drew upon sound theoretical reasoning and empirical evidence from the extant literatures in economics and strategy. Teece built on these foundations to help provide a working model and predictive framework for the exploitation and appropriation of innovative potential, the paper showing how the framework could be applied by strategists and policymakers to help formulate strategy and policy with the help of simple diagrams, 2 × 2 matrices, and flow charts.

Additional features of this paper worth noting include that the importance attributed to complementarities questioned (without explicitly mentioning) the lack of complementarities in Porter's (1980) industry-based 5-forces framework. Complementarities have proven the Achilles heel for Porter and/as it is arguable that modern network and business ecosystem-based strategy (to which Teece keeps contributing) is mostly about such complementarities. The second aspect is about international trade policy. Moreover, Teece's idea that the acquisition and leverage of complementary assets and capabilities can confer SCA to firms and that public policy can contribute, is in line with strategic trade-based ideas as in Krugman (1987) albeit of the minimally invasive type (i.e. in terms of supporting complementarities).

Dynamic Capabilities

While "Profiting from technological innovation" has made a lasting impact on the field of management in general and strategic management in particular, Teece's development of the notion of dynamic capabilities (DCs) has come to have an even more profound effect. Teece (2009: x) argues that:

> At minimum dynamic capabilities is a tool for integrating over fifty years of scholarship and empirical analysis in economics, sociology, behavioural decision theory, business history and strategic management itself. More ambitiously, it outlines a new theory of management which can be the cornerstone to a much deeper understanding of the business enterprise, competitive processes, competitive outcomes, and wealth creation in advanced post-industrial knowledge-based societies.

Teece and co-authors have written about the history and intellectual foundations for dynamic capabilities, including the relationship with transaction cost economics, evolutionary economics, neo-Schumpeterian and behavioral theory and the resource based view, and how it builds on and extends those traditions (see, e.g., Augier and Teece 2005, 2009; Teece 2009; Katkalo et al. 2010; Pitelis and Teece 2010).

The term dynamic capabilities was originally introduced into the literature in an early working paper co-authored with former students Gary Pisano and Amy Shuen (Teece et al. 1990), then developed further in Teece and Pisano (1994) and subsequently expanded into Teece et al. (1997). Teece also contextualized the concept by noting that earlier work in the resource-based theory (which included Rumelt 1984; Wernerfelt 1984; Amit and Schoemaker 1993; and some of his own work such as Teece 1980 and 1982), had tended to portray the enterprise as consisting of bundles of idiosyncratic competencies or resources (especially know-how) that could be difficult to trade or imitate, and whose ownership and utilization could generate competitive advantage at a point in time. However, he argued that sustainable advantage in modern globalized dynamic environments required more than just the ownership of such assets; it also required unique and difficult-to-replicate dynamic capabilities to "continuously create, extend, upgrade, protect, and keep relevant the enterprise's unique asset base" (2007: 1319). Teece also argued that for analytical and practical purposes DCs could be broken down into three categories: "(1) to sense and shape opportunities and threats, (2) to seize opportunities, and (3) to maintain competitiveness through enhancing, combining, protecting, and, when necessary, reconfiguring the business enterprise's intangible and tangible assets" (2007: 1319).

However, despite the efforts of Teece and others (see especially Helfat et al. 2007: 1–18) to clarify and operationalize the concept, the varieties of definitions and interpretations in the literature indicate that for some at least it is not a settled concept. Also, while the concept has been applied in a large variety of contexts and from a variety of disciplinary and methodological perspectives, this has been at the expense of a more focused and integrated development, which might have facilitated the ambitious agenda of making it both a recognized universal tool for managers and a general theory of management more realizable, at least in the medium term. To a large extent this is just to be expected; although the concept has been around since early-1990s, over half the Google Scholar citations to it have been made in just the last 5 years (to January 2016). Dynamic capabilities may be reasonably regarded as a concept that is still at the developmental stage, though to describe it as such is to run the danger of underplaying the significant impact it has already had on the field.

Some Other Contributions

Teece has made several other significant contributions to the field of strategy, including an early 1978 paper with Henry Armour, which was an early empirical contribution to the literature and perhaps the first to show a statistically significant positive linkage between organizational structure and financial performance. This was generalized across industries in Teece (1981b).

He also co-authored a contribution on the "virtual corporation" with Henry Chesbrough (Chesbrough and Teece 1996), which analysed the competitive opportunities – and dangers – of the "virtual corporation" "where businesses could outsource their business activities to a variety of partners. He and Chesbrough made a discriminating case for when innovation could be outsourced and when it could not, hence anticipating the literature on open innovation" (Chesbrough 2012).

More recently, Teece has revisited the core issues of the nature and essence of the firm and the MNE in terms of entrepreneurial shaping of the market (Augier and Teece 2009), and later market and business ecosystem creation and co-creation (Pitelis and Teece 2009, 2010; Teece 2014). This turned on its head the economics-based market failure approach and brings together core themes of entrepreneurship, strategy and IB, not least complementarities and business ecosystems that Teece helped found. It also built on Frank Knights and ► Edith Penrose's insights, extending the role of entrepreneurs to the simultaneous creation and evolution of markets and firms (Augier and Teece 2007: 182). In Katkalo et al. (2010) moreover the three main categories of DCs are linked explicitly to value creation and capture, a theme followed in subsequent work on business model innovation.

Teece (2010) contributed to the emergent stream of discourse in the strategy field on business models (which describe "the design or architecture of the value creation, delivery and capture mechanisms employed", p. 191) by relating them both to strategic conceptualizations and economic theory. Like much of his work, the article synthesized and integrated apparently disparate threads into a coherent narrative and sought to popularize and communicate complex but deceptively simple arguments without jeopardizing their essential integrity. His work on what he calls "next generation competition" (Teece 2012) also endeavors to explain how the nature of competition itself has changed, and how scholars and practitioners alike should reconceptualize it.

Public Policy

An economist at heart, Teece was throughout interested in anti-trust and public policy. He has made numerous contributions on anti-trust more relevant for this entry being arguably the article with Jorde (Jorde and Teece 1992). The article noted the importance of non-collusive aspects of inter-firm co-operation and their concomitant implications on innovation. These need to be taken into account for an informed anti-trust decision and legislation, lest an exclusive focus on market shares and concentration throws out the baby (value creation though non-collusive co-operation) with the bathwater (the increase in market share). Here too, Teece parts with the original IO-based approach of Bain (1956) and their strategy counterpart in Porter (1980).

Place in the Strategy Field and the Wider Context

Influences

As noted, Teece's Ph.D. thesis advisor Edwin Mansfield encouraged him to study technology transfer. As Teece recalls:

> As a graduate student at Penn in the early 1970s, I was fortunate to end up in his Ph.D. class on the economics of technological change. He opened my eyes to a set of issues for which I had no previous exposurey. No one at that time, including Ed, knew much about the topic. (Teece 2005: 17)

Later in his career Teece came to appreciate the contributions of Richard Nelson and Sidney Winter, Herbert Simon, James March and Nate Rosenberg and Alfred Chandler. These scholars are amongst Teece's acknowledged mentors (for more detailed accounts of the intellectual foundations for Teece's work see also Augier and Teece 2005, 2007, 2009). The 2005 paper explicates the different intellectual lines, including transaction cost theory, behavioural and evolutionary theory.

Reception of Work

The reception of Teece's work is manifested in several professional awards as well as in citations. Teece has been recognized in a study by Accenture as one of the world's top 50 living business intellectuals, defined in the study as "influential thinkers and writers on business management topics" (Accenture 2002). The list was produced on the basis of the sum or ranks in three categories; web hits, media mentions and scholarly citations. He was one of 30 US business professors named in the "A-List of Management Academics 2011," the list was compiled by Business Educators, a US-based private organization on the basis of professional accomplishments, renown in their field of expertise and dedication to sharing their knowhow with others (Haas 2011). Teece (2009) was selected as one of the best books of 2009 by the management magazine *Strategy + Business*.

With circa 95,000 citations in Google Scholar (5 January 2015) and an h-index of 90 (indicating that 90 of his publications had been cited at least 90 times each), the recognition of Teece's work is on a par with many a Nobel laureate. His most cited publications are "Dynamic capabilities and strategic management" (with Gary Pisano and Amy Shuen, 1997, with circa 25,000 citations already, also listed as the most cited paper in Economics and Business, 1995–2005, by *Science Watch* (Teece 2009: xi); and "Profiting from technological innovation y" (Teece 1986b). The latter was judged by the editors as one of the best papers published by Research Policy 1971–1991 and identified in 1999 by its editors as the most cited paper ever published in the journal. Both articles have been reprinted numerous times in other publications.

Science Watch (May 2008) ranked Teece in the top ten most-cited researchers world-wide in economics and business (1997–2007). He was identified as in the Top 10 Most Influential Scholars in Management Based on Citations, Academy of Management Perspectives (Aguinis et al. 2012).

In addition, Teece has received awards and numerous honorary doctorates, including from St Petersburg State University Russia (2000); Copenhagen Business School, Denmark (2004); Lappeenranta University of Technology, Finland (2004); and University of Canterbury New Zealand (2007). He is also a Fellow of the Strategic Management Society, and a recipient of the Eminent Scholar award (2013) and Honorary Fellow of the Academy of International Business. Finally, in addition to his academic activities Teece has played major roles in creating two service sector MNEs, first co-founding and chairing for two decades the Law and Economics Consulting Group (LECG) in 1988, and later founding and chairing the Berkeley Research Group (BRG) in 2010.

Summary

Teece has made a seminal contribution to the field of strategy and particularly in relation to where strategy thinking, is, or can be, rooted in the economic theory of the firm. The consensus in the field (supported by the weight of citations) would seem to be that, at least to date, his dominant contribution would be the notion of dynamic capabilities. But there is a paradox here in that although Teece approaches these issues from an economics perspective, his influence has been far greater outside economics rather than inside. This reflects self-imposed myopia in the economics profession in recent years and it is a fate shared by other leading economists in this area such as Penrose (1959) and Nelson and Winter (1982), Teece's recent re-engagement with Economic Journals (Teece 2015) seems to reflect a welcome change.

However, one emergent issue is that many of those who have subsequently pursued the notion of DCs in scholarly publications, teaching and practice are not themselves trained as economists and can have different interpretations of the nature, characteristics and significance of DCs. This further distancing from economics can be reinforced by the difficulties that can be encountered in this context in attempts to apply positivist hypothetico-deductive approaches, and the generally accepted need to invoke other methodologies that can be unfamiliar to economists. At best, this may lead to a vibrant and rich mix of multidisciplinary concepts, methodologies, applications and findings stimulated by Teece's seminal work in this field. At worst, some of the original sharpness, focus and insight here may be lost in its translation across disciplinary interfaces, in much the same way that the younger Teece (1977a, b) discovered unexpected and significant costs in transferring knowhow across international interfaces.

Teece's conceptualization of DCs has the potential to unify and integrate disparate approaches to the nature and functions of the firm in economics and management. The scope and variety of current efforts to develop and apply the concept across various areas of the field of management suggests that this is a promise that many think realistic and worth pursuing.

See Also

- ▶ Capturing Value from Advantages
- ▶ Hymer, Stephen Herbert (1934–1974): The MNE and International Business
- ▶ International Business
- ▶ Penrose, Edith T. (1914–1996)
- ▶ Transaction Cost Economics
- ▶ Williamson, Oliver E. (Born 1932)

References

Accenture. 2002. Accenture study yields top 50 'business intellectuals' ranking of top thinkers and writers on management topics. Accenture Institute for Strategic Change press release (22 May), Cambridge, MA.

Aguinis, H., I. Suárez-González, G. Lannelongue, and H. Joo. 2012. Scholarly impact revisited. *Academy of Management Perspectives* 26: 105–132.

Amit, R., and P.J.H. Schoemaker. 1993. Strategic assets and organization rent. *Strategic Management Journal* 14: 33–46.

Armour, H.O., and D.J. Teece. 1978. Organizational structure and economic performance: a test of the multidivisional hypothesis. *Bell Journal of Economics* 9: 106–122.

Augier, M., and D. Teece. 2005. Competencies, capabilities and the neo-Schumpeterian tradition. In *The Elgar companion to Neo-Schumpeterian economics*, ed. H. Hanusch and A. Pyka, 267–286. Cheltenham: Edward Elgar.

Augier, M., and D. Teece. 2007. Dynamic capabilities and multinational enterprise: Penrosian insights and omissions. *Management International Review* 47: 175–192.

Augier, M., and D. Teece. 2009. Dynamic capabilities and the role of managers in business strategy and economic performance. *Organization Science* 20: 410–421.

Bain, J.S. 1956. *Barriers to new competition*. Cambridge, MA: Harvard University Press.

Buckley, P.J., and M. Casson. 1976. *The future of the multinational enterprise*. London: Holms and Meier.

Chesbrough, H. W. 2012. Open innovation: where we've been and where we're going. *Research-Technology Management*, July–August, pp 20–27.

Chesbrough, H.W., and D.J. Teece. 1996. Organizing for innovation: when is virtual virtuous? *Harvard Business Review* 74: 65–73.

Grindley, P.C., and D.J. Teece. 1997. Managing intellectual capital: licensing and cross-licensing in electronics. *California Management Review* 39: 8–41.

Haas. 2011. Prof. David Teece makes a-list of management academics, *Haas Now*, January 17, University of California Berkeley. Available at: http://newsroom.haas.berkeley.edu/article/prof-david-teece-makes-list-management-academics.

Helfat, C.E., S. Finkelstein, W. Mitchell, M.A. Peteraf, H. Singh, D.J. Teece, and S.G. Winter. 2007. *Dynamic capabilities: understanding strategic change in organizations*. Oxford: Blackwell.

Hymer, S.H. 1976 [1960]. *The international operations of national firms: a study of direct foreign investment*. Cambridge, MA: The MIT Press.

Jorde, T.M., and D.J. Teece (eds.). 1992. *Antitrust, innovation, and competitiveness*. Oxford: Oxford University Press.

Katkalo, V.S., C.N. Pitelis, and D.J. Teece. 2010. Introduction: on the nature and scope of dynamic capabilities. *Industrial and Corporate Change* 19: 1175–1186.

Krugman, P.R. 1987. Is free trade passe? *Journal of Economic Perspective* 1: 131–144.

Nelson, R.R., and S.G. Winter. 1982. *An evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.

Penrose, E. 1959. *The theory of the growth of the firm*. Oxford: Basil Blackwell.

Pitelis, C.N., and D.J. Teece. 2009. The (new) nature and essence of the firm. *European Management Review* 6: 5–15.

Pitelis, C.N., and D.J. Teece. 2010. Cross-border market co-creation, dynamic capabilities and the entrepreneurial theory of the multinational enterprise. *Industrial and Corporate Change* 19: 1247–1270.

Porter, M. 1980. *Competitive strategy: techniques for analyzing industries and competitors*. New York: Free Press.

Rumelt, R.P. 1984. Towards a strategic theory of the enterprise. In *Competitive strategic management*, ed. R.B. Lamb. Englewood Cliffs, NJ: Prentice Hall.
Wernerfelt, B. 1984. A resource-based view of the firm. *Strategic Management Journal* 5: 171–180.
Williamson, O.E. 1975. *Markets and hierarchies: analysis and antitrust implications*. New York: Free Press.

Selected Works

Teece, David. 1977a. Technology transfer by multinational firms: the resource cost of transferring technological know-how. *The Economic Journal* 87: 242–261.
Teece, David. 1977b. Time–cost tradeoffs: elasticity estimates and determinants for international technology transfer projects. *Management Science* 23: 830–837.
Teece, David. 1980. Economies of scope and the scope of the enterprise. *Journal of Economic Behavior and Organization* 1: 223–247.
Teece, David. 1981a. The multinational enterprise: market failure and market power considerations. *Sloan Management Review* 22: 3–17.
Teece, David. 1981b. Internal organization and economic performance: an empirical analysis of the profitability of principal firms. *Journal of Industrial Economics* 30: 173–199.
Teece, David. 1982. Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior and Organization* 3: 39–63.
Teece, David. 1986a. Transaction cost economics and the multinational enterprise: an assessment. *Journal of Economic Behavior and Organization* 7: 21–45.
Teece, David. 1986b. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research Policy* 15: 285–305.
Teece, David. 1998. Capturing value from knowledge assets: the new economy, markets for know-how, and intangible assets. *California Management Review* 40: 55–79.
Teece, David. 2000. Strategies for managing knowledge assets: the role of firm structure and industrial context. *Long Range Planning* 33: 35–54.
Teece, David. 2005. Technology and technology transfer: Mansfieldian inspirations and subsequent developments. *Journal of Technology Transfer* 30: 17–33.
Teece, David. 2006. Reflections on 'profiting from innovation'. *Research Policy* 35: 1131–1146.
Teece, David. 2007. Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal* 28: 1319–1350.
Teece, David. 2009. *Dynamic capabilities and strategic management: organizing for innovation and growth*. Oxford: Oxford University Press.
Teece, David. 2010. Business models, business strategy and innovation. *Long Range Planning* 43: 172–194.
Teece, David. 2012. Next-generation competition: new concepts for understanding how innovation shapes competition and policy in the digital economy. *Journal of Law Economics and Policy* 9: 97–118.
Teece, David. 2014. A dynamic capabilities-based entrepreneurial theory of the multinational enterprise. *Journal of International Business Studies* 45: 8–37.
Teece, David. 2015. Dynamic capabilities and entrepreneurial management in large organizations: toward a theory of the (entrepreneurial) firm. *European Economic Review*. https://doi.org/10.1016/j.euroecorev.2015.11.006
Teece, David, and G. Pisano. 1994. The dynamic capabilities of firms: An introduction. *Industrial and Corporate Change* 3: 537–556.
Teece, David, G. Pisano, and A. Shuen. 1990. *Firm capabilities, resources and the concept of strategy*. In Consortium on competitiveness and cooperation working paper No. 90–9. Berkeley, CA: Center for Research in Management, University of California.
Teece, David, G. Pisano, and A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509–533.

Theory of the Firm

David J. Teece
Berkeley Research Group, LLC, Emeryville, CA, USA
Haas School of Business, University of California, Berkeley, Berkeley, CA, USA

Abstract

The theory of the firm is a broad topic area encompassing frameworks designed to answer a number of questions about firms, including why they exist, how their boundaries are determined, how the differing interests of owners and managers can be reconciled, how firms should be organized internally for efficiency and why performance outcomes differ between firms. Work in this field crosses disciplines such as economics and strategic management that start from fundamentally different assumptions about organizational behaviour. The theory adopted has practical implications for managers and policy-makers.

Definition The theory of the firm is a general topic encompassing models that seek to answer a number of questions about firms, including why they exist, what determines their boundaries, how the differing interests of owners and managers can be aligned, how firms should be organized internally for efficiency and why firms differ.

Theories of the firm are a cluster of economic and organizational models that seek to explain a number of fundamental questions in economics and strategic management. These include (1) why do firms exist? (2) what determines the boundary between the firm and the market? (3) how can owners control the firm's activities when owners and management are not the same? (4) how should firms be organized internally for efficiency and growth? and (5) how do individual firms develop and sustain competitive advantage? In most cases, the questions are dealt with in isolation, and authors generally apply their own disciplinary lens to arrive at an answer. This has made for a vast, and sometimes bewildering, literature.

The defining characteristics of firms are often seen differently in literatures that span economics, organization theory and strategic management. In one of the foundational and most influential articles in economics, Ronald Coase (1937: 388) highlighted the 'entrepreneur-co-ordinator, who directs production' as the defining characteristic of organizing economic activity within firms versus market contracting with individuals or other firms to get the job done. This paradigm was called into question by Alchian and Demsetz (1972). Others have argued that the firm was composed of 'a set of contracting relationships among individuals' (Jensen and Meckling 1976: 310) so that the difference with the market was one of degree rather than of kind. Masten (1988) disputed this formulation by noting that the legal system in which firms are embedded imposes 'substantial differences in the obligations, sanctions, and procedures' that govern internal and external relationships in ways that 'alter the incentives of actors . . . in a meaningful way' (p. 196). Williamson (1994) also notes that internal forces help to maintain authority within the firm because pressure on an employee who rejects the authority of management will come not only from management but also from other employees who see their fortunes as tied to those of the firm.

Existence of the Firm

In a well-functioning market economy, transactions amongst individuals could, in theory, be used to organize anything and everything. The question of why firms exist is thus non-trivial. For a long time, since at least Adam Smith's description of specialization in a pin factory (Smith 1776, I. 1.3), the dominant theory (often implicit in treatments of the firm) was technological determinism; firms exist because all parts of the production process need to be together.

The technological argument, however, does not stand up to rigorous scrutiny, because the assets and functions needed for production could, at least in theory, each be owned by separate firms or individuals, all of whom contract together and achieve the necessary coordination through a constellation of contracts. In fact, the increased use of outsourcing in recent decades has shown that the contracting counterfactual has some validity. The question is how much.

One of the first people to propose an alternative approach to a technological or internal process explanation was Ronald Coase (1937) mentioned earlier. Coase suggested that firms exist to economize on the transaction costs that would be incurred to form the equivalent network of contracts among independent entrepreneurs envisioned in the market-based counterfactual. Coase recognized, however, that firms as organizations have their own limitations, namely, that managers are limited in the number of activities they can manage efficiently. Combining these two concepts, Coase concluded that an activity would be integrated (i.e., internalized) by the firm when the cost associated with doing so is less than the costs that the firm would bear by having the same activity performed by an independent agent.

Oliver Williamson (1975) expanded on Coase's idea of contracting costs to create a predictive model. He noted that one or more parties to a contract become vulnerable because investments might be made in transaction-specific locational, physical, or human capital. The specificity that causes potential problems is transactional; that is, the investments made by one (or both) parties are specialized to the particular transaction such that they have much lower value in alternative uses. Thus, if Party A makes an idiosyncratic investment to support a contract with Party B, who later exploits ambiguities in the contract (remembering that no contract is airtight), then the value of A's investment could decline. In other words, asset specificity, either current or anticipated, imposes a potential loss in value and hence a transaction cost unless the activity takes place inside the firm, where incentives are better aligned.

Williamson's transaction cost framework thus assumes that there is a better coalignment of interests inside the firm than in a contracted relationship. Hence, internal organization is the preferred organizational arrangement when business requires large investment in capital that is purpose built and dedicated to the economic problem at hand.

A variant on Williamson's approach is the pure contracting, or property rights, approach, which doesn't rely on the use-specificity of assets. Instead, the property rights approach focuses on residual rights of control; the owner of an asset is the one who makes decisions about its disposition in circumstances not covered by the (incomplete) contracts that otherwise determine its use (Grossman and Hart 1986). The residual rights also enable the asset owner to extract profits in ex post negotiation. Since asset owners know this in advance, asset ownership determines the types of complementary, production-enhancing investments that will be made (e.g., the amount of effort that may be exerted). The property rights approach assumes that many such complementary investments cannot be supported by contractual arrangements. In some cases, to ensure that the optimal amount of complementary investments are made, it will be necessary for all assets to be owned by a single entity, namely, a firm.

While transaction cost and property rights theories highlight important considerations, they set aside other features that distinguish a firm as a way of organizing from a set of market-based contracts. To correct this deficiency, alternative theories have emerged, primarily in the strategic management field. These explain the existence of firms in terms of entrepreneurship, knowledge and capabilities.

Entrepreneurial theories of the firm (e.g., Sautet 2000) start from a more primitive initial state than the one assumed in the transaction cost or property rights approaches. In the Coase-Williamson framework, for example, markets, technologies and prices are assumed to exist already (Boudreaux and Holcombe 1989). In reality, entrepreneurs must first cut through uncertainty and create each market before preferences production and prices are specified that can lead to market activity, an observation that dates back to at least the work of Frank Knight (1921). However, entrepreneurship by itself can lead to the creation of countless small firms that can contract with each other to accomplish complex tasks.

Additional assumptions are needed to produce larger firms. Foss et al. (2007), for example, expand on a pure entrepreneurship model by adding multiple types of capital, so that the primary reason for the formation of firms is to allow entrepreneurs to experiment with different combinations of heterogeneous capital.

A different approach sees knowledge and/or organizational capabilities as the source of indivisibilities that give rise to firms. Theories in this vein view a firm, especially an established firm, as more than the sum of its parts more than what individual employers, managers, or machines bring to it at any point in time. Early representations of firms as knowledge-based entities can be found in Teece (1980, 1982), Winter (1988), Teece et al. (1990), Nonaka (1991), and Kogut and Zander (1992).

In a knowledge-based theory of the firm (e.g., Conner and Prahalad 1996), the organizational form of a transaction (internal or contracted) determines 'how the parties' starting knowledge endowments are blended and used' and 'how learning or developments occurring during the course of the work are taken into account' (Conner and Prahalad 1996: 484). In other words, knowledge is more likely to be freely shared and exploited within firms than between them. Teece (1976, 1977) also showed that international technology transfer costs were generally lower inside than across firms.

While the knowledge theory of the firm generally does not address the concerns about opportunism featured in the transaction cost and property rights frameworks, it fits relatively comfortably with a transaction cost theory. In some cases, knowledge considerations will be more prominent than risks associated with opportunism, and in other cases, the opposite will hold true.

An emerging theory of the firm that leverages such complementarities among various approaches is the ► dynamic capabilities framework (Teece et al. 1997; Teece 2010, 2014a). Dynamic capabilities reflect the firm's ability to integrate, build and reconfigure internal and external resources over time. Although it is not yet fully elaborated as a theory of the firm, the dynamic capabilities approach brings transaction costs, resources, and knowledge together in a way that can potentially explain not only why firms exist, but also, as will be discussed below, their scope and potential for growth and sustained profitability (► competitive advantage).

In explaining the existence of firms, the dynamic capabilities approach emphasizes asset availability, asset development through R&D and learning, asset combination and asset deployment. Many assets, especially intangible assets, are not traded much or at all. For example, a firm with a certain type of unpatentable know-how may not be able to license it to potential users without revealing so much that the user can employ the know-how without taking a licence (Arrow 1962). This and other types of market failure can drive firms to use ► business models that employ the technology internally rather than licensing it to others. To execute profitability on this approach, they must build complementary assets in-house and sell products that use the know-how in question.

An acceptable theory of the firm should be able to explain more than simply why firms exist in a market economy. At minimum it should also be able to explain firm boundaries (i.e., what's done inside and what's outsourced and/or left to others, how they are organized and how firms grow and prosper). Each of these is now addressed in turn.

Boundaries of the Firm

There are two levels of analysis at which the boundaries of the firm are typically considered. The amount of outsourcing versus in-house 'production' has received the most attention, but there is also another question about the range of businesses in which a firm will choose to compete, that is, the ► scope of the enterprise.

Early analyses of firm boundaries (e.g., Robinson 1934) attempted to answer relevant questions by analysing the optimal size of a single-product firm. One of the answers was that single-product firms cannot scale infinitely because, as they get larger, they tend to be slower at reaching decisions, which imposes performance penalties. This is the cost of bureaucracy. As organizations expand, the ability of top management to have all the information needed for effective decision making decreases. Multidivisional (M-form) structures help in this regard by pushing many operational decisions to lower levels in the organization. However, employees often have reasons to distort the information they provide to upper management in ways that they think will benefit them and not necessarily the organization (Williamson 1985; Rumelt 1987; Milgrom and Roberts 1990). As a result, a corporation's decision making is likely to become increasingly divorced from realities. Customer experience deteriorates with increased centralization and the deepening of managerial hierarchies. This effect may be mitigated if the organization's culture favours the sharing of information.

Coase's (1937) article touched on these issues by introducing the costs of transacting in the market against the cost of bureaucracy. He also introduced vertical and horizontal integration into the discussion, bringing it closer to the modern understanding of firm boundaries. In his approach, managers rank activities in terms of transaction costs. This decision rule was that managers should internalize transactions up to the point where the marginal cost of bureaucracy is equal to the marginal cost of transacting in the market.

Coase's analysis overlooks product-specific technological concerns. For example, some complementary activities have more need to be integrated than others and (vertical) integration is more likely to be preferred when unstructured (non-modular) technical dialogue is needed between two stages of production (Monteverde

1995). In many cases, however, complementarities do not require internalization to be successfully managed. This can occur because intrinsic or engineered modularity provides a clean interface that allows the complements to develop autonomously. Combining modular activities in a single firm is generally inefficient because managerial bureaucracy cannot replicate the high-powered incentives of arm's-length contracting. As insightful as Coase's work was, at best, the model in Coase (1937) implied an incomplete decision rule for firm size because it compared marginal transaction and bureaucracy costs without comparing these with marginal revenue (Kay 2015).

More than four decades after Coase's article, Williamson (1981, 1985) picked up the threads of Coase's argument and repositioned them in an analysis of discrete comparisons for each transaction a firm might conduct. To Coase's transaction costs, Williamson added asset specificity and certain institutional features of markets and hierarchies, providing a more discriminating account of the integration decision. In a world of ▶ incomplete contracts, that is, where all possible contingencies cannot be covered, asset specificity leads to recontracting hazards. The decision rule for managers is to integrate when efficient production requires making transaction-specific investments. Monteverde and Teece (1982) were the first to show that this framework had some predictive power.

To this relatively static view of firm boundaries, Teece (1982) added an element of firm expansion. Picking up on Penrose's (1959) idea of underutilized resources as the driver for firm growth, Teece noted the difficulty in transferring knowledge resources, which are often among a firm's most valuable, to outside licensees. To overcome these difficulties, a firm ready to expand will tend to apply its underutilized knowledge in-house and enlarge its range of activities, which implies that there will likely be at least a pair-wise linkage between a firm's business units (Teece et al. 1994).

A problem with the various transaction cost-oriented approaches to firm boundaries is that, in comparing internal organization with external contracting, they (implicitly) hold other things equal. Production costs, however, may depend endogenously on how a firms are organized. This can be true, for example, if demand is inadequate to justify internal manufacture of a component by one downstream user, but large enough to justify efficient production by a supplier who can sell to all downstream users.

Another weakness of transaction cost-focused theories of the firm is the (again implicit) assumption that the goal is merely to design firms efficiently. In strategic management, the dynamic capabilities framework recognizes that boundaries are unlikely to be efficient in a transaction-cost sense because innovation and growth are as important as efficiency. Whereas the transaction cost approach uses tight assumptions to allow specific predictions that hold true in aggregate, the dynamic capabilities approach to firm boundaries is more general and more prescriptive, with no prediction that growing firms will necessarily optimize their boundaries focusing single mindedly on efficiency.

One can blend transaction costs and capability considerations into a hybrid framework that has come to be known as 'profiting from innovation' (Teece 1986). The ▶ profiting from innovation (PFI) framework is applicable to business model selection and the design of organizational boundaries for any line of business new to the company. Entering a new market generally requires, in order to deliver a solution that customers will value, that the resources of the firm are combined with ▶ complementary assets, reflecting capabilities such as marketing and distribution, which the firm may not yet control. The firm must decide whether to contract, buy or build the complements that its business model requires.

The PFI framework highlights strategic considerations around appropriability focusing on where to position the firm not just to minimize transactions costs but to capture the most profits. The general rule is to figure out which parts of the value chain will become bottlenecks, own or reshape those, and access other assets through contractual arrangements.

From the transaction cost perspective, PFI considers high levels of asset specificity or hold-up risk by a potential complementor to be indicators of a bottleneck. If the complement is cospecialized with the focal innovation (i.e., there is bilateral asset specificity), then market contracting may raise the risk of value leakage.

A similar logic applies when the market for the supply of the complement is not highly competitive, or does not even exist. The firm owning a bottleneck asset may not only draw profits away from the focal firm; it may use its specialized knowledge of the asset to pace, redirect, or control the development of new capabilities in ways that impede the timely evolution of the focal firm's offering to consumers (Chesbrough and Teece 1996).

Ownership and Control of the Firm

A different dimension of the theory of the firm concerns its control. In sole proprietorships, partnerships, and closely held corporations, there is no meaningful distinction between owners and managers. In most large corporations, however, ownership belongs to a more or less fragmented group of shareholders, while day-to-day control is exercised by professional managers who may or may not own a significant number of shares in the firm. This raises the possibility that managers could choose to operate the firm in ways that benefit themselves rather than the shareholders. This possibly animates other theories of the firm.

Concerns about the potential for misallocation of resources by non-owner managers dates back to at least the work of Berle and Means (1932). In the 1960s, a flurry of books by economists such as Williamson (1964), Marris (1964), and Baumol (1967), expanded on the Berle and Means concern about incentive misalignment between managers and shareholders.

In the finance literature, Jensen and Meckling (1976) offered a different conceptualization of the firm based on its financial structure, that is, the balance between the firm's use of equity (stock) and debt (bonds). Adopting the logic of principal-agent theory (Ross 1973), they postulated that misalignments in the objectives and information sets of the principal (owners) and the agent (managers) impose agency costs such as contracting and monitoring expenses. Their solution relies on a trade-off between the agency costs of equity financing (which weakens the incentives for managers by reducing their ownership) and the agency costs of debt (which strengthens incentives for managers but can lead them to pursue overly risky strategies). Total agency costs are minimized when the marginal agency cost of additional debt equals the marginal agency cost of additional equity. Their framework produced vistas into the financial structures of firms, but fell far short of a full blown theory of the firm. Moreover, as Holmstrom and Tirole (1989) point out, it is unrealistic to model the debt-equity ratio as the best, or only, means of shaping management behaviour. Incentives such as stock options and performance bonuses clearly must play a role, and these have also been studied extensively using the principal-agent approach (e.g., Wang 1997).

An obvious potential problem with principal-agent theories of owners and management is that they fail to capture enough of the complexity of real firms. In most large public corporations, for instance, a board of directors sits between shareholders and management, exercising oversight on behalf of the former. Yet the board typically incorporates representatives of both shareholders and managers, making it only an imperfect instrument of shareholder preferences. And management itself is not a unitary 'agent' but rather a group composed of the CEO and a top management team, requiring that the potential detrimental effects of internal pay disparities be taken into account (Carpenter and Sanders 2002; Siegel and Hambrick 2005).

The dynamic capabilities approach provides the outlines of a theory of the firm that sidesteps the drawbacks of principal-agent models by emphasizing the role of managers in building and maintaining organizational capabilities and achieving continuous renewal (Teece 2012, 2014b). The risk of self-interested behaviour by managers is not ignored, but it is of secondary concern relative to building capabilities and orchestrating firm specific assets. Appropriate incentive systems and board oversight are recognized as desirable. However, the most important job of ▶ incentive design is to reward innovation and empower creative contributions from managers and 'expert talent' (Teece 2011).

Transactions costs, ▶ agency theory, and capability theory each have a role to play in a robust theory of the firm. Contractual problems need to be avoided; owners (i.e., shareholders and their representatives on the board) must find ways to prevent managerial excess and fraud without undermining the motivation of managers to guide the firm in hypercompetitive markets; capabilities need to be built and leverage to ensure profitability and long-run survival.

Internal Organization of the Firm

There is also a large literature on how firms should be organized internally. Here I will focus on aspects of organization that flow from models of the firm.

One of the fundamental theoretical divisions about the internal operations of firms concerns authority. In some theories of the firm, authority is taken as a given, because it serves as the obvious difference between the firm and the market. In other theories of the firm, the employees of the firm are just as much contractors with the firm's owners as any economic agent outside the firm would be, with the main difference being that inside the firm employees generally have no claim on the firm's productive assets apart from their own human capital. A further difference is that employers can specify the actions that employees must take on the date they are needed rather than at the start of the contract (Simon 1951).

To the extent that a firm is just a bundle of contractual relations between employees and owners, transaction cost and agency concerns apply. Agency models have explored a range of incentive design issues in the employment relation, such as the interactions between persuasion, authority and pay-for-performance (Van den Steen 2009).

The agency problem with employees is complicated by the fact that output in many instances results from group actions, making it difficult to match rewards to individual effort. This is known as 'the metering problem' (Alchian and Demsetz 1972). Close monitoring of individuals is one way to detect shirking, and various measures can help to reduce it.

Wages are often modelled as a means of reducing the likelihood of shirking and low effort by employees. In 'efficiency wage' models, firms are willing to pay more than the market clearing wage to reduce shirking and, hence, the need for monitoring (Yellen 1984). For the economy as a whole, the high wage leads to unemployment, the fear of which makes employees less likely to exercise low effort.

Instilling a sense of group loyalty is another tool that management can wield. The availability of opportunities for promotion can also increase the commitment and productivity of employees (Williamson et al. 1975; Prendergast 1993).

A different problem than shirking is the channeling of employee efforts toward unproductive internal lobbying. Even committed employees may seek to influence decisions about raises and access to resources. As a result, the organization may need to veer away from the organizational design that would be optimal in the absence of influence activities. Possible organizational adjustments include limiting communication by employees to management, pre-committing to seniority or other rules for promotions, and more carefully matching wages to the effort requirements for each job (Milgrom and Roberts 1988).

As was the case with ownership and control issues, the dynamic capabilities framework adopts a more forward-looking approach to incentive design. Organizational structures should be more 'organic' (decentralized, informal, expert-focused) than 'mechanistic' (centralized, bureaucratic, title-focused) in order to promote innovation and growth (Burns and Stalker 1961). Pay for performance is desirable, but an organizational culture that encourages knowledge creation and sharing, the ability of management to articulate a compelling vision, and employee autonomy is an important means for promoting innovative and entrepreneurial activities by a firm's employees.

Competitive Advantage of the Firm

The reasons why some firms are more successful than others are, oddly enough, more or less unexplored in economic theory, and the theory of the firm is no exception. Rather, economic theory and the theory of competitive markets proceeds as though all firms are the same, and that a 'representative' firm can somehow be identified or assumed.

Managers, whose particular decisions account for much of each firm's unique history and character, have been virtually absent from most economic analysis since Adam Smith (1776). In industrial organization economics, the cost or quality advantages that give some firms the ability to maintain price premiums are simply assumed to exist. Industrial organization theory does not have a positive framework to explain superior performance beyond socially negative forms of business conduct (e.g., conspiracies or raising rivals' cost). There is occasional reference to 'superior foresight' by management (e.g., Gilbert and Newbery 1982: 525), but no explanation of what that might entail. Patents and standards are among the few assets or activities that economists have recognized as potentially explaining superior firm-level performance. The transaction cost, property rights and agency cost approaches to the firm are also silent about how firms identify and exploit complementarities and develop competitive advantage. This is a remarkable state of affairs which ought to bother policymakers, who base much of antitrust policy on industrial organization models. The situation could be remedied and public policy would benefit if antitrust/competition policy scholars would tap into the very substantial scholarship outside the narrow confines of the field of industrial organization.

In the strategic management literature, the development of unique advantages by any number of means is recognized as one of the primary goals of managers. Moreover, in the field of strategic management there are now well-developed theories as to how they arise. In particular, the dynamic capabilities framework has emerged as one of the leading perspectives on developing and maintaining competitive advantage (Di Stefano et al. 2010).

In the dynamic capabilities framework, identifying opportunities, recognizing complementarities, securing combinations of assets that will enable the enterprise to address customer needs and developing hard-to-imitate business models are key to the firm's growth and survival. These activities implicate both managers and the capabilities of the organization as a whole (Teece 2014b).

Firm capabilities govern what the organization can potentially accomplish, not necessarily what it is currently producing. They result from a combination of prior investments and unique organizational histories. Sustained 'supernormal' profitability is possible because factor markets for certain types of assets (particularly intangibles and idiosyncratic physical and human assets) are not fully efficient, which in turn makes a firm's resources hard to imitate in full (Barney 1991).

For analytical convenience, capabilities can be divided into two types. Ordinary capabilities allow the organization to perform well-delineated tasks in the areas of administration, operations, and governance. Dynamic capabilities govern the organization's other activities and include high-level functions such as ► opportunity recognition, strategy implementation, and reconfiguration of the firm's resource base.

The strength of a firm's dynamic capabilities determines the speed and degree to which the firm's idiosyncratic resources can be aligned with the firm's changing needs over time. Strong dynamic capabilities, combined with difficult-to-imitate (idiosyncratic) resources and good strategy, can generate (and regenerate) competitive advantage, leading to sustained (durable) profitability, thereby increasing the chances of long-run survival.

Conclusion

A full understanding of the nature, organization and profitability of firms requires multiple theoretical perspectives. Economic and finance models tend to ignore research on the more qualitative aspects of organization such as knowledge creation. Capability based models from the field of (strategic) management tend to downplay the risks of opportunistic behaviour. While the results are not elegant, combining these theoretical perspectives can help scholars and managers alike understand today's very complex firms that are the heart and soul of modern capitalist economies.

See Also

- ▶ Agency Theory
- ▶ Business Model, the
- ▶ CEO Compensation
- ▶ Complementary Asset
- ▶ Co-specialization
- ▶ Decentralization
- ▶ Dynamic Capabilities
- ▶ Incentive Design
- ▶ Incomplete Contracts
- ▶ Make-or-Buy Decisions: Applications to Strategy Research
- ▶ M-Form Firms
- ▶ Opportunity Recognition
- ▶ Profiting from Innovation
- ▶ Resource-Based Theories
- ▶ Scope of the Enterprise
- ▶ Transaction Cost Economics
- ▶ Vertical Integration

References

Alchian, A.A., and H. Demsetz. 1972. Production, information costs, and economic organization. *American Economic Review* 62: 777–795.

Arrow, K.J. 1962. Economic welfare and the allocation of resources of invention. In *The rate and direction of inventive activity: Economic and social factors*. Princeton: Princeton University Press: NBER Report. Available at http://www.nber.org/chapters/c2144. Accessed 12 Feb 2012.

Barney, J.B. 1991. Firm resources and sustained competitive advantage. *Journal of Management* 17: 99–120.

Baumol, W.J. 1967. *Business behavior: Value and growth*. New York: Harcourt, Brace & World.

Berle, A.A., and G.C. Means. 1932. *The modern corporation and private property*. New York: Commerce Clearing House.

Boudreaux, D.J., and R.G. Holcombe. 1989. The coasian and knightian theories of the firm. *Managerial and Decision Economics* 10: 147–154.

Burns, T., and G.M. Stalker. 1961. *The management of innovation*. London: Tavistock.

Carpenter, M.A., and W.G. Sanders. 2002. Top management team compensation: The missing link between CEO pay and firm performance? *Strategic Management Journal* 23: 367–375.

Chesbrough, H., and D.J. Teece. 1996. Organizing for innovation: When is virtual virtuous? *Harvard Business Review* 74: 65–73.

Coase, R.H. 1937. The nature of the firm. *Economica* 4: 386–405.

Conner, K.R., and C.K. Prahalad. 1996. A resource-based theory of the firm: Knowledge versus opportunism. *Organization Science* 7: 477–501.

Di Stefano, G., M. Peteraf, and G. Verona. 2010. Dynamic capabilities deconstructed: A bibliographic investigation into the origins, development, and future directions of the research. *Industrial and Corporate Change* 19: 1187–1204.

Foss, K., N.J. Foss, P.G. Klein, and S.K. Klein. 2007. The entrepreneurial organization of heterogeneous capital. *Journal of Management Studies* 44: 1165–1186.

Gilbert, R.J., and D.M. Newbery. 1982. Preemptive patenting and the persistence of monopoly. *American Economic Review* 72: 514–526.

Grossman, S., and O. Hart. 1986. The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of Political Economy* 94: 691–619.

Holmstrom, B.R., and J. Tirole. 1989. The theory of the firm. In *Handbook of industrial organization*, vol. 1, ed. R. Schmalensee and R. Willig, 61–133. Amsterdam: North-Holland.

Jensen, M.C., and W.H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3: 305–360.

Kay, N.M. 2015. Coase and the contribution of 'the nature of the firm'. *Managerial and Decision Economics* 36: 44–54.

Knight, F. 1921. *Risk, uncertainty and profit*. New York: Augustus Kelley.

Kogut, B., and U. Zander. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* 3: 383–397.

Marris, R. 1964. *The economic theory of 'managerial' capitalism*. London: Macmillan.

Masten, S.E. 1988. A legal basis for the firm. *Journal of Law, Economics, and Organization* 4: 181–197.

Milgrom, P., and J. Roberts. 1988. An economic approach to influence activities in organizations. *American Journal of Sociology* 94: 154–179.

Milgrom, P., and J. Roberts. 1990. Bargaining costs, influence costs, and the organization of economic activity. In *Perspectives on positive political economy*, ed. J. Alt and K. Shepsle. Cambridge: Cambridge University Press.

Monteverde, K. 1995. Technical dialog as an incentive for vertical integration in the semiconductor industry. *Management Science* 41: 1624–1638.

Monteverde, K., and D.J. Teece. 1982. Supplier switching costs and vertical integration in the automobile industry. *Bell Journal of Economics* 13: 206–213.
Nonaka, I. 1991. The knowledge-creating company. *Harvard Business Review* 69: 96–104.
Penrose, E.T. 1959. *The theory of the growth of the firm*. London: Oxford University Press.
Prendergast, C. 1993. The role of promotion in inducing specific human capital acquisition. *Quarterly Journal of Economics* 108: 523–534.
Robinson, A. 1934. The problem of management and the size of firms. *Economic Journal* 44: 242–257.
Ross, S.A. 1973. The economic theory of agency: The principal's problem. *American Economic Review* 63: 134–139.
Rumelt, R.P. 1987. Theory, strategy, and entrepreneurship. In *The competitive challenge: Strategies for industrial innovation and renewal*, ed. D.J. Teece. Cambridge, MA: Ballinger.
Sautet, F.E. 2000. *An entrepreneurial theory of the firm*. London: Routledge.
Siegel, P.A., and D.C. Hambrick. 2005. Pay disparities within top management groups: Evidence of harmful effects on performance of high-technology firms. *Organization Science* 16: 259–274.
Simon, H. 1951. A formal theory of the employment relationship. *Econometrica* 19: 293–305.
Smith, A. 1776. *An inquiry into the nature and causes of the wealth of nations*. London: Strahan and Cadell. Available at: http://www.econlib.org/library/Smith/smWN.html.
Teece, D.J. 1976. *The multinational corporation and the resource cost of international technology transfer*. Cambridge, MA: Ballinger.
Teece, D.J. 1977. Technology transfer by multinational firms: The resource cost of transferring technological know-how. *Economic Journal* 87: 242–261.
Teece, D.J. 1980. Economies of scope and the scope of the enterprise. *Journal of Economic Behavior and Organization* 1: 223–247.
Teece, D.J. 1982. Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior and Organization* 3: 39–63.
Teece, D.J. 1986. Profiting from technological innovation. *Research Policy* 15: 285–305.
Teece, D.J. 2010. Technological innovation and the theory of the firm: The role of enterprise-level knowledge, complementarities, and (dynamic) capabilities. In *Handbook of the economics of innovation*, vol. 1, ed. N. Rosenberg and B. Hall. Amsterdam: North-Holland.
Teece, D.J. 2011. Human capital, capabilities and the firm: Literati, numerati, and entrepreneurs in the 21st-century enterprise. In *Oxford handbook of human capital*, ed. A. Burton-Jones and J.-C. Spender. Oxford: Oxford University Press.
Teece, D.J. 2012. Management and governance of the business enterprise: Agency, contracting, and capabilities perspectives. In *Oxford handbook of capitalism*, ed. D.-C. Mueller. Oxford: Oxford University Press.
Teece, D.J. 2014a. A dynamic capabilities-based entrepreneurial theory of the multinational enterprise. *Journal of International Business Studies* 45: 8–37.
Teece, D.J. 2014b. The foundations of enterprise performance: Dynamic and ordinary capabilities in an (economic) theory of firms. *Academy of Management Perspectives*. https://doi.org/10.5465/amp.2013.0116. Published online before print March 25, 2014.
Teece, D.J., G. Pisano, and A. Shuen. 1990. *Firm capabilities, resources, and the concept of strategy*, CCC Working paper No. 90-8. Berkeley: Center for Research in Management, University of California, Berkeley.
Teece, D.J., R. Rumelt, G. Dosi, and S. Winter. 1994. Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior and Organization* 23: 1–30.
Teece, D.J., G. Pisano, and A. Shuen. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18: 509–533.
Van den Steen, E. 2009. Authority versus persuasion. *American Economic Review: Papers and Proceedings* 99: 448–453.
Wang, C. 1997. Incentives, CEO compensation, and shareholder wealth in a dynamic agency model. *Journal of Economic Theory* 76: 72–105.
Williamson, O.E. 1964. *The economics of discretionary behavior: Managerial objectives in a theory of the firm*. Englewood Cliffs: Prentice Hall.
Williamson, O.E. 1975. *Markets and hierarchies: Analysis and antitrust implications*. New York: Free Press.
Williamson, O.E. 1981. The modern corporation: Origins, evolution, attributes. *Journal of Economic Literature* 19: 1537–1568.
Williamson, O.E. 1985. *The economic institutions of capitalism*. New York: Free Press.
Williamson, O.E. 1994. Visible and invisible governance. *American Economic Review* 84: 323–326.
Williamson, O.E., M.L. Wachter, and J.E. Harris. 1975. Understanding the employment relation: The analysis of idiosyncratic exchange. *Bell Journal of Economics* 6: 250–278.
Winter, S.G. 1988. On coase, competence, and the corporation. *Journal of Law, Economics, and Organization* 4: 163–180.
Yellen, J.L. 1984. Efficiency wage models of unemployment. *American Economic Review* 74: 200–205.

Theory X and Theory Y: HR Strategy

Richard P. Larrick[1] and Daniel C. Feiler[2]
[1]Duke University, Fuqua School of Business, Durham, NC, USA
[2]Dartmouth College, Hanover, NH, USA

Definition Douglas McGregor created the labels Theory X and Theory Y to capture two views of human motivation. The Theory X view assumes that employees must be monitored and controlled.

The Theory Y view assumes that employees work hard to accomplish important social and personal needs. McGregor argued that the assumptions managers make about motivation can become self-fulfilling.

In a 1957 article and 1960 book entitled *The Human Side of Enterprise*, Douglas M. McGregor described a basic tension in how managers and business scholars view the motivation of employees: Are they lazy, driven by money, needing to be tightly controlled and monitored? Or are they engaged, committed and interested in fulfilling themselves through work by contributing to their firm and society?

McGregor argued that a good deal of managerial practice was based on the former view, which he labeled as 'Theory X'. The assumptions of Theory X were that management is responsible for organizing and directing work, and 'without this active intervention by management, people would be passive – even resistant – to organizational needs' (1957: 23). He noted that less explicit assumptions tended to underlie Theory X (p. 23): 'the average man is by nature indolent . . . he lacks ambition, dislikes responsibility . . . and is inherently self-centered'. These assumptions lead managers to create rigid structures of evaluation, pay and control to manage 'indolent' workers. McGregor went on to argue, however, that money and job security are only the most basic needs. Drawing on earlier ideas developed by Abraham Maslow, McGregor argued that once basic needs have been fulfilled at work, employees crave to fulfil higher-order needs: to be accepted by others, to be independent and accomplish things, to be creative. This latter view he described as the Theory Y view. A central theme in his writing was that employees can often achieve higher levels of productivity when they are treated as responsible contributors to an organization rather than shirkers in need of prodding.

McGregor pointed to contemporary trends in management – decentralization and delegation, job enlargement and participation – as evidence that the Theory Y view had a growing presence in organizations, and his ideas anticipated a great deal of management theory and practice in the ensuing decades. Modern theories of job design (Hackman and Oldham 1976), intrinsic motivation (Deci and Ryan 1985) and organizational justice, especially procedural and interactional fairness (Lind and Tyler 1988), are heirs of this view. The recent interest in adding psychology to economics in the new subfield 'behavioural economics' can be viewed as a Theory Y correction of a Theory X view.

Management thinking tends to go through cycles, with the tenor of theories swinging from one emphasis to another (Abrahamson and Eisenman 2008), and Theory X and Theory Y capture a key dimension along which perspective shifts. This raises an interesting question of which perspective is true. Although subsequent researchers sometimes took one side or the other in their theorizing, a fundamental insight in McGregor's work was not a claim about the truth of each view (although he firmly believed in the assumptions of Theory Y), but that the unexamined assumptions of Theory X easily become self-fulfilling (Heath 1999). McGregor noted (McGregor 1957: 24) that 'human behavior in industrial organization today' corresponds to Theory X, but 'this behavior is not a consequence of man's inherent nature'. It is a consequence of 'management philosophy, policy, and practice', leading workers to behave exactly as predicted – with indolence and passivity. By acting on their pessimistic assumptions, managers evoke the behaviours they expect and arbitrarily confirm their initial pessimism. He concludes that 'it would seem that we are caught in a web of our own weaving' (p. 88). This theme is an enduring contribution of McGregor's work, and is reflected in a vibrant stream of current work showing the limitations of a purely Theory X perspective on employee behaviour (Ferraro et al. 2005; Markle 2011).

T

Because McGregor did not propose a specific theory of motivation but a summary of competing perspectives on motivation, his work did not generate directly testable hypotheses. Nevertheless, his proposal captured basic truths that will endure in organizations and will underpin future management research: employees are motivated by a range of interests; a focus on money and control ignores important motivations; and the assumptions that managers and scholars make about employee motivation can ultimately be self-reinforcing.

See Also

- ▶ General Management
- ▶ Human Resources
- ▶ Incentive Design
- ▶ Incentives
- ▶ Organizational Culture
- ▶ Organizational Learning

References

Abrahamson, E., and M. Eisenman. 2008. Employee-management techniques: Transient fads or trending fashions? *Administrative Science Quarterly* 53: 719–744.

Deci, E.L., and R.M. Ryan. 1985. *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Ferraro, F., J. Pfeffer, and R.I. Sutton. 2005. Economics language and assumptions: How theories can become self-fulfilling. *Academy of Management Review* 30: 8–24.

Hackman, J.R., and G.R. Oldham. 1976. Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance* 16: 250–279.

Heath, C. 1999. On the social psychology of agency relationships: Lay theories of motivation overemphasize extrinsic incentives. *Organizational Behavior and Human Decision Processes* 78: 25–62.

Lind, E.A., and T.R. Tyler. 1988. *The social psychology of procedural justice*. New York: Plenum.

Markle, A.B.. 2011. Dysfunctional learning in decision processes: The case of employee reciprocity. *Strategic Management Journal* 32: 1411–1425.

McGregor, D.M. 1957. The human side of enterprise. *The Management Review* 46: 22–28 and 88–92.

McGregor, D.M. 1960. *The human side of enterprise*. New York: McGraw-Hill.

Time-Based Competition

George Stalk
Boston Consulting Group, Toronto, ON, Canada

Abstract

Time-based competition is the use of speed to meet the needs of customers faster than can ▶ competitors and in ways that are difficult for competitors to match. Many executives agree that 'time is money', but few manage time as strictly as they manage money. Competitive time advantages can be used in strategies to raise prices, increase variety and quality, to re-segment the market, to gain market share and to grow profitably.

Definition Time-based competition is the use of speed to meet the needs of customers faster than can competitors and in ways that are difficult for competitors to match.

Time-based competition is the use of speed to meet the needs of customers faster than can ▶ competitors and in ways that are difficult for competitors to match (Stalk 1988). Speed can be used in the order-to-cash cycle, the innovation-to-cash cycle, the service-to-cash cycle and many more. Speed can also be used in capital investment programmes to build a business faster or alter ▶ business model, the more quickly than can competitors.

Time-based competition is not re-engineering. Rather, it is re-engineering for speed, not costs (although overall costs very often are reduced) as well as a strategy for using the speed advantages to raise prices, increase variety and quality, to re-segment the market, to gain market share and to grow profitably.

Key Elements of Time-Based Competition

A company's operations are subject to a challenging set of principles known as the rules of response. When managers appreciate the significance of these rules and use them to their benefit, they can change strategies and achieve startling gains.

The rules work as follows:

The Time Productivity Rule

Across a spectrum of businesses, the time required to execute a service or to order, manufacture and deliver a product is far less than the actual time the service or product spends in the company's value-delivery system.

For example, a manufacturer of heavy vehicles takes 45 days to prepare an order for assembly but only 16 h to create each product. The vehicle is actually being worked on less than 1% of the time it and its order spend in the operations system.

This rule highlights the poor time productivity of the majority of organizations. Most products and many services are actually receiving value for only 0.05–5% of the time they spend in the value-delivery systems of their companies.

The Time Lost Rule

During the 95–99.5% of the time a product is not receiving value, it is waiting. Time is lost waiting for three things:

- Completion of the batch the product is part of,
- Completion of the batch ahead of the batch the product is part of,
- Management getting around to making and executing the decision to send the batch on to the next step.

Generally, time lost is divided almost equally among these categories.

This lost time is affected very little by working harder, but working smarter has tremendous impact. Companies that reduce the size of the batches they process – whether physical goods or packets of information – and streamline the workflow will significantly reduce the time lost.

When a manufacturer of hospital equipment reduced standard lot sizes by 50%, the time required to manufacture the product declined by 65%. Then the company streamlined production flows to reduce material handling and lowered the number of intermediate steps that required scheduling. As a result, total time fell by another 65%.

The manufacturer's time productivity increased more than 200% – to 7% from 3%.

The Time Reward Rule

Companies that attack the consumption of time in their value-delivery system experience remarkable performance improvements. For every quartering of the time needed to provide a service or product, the productivity of labour and of working capital can often double. These gains result in as much as a 20% reduction in costs.

One US manufacturer of consumer durable products has reduced its time interval from 5 weeks to slightly more than a week. Labour and asset productivity has more than doubled. Costs are down considerably and profit is approaching extraordinary levels.

In addition, companies that cut time consumption turn the basis of competitive advantage to their favour. Growth rates of three times the industry average and profit margins of twice the usual levels can be achieved.

Take the example of a manufacturer of a prefinished building material, which reduced the time required to meet any and all customer orders to less than 10 days. Most orders can be on the customer's site 1–3 days from when they were placed. Meanwhile, competitors require 30–45 days to fill any order.

The fast-response competitor has grown more than 10% a year for the past 10 years, becoming the market leader. The average industry growth rate has been less than 3% over the same period. The pretax return on net assets of the fast-response competitor is 80% – more than double the average.

Companies that provide their customers with the value they want end up growing faster and are more profitable than their competitors. When a business can deliver value twice as quickly as its competitors, it will grow twice as fast as the market and it will be three times more profitable.

These rules of response apply to both service and product businesses. Managers who take the time to understand how the rules apply and aggressively use them to their advantage are:

- Growing much faster and more profitably than their competitors,
- Becoming closer and more essential to their customers,
- Taking leadership positions in their industries. Perhaps the greatest benefit of all is the excitement of new growth in core businesses.

Time-based competition is discussed in great detail in Stalk (1990) and Blackburn (1990).

Reflections on Time-Based Competition

Some believe that Western business has moved beyond time as a source of competitive advantage to other, more important, sources. In part this is true, but time is an even more important performance metric today than it was in the late 1980s. Product life-cycles are continuously shrinking and products and business models are morphing ever faster – witness cell, smart, tablets and more (see Stalk 2006).

On the grandest scale is global logistics. Because of the rise of Asia, especially China, as the centre of global manufacturing, supply chains are lengthening in distance and in time and are becoming more complex. If one also considers increasing congestion and delays, it is clear that the volatility of all aspects of the supply chain is worsening. The principles of time-based competition will enable some companies to gain an edge over their slower competitors (see Mercier and Stalk 2011).

The founder of the Boston Consulting Group, Bruce Henderson (1978, personal communication), said: ‘You don’t have to be the best. You just have to be better than your competitors.’ Having a fast and reliable supply chain when competitors are struggling is a huge advantage. Companies in this enviable position can:

- Commit to quick-turnaround orders,
- Be more flexible when customers want to change their order mix and volume,
- Increase variety, freshness and turnover of store offerings,
- Gain leverage with key suppliers and ensure preferential treatment,
- Squeeze competitors by lowering prices – or respond to changes in input costs more quickly,
- Give customers better terms and service to strengthen loyalty and increase share of wallet.

See Also

► Business Model, the
► Business Strategy
► Competitiveness
► Competitors
► Dynamic Capabilities

References

Blackburn, J. 1990. *Time-based competition: The next battleground in American manufacturing*. Homewood: Irwin.

Mercier, P., and G. Stalk. 2011. *The coming global gridlock: Is your supply chain ready?* Boston: Boston Consulting Group.

Stalk, G. 1988. Time – the next source of competitive advantage. *Harvard Business Review* 66: 41–51.

Stalk, G. 1990. *Competing against time*. New York: Free Press.

Stalk, G. 2006. Hardball innovation. *Research Technology Management* 49: 20–29.

Top Management Teams

Sydney Finkelstein
Dartmouth College, Tuck School of Business, Hanover, NH, USA

Abstract

Research on top management teams (TMTs) has become a central feature of work in strategic leadership in particular, and strategic management in general. The TMT unit of analysis allows for a comprehensive assessment of how senior executives interact, as well as when and how they influence key organizational outcomes. Studies of TMTs can provide more insight into key organizational outcomes than investigations of CEOs alone. Research on TMTs pays special attention to strategic decisions, which emanate from a group of top managers interacting as social and political creatures, and the resulting strategic choices and firm performance that ensues.

Definition A top management team (TMT) is the collection of senior executives at the strategic apex of an organization, with primary responsibility for a firm’s strategic direction and performance. As such, TMTs are a key unit of analysis in studies of organizational life.

Why Should Anyone Study Top Management Teams?

Scholars have been drawn to the study of top management teams (TMTs) for five main reasons. First, as an aggregation of subunits and individuals, organizations have multiple goals that are often in conflict (Cyert and March 1963). The existence of these multiple goals, and hence of multiple preferences, at the top of organizational hierarchies is likely to affect how organizations strive towards organizational outcomes, as well as the characteristics of those outcomes. Second, almost all descriptions of ▶ strategic decision-making processes typically emphasize the relevance of stages, sequences and processes that involve a group of top managers interacting towards desired ends (Mintzberg et al. 1976). Indeed, the TMT is at the strategic apex of an organization (Mintzberg 1979); it is the executive body most responsible for ▶ strategic decision-making and, by extension, for such fundamental organizational outcomes as firm strategy, structure and performance. Third, the interactions among top managers, including power distributions, decision processes, integration and fragmentation, create outcomes of interest to strategy research.

Fourth, there is clearly some amount of role differentiation in most, if not all, top management groups. For example, in the United States, Sarbanes–Oxley legislation requires that a chief financial officer, along with the CEO, personally certify the accuracy of accounting statements prior to sending them to shareholders and filing them with the Securities and Exchange Commission. Thus, some specific responsibilities of executives other than the CEO have been legally mandated for public companies.

Fifth and most important, evidence suggests that studying TMTs, rather than CEOs alone, provides better predictions of organizational outcomes (Ancona 1990). For example, in a series of tests of upper echelon hypotheses, Finkelstein (1988) reported far stronger results from using the TMT, rather than the CEO, as the level of analysis. Other studies have similarly demonstrated that significant variance in organization-level outcomes can be explained by examining the attributes of executives beyond the CEO (Zhang and Rajagopalan 2004).

For these reasons, whether one refers to such groups as dominant coalitions (Cyert and March 1963), ‘inner circles’ (Finkelstein 1992), top management groups (Hambrick 1994) or top management teams (Carpenter et al. 2004), there is much to gain from focusing on the relatively small constellation of executives at the top of an organization.

What Is in the DNA of Top Management Teams?

Although the term ‘top management team’ is now widely used, it is not uncommon for individual pieces of research to emphasize different aspects of what is, in essence, a multi-dimensional construct. A top management team has three central conceptual elements: composition, structure and process. Composition refers to the collective characteristics of top team members, such as their values, cognitive bases, personalities and experiences. Although these characteristics can be considered in terms of both the central tendency of the team and the heterogeneity of the team, most researchers have focused on the latter. In addition, TMT heterogeneity encompasses both psychological factors (values, beliefs, cognitions) and aspects of executive experiences (age, tenure, functional background, education).

The structure of a top team is defined by the roles of members and the relationships among those roles. Central to this definition is the role interdependence of team members, an important construct that surely has significant consequences for how strategic decisions are made (Hambrick 1994). We define role interdependence as the degree to which the performance of the firm depends on information- and resource-sharing, as well as other forms of coordination within the TMT. For example, a TMT consisting of heads of functional areas typically has more role interdependence than one made up of heads of autonomous business units.

The third major conceptual element of a TMT is its processes. By processes, we mean the nature of interaction among top managers as they engage in strategic decision-making. The two most often studied process dimensions are social integration and consensus. Social integration is defined as 'the attraction to the group, satisfaction with other members of the group, and social interaction among the group members' (O'Reilly et al. 1989: 22) and is one of the most studied of process constructs. Consensus within a TMT is 'the [extent of] agreement of all parties of a group decision' (Dess and Origer 1987: 313).

All three conceptual elements – composition, structure and process – are related to the social make-up and interactions of the top team in the process of making strategic decisions. Strategic decisions are not made in a vacuum; rather, they emanate from a group of top managers interacting as social and political creatures. The nature of these interactions and their effects on both strategic decision-making and organizational outcomes are of central importance.

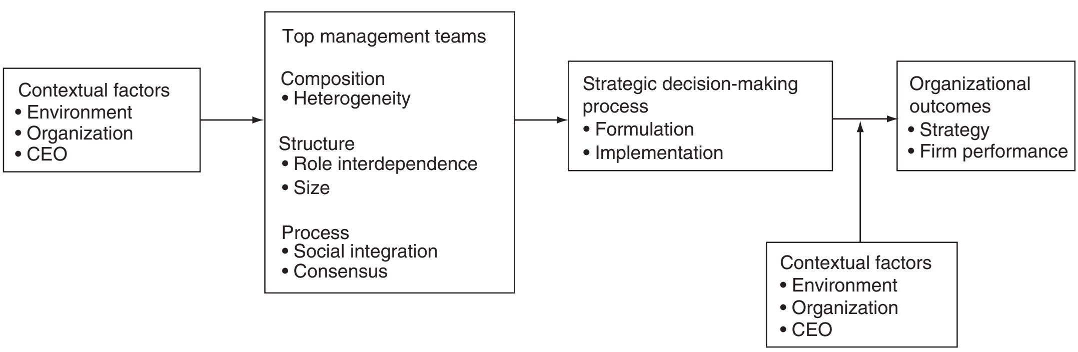
We believe these issues can be best understood by adopting the framework shown in Fig. 1. At the centre of this framework is the TMT, characterized in terms of a set of conceptual constructs: heterogeneity (TMT composition), role interdependence and team size (TMT structure), and social integration and consensus (TMT process). We focus on these constructs, in particular, because they are central to both strategic decision-making and social relations within TMTs, and they have been the subject of considerable theoretical interest among scholars for some time.

The framework suggests how each of these facets of TMTs is interrelated. The model also encompasses the effects of contextual conditions on TMTs. These contextual factors include the environment, the organization and the CEO. Finally, Fig. 1 shows how TMTs are associated with the strategic decision-making process and the organizational outcomes that arise from this process.

How Do TMTs Affect Firm Strategy?

There have been many studies over the years on this very question. As we will see, however, the jury is still out on just how much progress has been made. Most of these studies have focused on the effects of TMTs on organizational innovativeness, the interdependence of diversification posture, and strategic change.

We will highlight two classic studies on the associations among demographic heterogeneity, team size and organizational innovation. Arguing that demographic heterogeneity proxies for cognitive heterogeneity within a TMT, Bantel and Jackson (1989) found that functional heterogeneity was positively associated with administrative innovation in a sample of 199 banks in a Midwestern US state. But the heterogeneity of team members along other demographic dimensions,



Top Management Teams, Fig. 1 A model of top management teams (From Finkelstein et al. (2009))

such as age, tenure and educational specialization, did not significantly predict administrative innovations.

In another study on innovation, O'Reilly and Flatt (1989) used multiple measures of organization innovation, as well as both age and tenure heterogeneity of TMTs, to test related hypotheses. Of the eight different models tested, three yielded negative and significant results, indicating that homogeneous TMTs were more innovative. No consistent pattern of results arose across the two studies. Of the 16 different models tested across articles, 1 indicated that heterogeneity was a positive predictor of innovation and 3 suggested the opposite.

There have been dozens of other studies focused on the direct relationship between elements of a TMT and firm strategy, but the hallmark of this work is, alas, rather inconsistent results. Why? One explanation may be that attempts to relate TMT heterogeneity and strategic choices directly are assuming a connection that is more distant than commonly recognized. Demographic heterogeneity is associated with cognitive heterogeneity, both of which increase the number of strategic alternatives considered by a TMT and the quality of the evaluation of those alternatives. Rigorous strategy formulation, in turn, is expected to lead to better informed decisions. Using this logic to predict strategic outcomes is subject to three potential drawbacks.

First, and as is the case for predictions based on the central tendencies of TMTs, there are several logical stages between TMT composition and strategic choice that can disrupt or attenuate expected associations (Cannella and Holcomb 2005). For example, the strategic decision-making process is complex and ambiguous, numerous contextual conditions can affect the process through which strategic choices are selected and implemented, and many of these same contextual factors are often direct determinants of strategic choices as well. Hence, while TMTs undoubtedly affect strategic outcomes, our ability to empirically detect this relationship may be limited.

Second, the logical sequence we outlined above does not link TMT heterogeneity to strategic choices as much as it relates heterogeneity to the quality of strategic decisions. There is a big difference between predicting rigorous strategy formulation and predicting specific strategic outcomes, which suggests that measures of cognitive heterogeneity should not be any better predictors of strategy, since heterogeneity, whether measured demographically or cognitively, is potentially far removed from specific strategic outcomes.

Finally, there is a point that is seldom noted in the literature but which may be quite telling. Logically, a significant difference exists between how TMT heterogeneity and TMT average tendencies are expected to affect strategy. Because the extent to which a TMT is characterized by a particular compositional attribute defines its orientation or preference set (Finkelstein 1988), this attribute can more easily be translated into specific strategic outcomes than is true for TMT heterogeneity. For example, TMTs dominated by executives with sales and marketing experience will perceive and interpret information in such a way that they will be more likely to prefer such strategies as product innovation and differentiation (Hambrick and Mason 1984). In contrast, and as we have seen, TMT heterogeneity affects the *process* of making strategic decisions much more than it does the *content* of those strategies. Hence, we should not necessarily expect heterogeneity to have a direct impact on strategy content.

The findings we summarize here suggest that direct relationships between TMT heterogeneity and strategic choices are unlikely to be robust. Rather, it may be that TMT heterogeneity and social integration interact during strategic decision-making, potentially affecting how the formulation and implementation processes come out. As a result, it seems important to study the relationships among TMT heterogeneity, social integration and strategic decision-making as a first step before attempting to predict strategic outcomes.

How Do TMTs Affect Firm Performance?

Given some of the problems in empirically establishing linkages between TMT interaction processes and strategic choices, it would not be surprising if studies of the association between the distributional properties of TMTs and firm performance were even more problematic. To some extent, this is reflected in the often inconsistent findings that emerge from this work. In contrast to studies of strategic outcomes, however, some research predicting firm performance has also incorporated contingency factors, such as industry change or turbulence, that have the potential to strengthen results.

For example, the positive effects of TMT heterogeneity on firm performance in 'high velocity' or turbulent environments in several studies (e.g., Kor 2003) may help point the way to a clearer understanding of what heterogeneity among top managers really means. TMT heterogeneity promotes a more rigorous strategy formulation process by increasing the number of feasible strategic alternatives under consideration and the quality of their evaluation. In fast-changing, dynamic environments, managerial work becomes more fragmented (Mintzberg 1973), information-processing requirements increase (Hambrick et al. 2005), and new opportunities and crises necessitate greater adaptive capabilities (Galbraith 1973) – all of which place a higher premium on the generation of multiple and novel solutions. It is precisely in the most unstable environments that TMT heterogeneity is most valuable.

In contrast, we might expect strategy implementation to be more salient than strategy formulation in stable environments because the strategic challenge is less in developing new ideas than it is in preserving established procedures (Tushman and Romanelli 1985). TMT cooperation and stability become more important when environments are more stable (Nutt 1987), suggesting that integrated TMTs may be preferred. Hence, in stable environments, TMT social integration, rather than heterogeneity, may be related to firm performance.

Beyond the moderating role of organizational environments, other contingency factors may help explain how and when TMT heterogeneity affects firm performance. In fact, the contextual conditions that give rise to different configurations of TMTs may themselves often operate as moderating forces on firm performance. For example, because unstable environments impose demands on how organizations should structure their TMTs, firms promoting TMT demographic heterogeneity in unstable environments should do better. More generally, those TMTs that are organized in line with environmental contingencies are likely to perform better.

Research on TMTs has held great promise for years, but only some of this potential has been fulfilled. This work needs more complex frameworks of TMTs that recognize the role of senior executives in strategic decision-making, along with the moderating role of such important contextual influences as the environment, the organization and the CEO. In addition, much more work is needed on such basic aspects of TMTs as their boundaries and determinants. This stream of work may no longer be new, but as it enters its adolescence there remain many key questions to be examined.

See Also

▶ CEO Compensation
▶ Managerial Discretion
▶ Strategic Decision-Making
▶ Succession Management
▶ Upper Echelons Theory

References

Ancona, D.G. 1990. Top management teams: Preparing for the revolution. In *Applied social psychology and organizational settings*, ed. J.S. Carroll. Hillsdale: Erlbaum.

Bantel, K.A., and S.E. Jackson. 1989. Top management and innovations in banking: Does the composition of the top team make a difference? *Strategic Management Journal* 10: 107–124.

Cannella Jr., A.A., and T.R. Holcomb. 2005. A multilevel analysis of the upper-echelons model. In *Research in multi-level issues*, ed. A. Dansereau and F.J. Yammarino. Oxford: Elsevier.

Carpenter, M.A., M.A. Geletkanycz, and W.G. Sanders. 2004. The upper echelons revisited: Antecedents, elements, and consequences of top management team composition. *Journal of Management* 60: 749–778.

Cyert, R.M., and J.G. March. 1963. *A behavioral theory of the firm*. Englewood Cliffs: Prentice Hall.
Dess, G.G., and N.K. Origer. 1987. Environment, structure, and consensus in strategy formulation: A conceptual integration. *Strategic Management Journal* 8: 313–330.
Finkelstein, S. 1988. Managerial orientations and organizational outcomes: The moderating roles of managerial discretion and power. Unpublished Ph.D. dissertation, Columbia University, New York.
Finkelstein, S. 1992. Power in top management teams: Dimensions, measurement, and validation. *Academy of Management Journal* 35: 505–538.
Finkelstein, S., D.C. Hambrick, and A. Cannella. 2009. *Strategic leadership: Theory and research on executives, top management teams, and boards*. Oxford: Oxford University Press.
Galbraith, J.R. 1973. *Designing complex organizations*. Reading: Addison-Wesley.
Hambrick, D.C. 1994. Top management groups: A conceptual integration and reconsideration of the 'team' label. In *Research in organizational behavior*, ed. B.M. Staw and L.L. Cummings. Greenwich: JAI Press.
Hambrick, D.C., and P. Mason. 1984. Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review* 9: 193–206.
Hambrick, D.C., S. Finkelstein, and A.C. Mooney. 2005. Executive job demands: New insights for explaining strategic decisions and leader behaviors. *Academy of Management Review* 30: 472–491.
Kor, Y.Y. 2003. Experience-based top management team competence and sustained growth. *Organization Science* 14: 707–719.
Mintzberg, H. 1973. *The nature of managerial work*. New York: Harper & Row.
Mintzberg, H. 1979. *The structuring of organizations: The synthesis of the research*. Englewood Cliffs: Prentice Hall.
Mintzberg, H., D. Raisinghani, and A. Theoret. 1976. The structure of unstructured decision processes. *Administrative Science Quarterly* 21: 246–275.
Nutt, P.C. 1987. Identifying and appraising how managers install strategy. *Strategic Management Journal* 8: 1–14.
O'Reilly III, C. A., and S. Flatt. 1989. *Executive team demography, organizational innovation and firm performance*. Working paper, University of California, Berkeley.
O'Reilly III, C.A., D.F. Caldwell, and W.P. Barnett. 1989. Work group demography, social integration, and turnover. *Administrative Science Quarterly* 34: 21–37.
Tushman, M.L., and E. Romanelli. 1985. Organizational evolution: A metamorphosis model of convergence and reorientation. In *Research in organizational behavior*, ed. L.L. Cummings and B.M. Staw. Greenwich: JAI Press.
Zhang, Y., and N. Rajagopalan. 2004. When the known devil is better than an unknown God: An empirical study of the antecedents and consequences of relay CEO succession. *Academy of Management Journal* 47: 483–500.

Trade Creation/Diversion

Elitsa R. Banalieva
Northeastern College of Business Administration, Boston, MA, USA

Definition Trade creation occurs when tariff reduction within a regional trade agreement (RTA) displaces a member's higher-cost production with lower-cost imports from another member (Viner 1950). Trade diversion occurs when a member's lower-cost imports from non-members are displaced by higher-cost imports from another member (Viner 1950).

Regional trade agreements (RTAs) have been gaining in importance: the number of notified RTAs has been rising since the 1990s, reaching 575 as of 31 July 2013 (World Trade Organization 2013). Additionally, '[m]ore than half of world trade' occurs within RTAs (Cipollina and Salvatici 2010: 63). The literature suggests that the reduction in trade barriers in an RTA tends to be welfare-improving. The access to larger markets and lower-cost suppliers in the RTA leads to greater economies of scale. Moreover, the increased competition in the RTA induces firms to become efficient at serving customers.

Research has used different methodologies to detect RTAs' net effect on trade (Grimwade et al. 2011) but the most common approach is the gravity model. Its simplest form explains trade between two countries as a function of their GDPs (larger economies trade more) and distance (a proxy for transaction costs), but this simple model has evolved to include additional variables such as land-locked location, contiguity, colonial ties and psychic distance, among other factors (Gnosh and Yamarik 2004).

The gravity model captures trade-creating and diverting RTA effects with three variables: *Intra-RTA Trade* (D1) equals one if both countries belong to the same RTA, zero otherwise; *RTA Imports from Non-Members* (D2) equals one if the importing country is an RTA member and the exporting one is not, zero otherwise; and *RTA Exports to Non--Members* (D3) equals one if the exporting country is an RTA member and the importing one is not, zero otherwise (Carrère 2006). If the coefficients of D1 $>$ 0 and D2 $\geq$ 0 (D3 $\geq$ 0), there is pure trade creation in imports (exports). If the coefficients of D1 $>$ 0 and D2 $<$ 0 (D3 $<$ 0), there is trade diversion in imports (exports) (Carrère 2006). A recent meta-analysis found that '*ex post* empirical estimates of an influence of RTAs on trade flows are positive', with an average of 40% increase in trade (Cipollina and Salvatici 2010: 78).

Most of the RTA analysis has been at country and industry levels: '[s]trategic management scholars have . . . traditionally stayed away from research topics that require . . . interest . . . in trade theory' (Brahm 1995: 87). The dearth of RTA research at the firmlevel has shown that firms increase their regional investment, upgrade their technological capabilities, improve their technical efficiency from an increasing home region focus, and adopt regional staffing policies due to deeper regional integration (Fratianni et al. 2011; Banalieva et al. 2012).

A useful way to extend this research is by separating the effects of degree of trade creation (the current focus of the literature) from speed of trade creation (the number of years it takes an RTA to become trade-creating since its formation) and rhythm of trade creation (the variation in the rate of trade creation over time). Research can use continuous measures such as, for example, the intraregional trade intensity index to capture the effect of speed and rhythm of trade creation on firms' strategies and performance. Faster or more sudden trade creation in an RTA imposes accelerating adaptation demands on RTA-based firms, so smaller firms likely face greater adaptation constraints. Consequently, the moderating roles of slack (high- and low-discretion), marketing and innovation resources on firms' relationship between the speed or rhythm of trade-creation and performance could also be further analysed.

See Also

- ▶ Industrial Policy
- ▶ International Political Economy
- ▶ Regional Development
- ▶ Transaction Cost Economics

References

Banalieva, E., M. Santoro, and J. Jiang. 2012. Home region focus and technical efficiency of multinational enterprises: The moderating role of regional integration. *Management International Review* 52: 493–518.

Brahm, R. 1995. National targeting policies, high technology industries, and excessive competition. *Strategic Management Journal* 16: 71–91.

Carrère, C. 2006. Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model. *European Economics Review* 50: 223–247.

Cipollina, M., and L. Salvatici. 2010. Reciprocal trade agreements in gravity models: A meta analysis. *Review of International Economics* 18: 63–80.

Fratianni, M., F. Marchionne, and C. Oh. 2011. A commentary on the gravity equation in international business research. *Multinational Business Review* 19: 36–46.

Gnosh, S., and S. Yamarik. 2004. Are regional trade arrangements trade creating? An application of extreme bounds analysis. *Journal of International Economics* 63: 369–395.

Grimwade, N., D. Mayes, and J. Wang. 2011. Estimating the effect of integration. In *International handbook on the economics of integration*, vol. 3, ed. M.N. Jovanovic. Cheltenham: Edward Elgar.

Viner, J. 1950. *The customs union issue*. New York: Carnegie Endowment.

World Trade Organization. 2013. Regional trade agreements. Available at http://www.wto.org/english/tratop_e/region_e/region_e.htm. Accessed 26 Nov 2013.

Trade Dress

Edward F. Sherry
Berkeley Research Group, Emeryville, CA, USA

Definition 'Trade dress' is a legal term referring to the overall visual look and feel of a product line, packaging or business.

'Trade dress' is a legal term referring to the overall visual look and feel of a product line, packaging or business. In addition to a product's physical appearance – encompassing factors such as size, shape, design, colour and texture – trade dress may also refer to the manner in which a product is packaged, labelled, presented or advertised, including the use of distinctive graphics and marketing strategies.

Trade dress is generally protectable as a form of intellectual property akin to, but distinct from, ► trademark (which protects a name or logo used in commerce). In the US, trade dress is protected at the federal level under Section 43(a) of the Lanham Act (15 U.S.C. § 1051 ff). Trade dress can be registered with the US Patent and Trademark Office (whether on the Primary Register or Secondary Register), although it can also be protected, even if not registered. Trade dress is protected at the state level under similar state statutes, under general unfair competition statutes and under certain common law doctrines (such as the law of 'passing off').

Trade dress law, its relationship to other laws (e.g., consumer protection law and patent law), the extent to which trade dress is legally protectable, the nature of that protection, and the remedies available for trade dress infringement all vary by country. Information about trade dress laws in various countries is available in *Trade Dress: International Practice and Procedures*, a searchable database of country profiles on trade dress protection and enforcement published by the International Trademark Association.

To qualify for trade dress protection, the trade dress elements in question must be 'non-functional' and must be 'inherently distinctive', so that the trade dress has acquired 'secondary meaning'. US courts have identified eight factors to be assessed in determining whether there has been misappropriation of another's trade dress:

1. The strength of the plaintiff's mark;
2. The relatedness of the goods;
3. The similarity of the marks;
4. Evidence of actual confusion;
5. The marketing channels used;
6. The likely degree of purchaser care;
7. The defendant's intent in selecting the mark; and
8. The likelihood of expansion of the product lines.

The most commonly stated rationale for trade dress protection is to prevent or reduce *consumer* confusion as to the source or quality of the product in question. Traditionally, an 'ordinary buyer' test is used to determine whether there is the requisite likelihood of confusion.

But trade dress protection serves other purposes as well: encouraging customers to rely on distinctive trade dress when evaluating the quality of merchandise; protecting the economic investments made by firms in distinguishing their products from those of competitors; protecting the good will and reputation that may be associated with the trade dress of particular products or companies; preventing free-riding on the investments of others; and encouraging competition by encouraging firms to associate their own trade dress with the nature and quality of the products they sell.

From a strategic management perspective, developing a distinctive form of trade dress provides a useful and legally protectable mechanism to differentiate one's product from competitors' products and to cement the product's image in consumers' eyes. However, because the bounds of what is and what is not protected by trade dress law are often uncertain – how similar does a competitor's product's 'look and feel' have to be before it is deemed to infringe the plaintiff's trade dress? – the degree to which trade dress law provides effective competitive protection is also uncertain.

See Also

► Patents
► Trademark

Trade Secret

Liang Guo
Department of Marketing, The Econometric Society, Chinese University of Hong Kong, Kowloon, Hong Kong, China

Abstract

This entry provides a summary of issues on trade secret both as a legal concept and as a research subject. First, it discusses the general standards that are commonly used in the

definition of trade secret and the precautionary measures and the legal system for trade secrecy. It also compares the advantages and disadvantages of the legal protection of trade secrets with that of ▶ patents, while pointing out the supplemental role of trade secret. Next, two streams of academic literature are reviewed, on firms' strategic incentives for disclosing confidential information to competitors and to channel partners respectively. The issues examined in the economics of trade secret law are also briefly addressed.

Definition Trade secret consists of any proprietary knowledge or information which is not publicly known or reasonably ascertainable but which is valuable to the secret holder.

The subject matter of trade secret encompasses advertising strategies, client lists, consumer profiles, designs, formulas, manufacturing processes and sales methods. Trade secrets are consistently regarded by most European and American technology managers as more important than ▶ patents in the appropriation of returns from technical ▶ innovation (Arundel and Kabla 1998; Cohen et al. 2000; Arundel 2001; Png 2011). Although the precise language and interpretation varies across legal systems, there are three general standards that are common to the definition of trade secret:

(a) The information must be kept confidential, that is, is not generally known, or readily accessible, to the public.
(b) The information must confer some economic value on its holder, and the value must be derived specifically from its not being generally known (i.e., not just from the value of the information itself).
(c) The owner must make reasonable efforts to maintain the confidentiality of the secret in question.

Owners of a trade secret can take a variety of (civil and commercial) precautionary measures to preserve confidentiality, for example, non-disclosure agreements within employee contracts and non-compete clauses in vendor supply or licence arrangements. Violators of these agreements may be liable for financial penalties.

The protection of trade secrets from being misappropriated is an important legal aspect in many societies to safeguard overall economic vitality. Misappropriators of trade secrets can incur legal liability, subject to the holder's obligation to take reasonable steps to maintain its secrecy. In the United States, trade secrets are under state jurisdiction, and were historically protected by common law (Png 2011). About 46 states have adopted the Uniform Trade Secrets Act (UTSA) as the basis for trade secret law. Another relevant legislation is the Economic Espionage Act of 1996, under which the theft or misappropriation of trade secrets constitutes a federal crime.

The advantages conferred by the legal protection of trade secrets over that of patents include exemption from any procedural formalities. Consequently, trade secret protection involves no registration costs, can have immediate effect, and does not lead to the disclosure of the confidential information. This effectively permits perpetual monopoly status and unlimited protection as long as appropriate efforts are made to maintain secrecy and the secret has not been made known to the public, whereas patents normally specify a limited period of protection (usually 20 years). Moreover, trade secret protection has a broader scope and lower qualification criteria, covering works in progress and extending beyond technical innovations to business ideas, marketing concepts, customer lists and so on.

The downside of trade secret is that the protection is generally weaker than that for patents. It does not provide a minimum protection period, nor does it provide protection against accidental disclosure, independent discovery or reverse engineering. In addition, the enforceability of trade secret protection is relatively low and the administration costs are likely high. The costs of confidentiality maintenance for the holder of a trade secret can also be potentially high. Moreover, there may be a risk of legitimate patenting of the same invention by a third party.

Trade secret laws usually supplement patent laws (Friedman et al. 1991). The holder of a patentable invention can choose whether to resort to trade secret in order to appropriate the returns from its invention, by gauging the overall advantages and disadvantages relative to those of patenting. For example, trade secrecy can be chosen, when the estimated likelihood is relatively low that the information can be legitimately acquired by a third party within the period of patent protection, or when the value of the invention is believed to be too small relative to the registration cost of patenting. Nevertheless, trade secrecy can complement patent protection, especially during the period prior to the formal grant of the patent, or for ongoing/additional innovation that is not fully covered or disclosed within patent registration.

There is a substantial academic literature on the benefits a firm can derive from sharing confidential demand or cost information with competitors and thus from influencing their behaviours to its own advantage (e.g., Novshek and Sonnenschein 1982; Gal-Or 1986; Shapiro 1986; Villas-Boas 1994). Vives (1984) and Raith (1996) show that the equilibrium incentives of competing firms to share confidential information are affected by the nature of the product (substitutes or complements) and the nature of competition (Cournot or Bertrand). Sheppard (1996) points out that a firm can benefit from licensing its proprietary technology to competitors, because the increasing competition can allow the firms to make a credible quality commitment and thus enhance industry demand.

There is another stream of academic studies in operation management and in marketing that examine how information-sharing can improve the efficiency of a supply chain by reducing production, logistical or inventory-related costs (e.g., Cachon and Fisher 2000; Kulp et al. 2004; Iyer et al. 2007; He et al. 2008). This literature highlights the strategic incentives to exchange confidential information in order to mitigate distorted behaviours from other channel partners (i.e., vendors, retailers). Recent developments in this literature investigate how these strategic incentives are influenced by the channel structure (e.g., Li 2002; Gal-Or et al. 2008), and by whether the sharing mechanism is mandatory or voluntary (Guo 2009; Guo and Iyer 2010).

The economics of trade secret law is primarily concerned about the legal arrangements under which the private sector's incentives for *ex ante* investments on innovation and for *ex post* sharing are compatible with the socially optimal outcome. In comparison to patent law, although the protection of trade secret may lead to excessive duplication of *ex post* inventive efforts, it may discourage too many *ex ante* investments due to its relatively weaker protection (Friedman et al. 1991). Another question is why certain conduct to unmask competitors' trade secrets (e.g., accidental disclosure, independent discovery and reverse engineering) are permitted while others (e.g., industrial espionage, breach of agreement, bribery) are prohibited. Relative to the unlawful conducts, the lawful ones are more costly to prevent, involve greater social benefits of information-sharing, and/or can induce less defensive (and socially suboptimal) efforts from the trade secret's owners.

See Also

▶ Competitive Advantage
▶ Information and Knowledge
▶ Innovation
▶ Patents

References

Arundel, A. 2001. The relative effectiveness of patents and secrecy for appropriation. *Research Policy* 30: 611–624.

Arundel, A., and I. Kabla. 1998. What percentage of innovations are patented? Empirical estimates for European firms. *Research Policy* 27: 127–141.

Cachon, G.P., and M. Fisher. 2000. Supply chain inventory management and the value of shared information. *Management Science* 46: 1032–1048.

Cohen, W.M., R.R. Nelson, and J.P. Walsh. 2000. *Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not)*, NBER Working paper, vol. 7552. Cambridge, MA: NBER.

Friedman, D.D., W.M. Landes, and R.A. Posner. 1991. Some economics of trade secret law. *Journal of Economic Perspectives* 5: 61–72.

Gal-Or, E. 1986. Information transmission: Cournot and Bertrand equilibria. *Review of Economic Studies* 53: 85–92.

T

Gal-Or, E., T. Geylani, and A. Dukes. 2008. Information sharing in a channel with partially informed retailers. *Marketing Science* 27: 642–658.
Guo, L. 2009. The benefits of downstream information acquisition. *Marketing Science* 28: 457–471.
Guo, L., and G. Iyer. 2010. Information acquisition and sharing in a vertical relationship. *Marketing Science* 29: 484–506.
He, C., J. Marklund, and T. Vossen. 2008. Vertical information sharing in a volatile market. *Marketing Science* 27: 513–530.
Iyer, G., C. Narasimhan, and N. Niraj. 2007. Information and inventory in distribution channels. *Management Science* 53: 1551–1561.
Kulp, S.C., H.L. Lee, and E. Ofek. 2004. Manufacturer benefits from information integration with retail customers. *Management Science* 50: 431–444.
Li, L. 2002. Information sharing in a supply chain with horizontal competition. *Management Science* 48: 1196–1212.
Novshek, W., and H. Sonnenschein. 1982. Fulfilled expectations Cournot duopoly with information acquisition and release. *Bell Journal of Economics* 13: 214–218.
Png, I. P. L. 2011. Law and innovation: Evidence from the uniform trade secrets act. Working paper, National University of Singapore.
Raith, M.A. 1996. A general model of information sharing in oligopoly. *Journal of Economic Theory* 71: 260–288.
Shapiro, C. 1986. Exchange of cost information in oligopoly. *Review of Economic Studies* 53: 433–446.
Shepard, A. 1996. Licensing to enhance demand for new technologies. *RAND Journal of Economics* 18: 360–368.
Villas-Boas, J.M. 1994. Sleeping with the enemy: Should competitors share the same advertising agency? *Marketing Science* 13: 190–202.
Vives, X. 1984. Duopoly information equilibrium: Cournot and Bertrand. *Journal of Economic Theory* 34: 71–94.

Trademark

David Soberman
University of Toronto, Rotman School of Management, Toronto, ON, Canada

Abstract

In this article, we start by discussing how trademarks are used around the world. We then provide a rationale for the existence of trademarks (how they provide value) and continue by examining the relationship between trademarks and brands. Finally, we conclude by discussing how firms protect their trademarks and the types of trademark misuse that are of significant concern to trademark owners.

Definition A trademark (which can also be written as trade mark) is a word, symbol or emblem that a firm uses to identify itself (or its products) within a category as being distinct (and different) from competitors. Typically, the names of products are trademarks, for example, Coke or BlackBerry, but symbols or emblems such as the Nike swoosh have the same protection (Liodice 2010).

Global Use of Trademarks

Today in most democratic countries, trademarks are registered with the government; however, this is not strictly required, at least in countries governed by Common Law (countries governed by Common Law, which has its origins in medieval England, include most members of the Commonwealth of Nations and the United States). In particular, there is a tradition of recognizing the ownership of a word, symbol or emblem by a firm given that the firm is (and has been) active in using and trading products or services identified by the trademark in question. Many countries also belong to a convention which accords trademark privileges to a company that intends to launch an existing product or service into a new country. Trademarks sometimes become identified with a category as is the case with products such as Aspirin and Kleenex. Firms even engage in activity to prevent this (Carlton and Perloff 2000). In contrast to patents and copyright, registered trademarks are recognized indefinitely though the registration must be renewed on a regular basis (Smyth et al. 2010).

Value of Trademarks

The value of trademarks is that they facilitate the identification of products and services as coming from specific firms. In fact, trademarks are also useful for distinguishing between products that come from the same firm: for example, the Golf, the Jetta, the Passat and the Touareg are all trademarks of Volkswagen but they refer to products that have different characteristics. There are two reasons why the rapid identification of products is valuable, each of which is related to a type of differentiation that exists within markets.

First, many markets are characterized by horizontal differentiation. That is, within a category, there are consumers who want some attributes and not others (the alternatives are of equal quality but are different). In this case, the trademark assists consumers in finding the product that meets their needs. For example, Red Lobster, a well-known trademark, is known as the family seafood restaurant. This suggests that Red Lobster is ideal for families who want to eat seafood (Group 1). However, it also suggests that Red Lobster is not ideal for people who do not like (or are allergic to) seafood (Group 2). Thus, the Red Lobster trademark helps both groups find the 'right' restaurant. When a consumer is able to consistently find a product that meets her needs (and avoid the disappointment of a product that does not), she is willing to pay more (this example is adapted from Soberman 2003).

Second, in some markets, products are vertically differentiated, that is, products are 'perceived' to offer different levels of quality. A well-known trademark may evoke feelings of confidence, trust, security, efficacy and performance. For example, Bayer aspirin can be sold at up to two times the price of generic aspirin due to the strength of the Bayer trademark (Peter et al. 2008).

The value of a trademark is thus embodied in the additional amount that consumers are willing to pay for a product or service so identified. In fact, in legal terms, trademarks are valuable because of the good-will that they represent to the owner of the trademark. Goodwill is defined as the benefit and advantage of the good name, reputation and connection of a business (Smyth et al. 2010).

Relationship to Brands (and Branding)

► Brand is the general term used for the trade name of a specific product or service sold by a company. Most well-known brands are indeed trademarks. That is, they enjoy legal protection against the unauthorized use by another firm to market products in the same category (owners to a trademark can license or assign the rights to another firm). Trademark is narrower than 'brand' because identifying nicknames, symbols, monikers or emblems associated with a specific brand may also be trademarks. As noted earlier, Nike's swoosh is often used in isolation to distinguish Nike products from competitors, and nicknames such as 'the Silver Bullet' and 'the Real Thing' may be used as a substitute for the brand itself (Coors Light and Coca-Cola respectively). The literature on branding is extensive and brands represent more than the physical product to most consumers (Keller 2008). The trademark is an important aspect of branding that allows a firm to capture value from the creation of a valuable brand.

Protection of Trademarks

When a company makes unauthorized use of a trademark, that company has committed a tort known as passing off. In this situation, the owner of the trademark can take legal action to stop the unauthorized use of the trademark and claim damages that may have been incurred (Smyth et al. 2010). In addition, the government may act unilaterally to stop the unauthorized use of a trademark, especially when the use constitutes a public hazard (this is especially the case in categories like pharmaceuticals).

The Inappropriate Use of Trademarks

As the marketing of North American and European goods has expanded globally so has the inappropriate use of trademarks. This problem is particularly acute in Asia where genuine goods face competition from counterfeit products that carry well-known trademarks (Lasserre and Schütte 1999). In the absence of enforcement or efficient legal remedies, the use of another's trademark is both attractive and lucrative to those misusing the trademark. However, it is damaging to the rightful owner. The absence of redress ultimately hinders commerce.

See Also

- ▶ Brand
- ▶ Intangible Assets

References

Carlton, D.W., and J.M. Perloff. 2000. *Modern industrial organization*, 3rd ed. Reading: Addison-Wesley.

Keller, K.L. 2008. *Strategic brand management*. London: Pearson Education.

Lasserre, P., and H. Schütte. 1999. *Strategies for Asia Pacific*. London: Palgrave Macmillan.

Liodice, B. 2010. A look back at 10 ideas that changed the marketing world. *Advertising Age* 81: 14.

Peter, J.P., J.H. Donnelly Jr., and M.B. Vandenbosch. 2008. *A preface to marketing management*. Toronto: McGraw-Hill Ryerson.

Smyth, J.E., D.A. Soberman, A.J. Easson, and S.A. McGill. 2010. *The law and business adminstration in Canada*. Toronto: Pearson Canada.

Soberman, D.A. 2003. The role of differentiation in markets driven by advertising. *California Management Review* 45: 1–17.

Transaction Cost Economics

Jackson Nickerson[1] and C. James Yen[2]
[1]Washington University in St Louis, St Louis, MO, USA
[2]Washington University at St Louis, Olin Business School, St Louis, MO, USA

Abstract

This entry on transaction cost economics (TCE) outlines the core elements and arguments of TCE: fundamental transformation and the discriminating alignment hypothesis. There then follows a discussion, from a transaction cost viewpoint, of the theoretical questions concerning four strategic decisions: (1) make-or-buy decisions; (2) the decisions of internal organizational structures; (3) partnership decisions; and (4) diversification decisions. Finally, there is an exploration of recent theoretical attempts to integrate TCE and capabilities-based theories. While admittedly not being comprehensive or balanced, the entry provides useful contact points to the transaction cost economics literature in strategic management.

Definition Transaction cost economics (TCE) utilizes comparative institutional analysis to examine the governance of a transaction. Assuming bounded rationality and opportunism, TCE maintains that transaction costs could arise if a bilateral dependency occurs between transacting parties. TCE predicts a discriminating and economizing alignment between transaction attributes including asset specificity, frequency, and uncertainty and governance.

Strategic Decisions Through a Transaction Cost Economics Lens

Many managerial decisions involving the boundaries, structures and, ultimately, governance of the firm are crucial in offering an explanation of the heterogeneous performance between business organizations. Strategy research therefore emphasizes the understandings of the antecedents as well as the performance implications of various managerial decisions. Transaction cost economics has proven to be a useful theory in the understanding of many such decisions and their performance implications from both a positive and a normative perspective. Because of the breadth of its applicability, we narrow our focus to four strategic decisions spanning intrafirm and interfirm analyses. These four central issues are: (1) whether managers should vertically integrate or outsource particular activities; (2) how internal organizational structures should be designed and when they should be subject to alteration; (3) how managers structure and govern the partnerships when there are needs for collaboration to transfer knowledge; and (4) whether and into which businesses firms should diversify. Before considering these four managerial issues, there should be some analysis of the core elements and arguments of transaction cost economics.

Fundamental Transformation

The 'fundamental transformation' is central to the understanding of value creation and value capture from a transaction cost perspective. Williamson (1985) describes the fundamental transformation as a situation in which 'what was a large numbers bidding condition at the outset is effectively transformed into one of bilateral supply thereafter' (Williamson 1985: 61). Once one or both parties have made durable co-specialized investments, they cannot easily choose a different trading partner without incurring significant adjustment costs due to asset specificity, which creates an ongoing dependency relation.

TCE contributes to the understanding of value creation by pointing out that firms, in order to create value in some unique way, invest in co-specialized assets for a transaction. These investments in co-specialized assets create value either by enhancing quality or by reducing the production costs associated with the transaction. It is precisely these co-specialized investments that give rise to the fundamental transformation and pose exchange hazards to one or both parties in the transaction. For instance, the dependency relation created by co-specialized investments typically encounters uncertainty in the transaction. When uncertainty manifests, it can create the need for both parties to adapt to these new circumstances. TCE assumes that trading partners may be opportunistic, which implies that the parties may engage in *ex post* haggling or hold-up, and fail to make necessary adjustments. These maladaptation costs may reduce value or even affect which partner captures the value generated by a transaction.

TCE identifies an essential paradox embedded in a would-be value-creating transaction – that is, the value created by specific investments might be destroyed by exchange hazards. Such transaction costs arising from the fundamental transformation are thus viewed as strategic in terms of both making these investments and governing them so that value is neither expropriated nor destroyed. Consequently, TCE offers prescriptions about how firms align the level of co-specialized investments with *governance structures* that enable the investments in co-specialized assets as well as protecting those investments as the adaptation needs arise. Put differently, transaction cost economics is concerned with finding a discriminating alignment between exchange partners to economize on adaptation costs.

Discriminating Alignment Hypothesis

Williamson boldly argues that economizing 'is the best strategy' (Williamson 1991a). In this sense the term 'economizing' means designing governance arrangements to align with the transaction attributes in order to economize on transaction costs that arise from the fundamental transformation. Williamson (1991b) explicitly posits this relationship between transaction attributes and governance structures as the discriminating alignment hypothesis.

Three steps are involved in the discriminating alignment hypothesis (Williamson 2005). First, identify the exchange attributes of asset specificity, uncertainty and frequency that make some transactions simple and others complex. Of these three attributes, asset specificity is the main locomotive because it leads to the fundamental transformation. Second, specify the costs and competencies of alternative modes of governance such as the differences between market (spot markets and simple contracting), hybrid (variations of complex contracting) and hierarchy (vertical integration). For example, market mechanisms excel at autonomous adaptations (i.e., exchange partners adapt by exiting the exchange relation and seeking out alternative partners) because price signals are the only information necessary to adapt to changing circumstances. In cooperative adaptations, however, hierarchies surpass markets because they require coordinated responses (i.e., mutual adaptations often involving incremental, coordinated and co-specialized adjustments). In these instances, hierarchy, while costly, can support coordinated adjustments whereas the market mechanism cannot easily support such adaptation.

Finally, the discriminating alignment hypothesis predicts that transactions are aligned with governance structures so as to realize a transaction cost economizing result. In particular, matching markets with low levels of co-specialization, complex contracts with moderate levels of co-specialization and hierarchy with high levels of co-specialization economize on transaction costs. The discriminating alignment leads to TCE's main performance implications.

Performance Implications of the Discriminating Alignment Hypothesis

TCE maintains that firms that match transaction attributes to governance alternatives in an economizing way enjoy performance benefits in terms of survival and, implicitly, profits compared to those firms that do not organize based on TCE predictions. Although still small, there has been an increase in number of empirical studies emphasizing economic performance or firm survival in TCE (Macher and Richman 2008). Moreover, while many other managerial theories also tackle the antecedents of performance, TCE research has been among the perspectives that have given serious consideration to the endogeneity problem that is inherent in empirical studies in strategy. For example, Masten et al. (1991) provided the first estimates of economic performance at a transaction level in shipbuilding components. They found that overall organization costs in ship construction were lower when transactions and organizational forms were aligned according to the discriminating alignment hypothesis. In developing their empirical estimates they also utilized econometric methods that statistically accounted for the endogeneity problem inherent in doing comparative contractual analysis that accounts for the selection of discrete organizational forms (see also Masten 1993; Shaver 1998; Hamilton and Nickerson 2003).

In addition to examining the effect of alignment on economic performance, research has explored the extent to which transaction misalignment impacts on firm survival. For example, Silverman et al. (1997) and Nickerson and Silverman (2003) studied the discriminating alignment of the employment relation in the US trucking industry for more than a decade following its deregulation. Their empirical analyses are among the first to show increased mortality when firms do not adhere to operating policies consistent with the principles of transaction cost minimization. Moreover, in an analysis of data from the early US auto industry, Argyres and Bigelow (2007) found that the transaction misalignment has an impact on firm survival during the shakeout stages of the industry but has no significant impact during the pre-shakeout stages. This finding indicates that the industry life cycle is an important moderator for the performance implications of the discriminating alignment hypothesis.

In sum, these studies provide initial support for the performance implications of TCE's discriminating alignment hypothesis. The following sections explore the performance implications in greater detail and tackle the four strategic decisions discussed above from a transaction cost point of view.

Make-or-Buy Decisions

Following Coase's question of why firms exist, Williamson proposed to treat market and hierarchy as two alternative governance mechanisms (Williamson 1975, 1985). Adopting the transaction as the unit of analysis, Williamson assessed the make-or-buy decision as trade-offs between these two alternative governance modes. Since that time a variety of studies into strategic management have considered the issue of make-or-buy decisions. For instance, Leiblein et al. (2002) examined such decisions and found that the misaligned governance choices decrease the technological performance in the context of the production of semiconductor devices. While this study offered a direct empirical research on the performance implication of discriminating alignment hypothesis, it, like most studies, did not estimate the profitability of the make versus buy decision at the transaction level.

To our knowledge, Mayer and Nickerson (2005) provided the first estimates of profitability at a transaction level. They applied the discriminating alignment hypothesis to explain why firms organize the knowledge workers as employees versus independent contractors and to predict the performance implications of such a choice. Through an examination of the contracts of an information technology company, they identified three major contracting difficulties, namely, expropriation concerns, measurement costs and interdependence. They also assessed the alignment implications for profitability when governing the transaction through integration or outsourcing for 190 information technology service projects. Using a two-stage switching regression model, their analysis showed that projects aligned according to their version of the discriminating alignment hypothesis are, on average, more profitable than misaligned ones and that firm capability impacts organizational choice but not profitability (Mayer and Nickerson 2005).

Decisions of Internal Organizational Structures

The choice of organizational structures is critical to the implementation of strategy and it is illustrated by Chandler's (1962) famous argument of 'structure-follows-strategy'. Two problems related to the choice of organizational structures are explored in this section. One is the adoption of multidivisional structures (M-form) and the other is the choice between centralization and decentralization. First, Chandler (1962) identified the M-form as a crucial organizational innovation to fulfil the strategy of serving diversified product markets. Following Chandler, Williamson (1975) maintained that large corporations organized as an M-form enjoy performance advantages over those organized as the traditional centralized functional form (U-form) because M-form, unlike U-form, can enable clear responsibility and hence can better motivate the managers and employees in the sub-autonomous divisions. Armour and Teece (1978) provided one of the earliest empirical tests on Williamson's M-form hypothesis, which is essentially a version of the discriminating alignment hypothesis.

Second, managers can influence employees' motivations by altering the formal structures of organizations and the authority of decision-making, which is most often done by adopting a structure of either centralization or decentralization. In one study Argyres and Silverman (2004) explored how the decision to centralize a firm's R&D activities influences the type of innovation it produces. While they did not assume that more impactful or broader innovations are more profitable to the firm, they found that innovations generated by centralized R&D have a larger and a wider impact on subsequent technological actions because the centralization of research reduces internal coordination costs across units involved in the R&D.

Using a transaction cost lens Nickerson and Zenger (2002) constructed a quite different theory to explaining the decisions to adopt a structure of either centralization or decentralization. Observing that some organizations vacillate between centralization and decentralization, they developed a theory of structural modulation. Given the assumptions that governance choices are discrete and that the desired functionality lies in-between two discrete governance modes, they argued that efficiency may dictate modulating between discrete governance modes in response to a stable set of exchange conditions. Their finding is contrary to the traditional proposition of static match between organizational forms and environmental conditions, market strategies or exchange conditions (Nickerson and Zenger 2002).

T

Partnership Decisions

The explosion of interfirm activities since the 1970s has sparked much research into the structures and governance of partner relationships – what Williamson (1991b) refers to as hybrids. A joint venture is one of the focal forms of these hybrid modes of governance. A central focus of this research stream into hybrids is the exchange of knowledge, which incurs different considerations than transactions involving physical assets due to the distinct attributes of knowledge, such as tacitness and complexity (e.g., Mayer and Nickerson 2005). Hennart (1988: 369) argued that joint ventures can be 'explained as a device to bypass inefficient markets for intermediate inputs' and he also identified one of the critical non-marketable inputs as firms' know-how or knowledge. Assuming that the efficient production requires two non-marketable elements of know-how held by two firms, Hennart predicted that these two firms will form a joint venture to govern the combination of two types of knowledge.

Oxley (1997) discussed an appropriability hazard in the market for know-how in the context of interfirm alliances aiming to transfer technology between two parties. She explored how expropriation concerns for unprotected, uncodified, embedded knowledge influence governance choice. Following the market–hierarchy continuum tradition, she ranked multiple governance arrangements, such as technology licensing and equity joint venture, in a spectrum with market and hierarchy as the two extremes. She then identified the four alliance characteristics that lead to a more hierarchical governance mode. Namely, when an alliance involves: (1) design rather than production activities; (2) a broader range of products or technology; (3) a wider geographical area; and (4) more partner firms, a more hierarchical governance mode will be chosen by alliance firms. In sum, while firms collaborate with other organizations in order to transfer knowledge between one another, they try to minimize the possibility of the exploitation of knowledge by other parties.

There thus emerged a tension between knowledge sharing and expropriation. Since knowledge transfer requires close personal contact in order to perform teaching, demonstration, and presentation to deliver the 'tacit' know-how, this intimate communication channel can be easily exploited by both parties in order to obtain each other's knowledge in those areas not covered by the original transfer agreement (Mayer and Nickerson 2005). Heiman and Nickerson (2004) empirically tested the tension between knowledge sharing and knowledge expropriation. They argued that, in order to collaborate with partner firms, managers craft knowledge management practices to economize on human cognitive limits and then facilitate knowledge transfer. Although these practices are effective in facilitating knowledge transfer when knowledge is tacit and problem-solving is complex, they create further expropriation concerns as the opportunistic behaviours of partners arise. They found that with the adoption of knowledge management practices, firms are more likely to utilize equity-based governance to mitigate expropriation hazards.

Finally, the selection of partners is a critical aspect of interfirm relationships. Hoetker (2005) utilized insights from TCE, interfirm relations and the capability-based view and integrated them into a model to simultaneously consider constructs of uncertainty, trust and capability in order to choose a unique supplier for a technologically innovative component. He found that when the level of uncertainty posed by the desired innovative component is low, the selection decisions are made primarily on the basis of the technical capabilities possessed by a supplier. However, as the level of uncertainty increases, considerations of past relationships and trust are accorded greater weights than technical capabilities. In the extreme case, the uncertainty is so high that firms would make the innovative component in-house. His model links partnership decisions with make-or-buy decisions and shows that hybrids and hierarchies are both alternative governance arrangements for managers.

Diversification Decisions

Managerial research on diversification has a long and fruitful tradition under the approach of the resource-based view of the firm (RBV). Managers' drive to utilize the unused productive services of resources is seen as a major force behind the internal growth of the firm (Penrose 1959). When a firm has an unused or underutilized resource, diversification into other businesses that appreciate the value of the resource is considered a critical means to create and capture value.

Therefore, diversifying into ‘related’ businesses is recommended because related businesses will utilize existing resources. However, an alternative governance mode for exploiting the underutilized resource is to contract out to other firms that need it: for example, a licensing contract could be crafted to transfer the intangible know-how to other firms (Teece 1986). Therefore, while empirical RBV traditionally assumes that resources are difficult to transfer due to contractual hazards and hence can only be utilized by diversification, TCE offers a theory of comparative governance choices to analyse those contractual hazards before diversification decisions are made.

Silverman (1999) developed the first theoretical (see also Teece 1982) and empirical study to incorporate TCE’s constructs of contractual hazards in exploring the directions of a firm’s diversification behaviours. He hypothesized three factors that will predict a firm’s diversification directions. Specifically, a firm is more likely to diversify into a business: (1) as the feasibility of licensing its technological resources in that business decreases; (2) as the need for secrecy to appropriate returns to its technological resources in that business increases; and (3) as the degree of tacit knowledge associated with its technological resources in that business increases. In TCE terms, if the alternative governance mode such as licensing is not available to the focal firm, it will tend to utilize the resources by diversification. Moreover, if the technological resources embody knowledge attributes, such as tacitness, that increase the concerns for knowledge expropriation, the firm will tend to utilize the resources by diversification. By incorporating TCE with RBV, Silverman (1999) contributed to both managers’ and scholars’ understanding of into which industry firms should diversify. In sum, a firm’s diversification decision is influenced by the severity of hazards surrounding contractual alternatives to diversification.

TCE and the Capability-Based View of the Firm

Transaction cost economics provides a useful perspective to understand the four managerial decisions discussed above. However, these four issues are also addressed by other managerial perspectives such as the capabilities/resource-based view of the firm. Debates among researchers revolve around the predictions and prescriptions of these different theories. Thus, in this final section, we discuss two recent theoretical developments that aim to integrate TCE and capability-based theories.

The first development comes from research on the antecedents of boundary decisions. In his 1985 book, *The Economic Institutions of Capitalism*, Williamson maintained that firms should organize the transactions within firm boundaries when asset specificity is high, assuming that firms have the same capabilities. In the TCE framework differing capabilities are accounted for by incorporating a shift parameter reflective of differing production costs (Williamson 1991b). Moving beyond the shift parameter approach requires an exploration of the origins of firm capabilities. Previous boundary decisions – whether to use market or hierarchy to govern the activities – may be the drivers of the generation of firm capabilities because they represent the resulting investments in co-specialization. Within this perspective, prior boundary decisions and the resulting capabilities are themselves the results of the considerations of minimizing transaction costs.

The second development in the intersection of TCE and capability-based theories focuses on the capabilities of governing activities. Whereas the capabilities-based theories of the firm have tended to emphasize that the key firm capabilities are technological in nature, scholars are now exploring the possibility that firms develop capabilities for governing activities in ways that are similar to those in which they develop production capabilities. This research stream combines insights from both TCE and capabilities/resource-based view of the firm. For example, in the previous section on Diversification, if licensing is an option to the focal firm, then diversification may not be chosen to utilize the technological resource (Silverman 1999). Moreover, Mayer and Argyres (2004) emphasized the learning ability between two cooperating firms to design contracts and stressed the role of formal contracts as a repository of knowledge within and among firms. They also argued that the ability to design formal contracts is essentially a critical ability to deal with the fast-changing environment and that this contracting ability increases with the familiarity of interacting parties (Mayer and Argyres 2004). In sum, when facing difficult business decisions, managers and strategy researchers can benefit from these approaches to understanding the interaction between capabilities and transaction cost perspectives to make better judgements of strategic decisions.

See Also

► Capability Development

Acknowledgements The authors wish to thank Nick Argyres and Lamar Pierce for their comments and suggestions.

References

Argyres, N., and L. Bigelow. 2007. Does transaction misalignment matter for firm survival at all stages of the industry life cycle? *Management Science* 53: 1332–1344.

Argyres, N., and B. Silverman. 2004. R&D, organization structure, and the development of corporate technological knowledge. *Strategic Management Journal* 25: 929–958.

Armour, H.O., and D.J. Teece. 1978. Organization structure and economic performance: A test of the multidivisional hypothesis. *Bell Journal of Economics* 9: 106–122.

Chandler Jr., A. 1962. *Strategy and structure: Chapters in the history of the American industrial enterprise*. Cambridge, MA: The MIT Press.

Hamilton, B., and J.A. Nickerson. 2003. Correcting for endogeneity in strategic management research. *Strategic Organization* 1: 53–80.

Heiman, B., and J. Nickerson. 2004. Empirical evidence regarding the tension between knowledge sharing and knowledge expropriation in collaborations. *Managerial and Decision Economics* 25: 401–420.

Hennart, J. 1988. A transaction costs theory of equity joint ventures. *Strategic Management Journal* 9: 361–374.

Hoetker, G. 2005. How much you know versus how well I know you: Selecting a supplier for a technically innovative component. *Strategic Management Journal* 26: 75–96.

Leiblein, M.J., J.J. Reuer, and F. Dalsace. 2002. Do make or buy decisions matter? The influence of organizational governance on technological performance. *Strategic Management Journal* 23: 817–834.

Macher, J.T., and B.D. Richman. 2008. Transaction cost economics: An assessment of empirical research in the social sciences. *Business and Politics* 10: 1–63.

Masten, S.E. 1993. Transaction costs, mistakes, and performance: Assessing the importance of governance. *Managerial and Decision Economics* 14: 119–129.

Masten, S.E., J.W. Meehan, and E.A. Snyder. 1991. The costs of organization. *Journal of Law, Economics, and Organization* 7: 1–25.

Mayer, K., and N. Argyres. 2004. Learning to contract: Evidence from the personal computer industry. *Organization Science* 15: 394–410.

Mayer, K., and J. Nickerson. 2005. Antecedents and performance implications of contracting for knowledge workers: Evidence from information technology services. *Organization Science* 1: 225–242.

Nickerson, J., and B.S. Silverman. 2003. Why firms want to organize efficiently and what keeps them from doing so: Inappropriate governance, performance, and adaptation in a deregulated industry. *Administrative Science Quarterly* 48: 433–465.

Nickerson, J., and T. Zenger. 2002. Being efficiently fickle: A dynamic theory of organizational choice. *Organization Science* 13: 547–566.

Oxley, J.E. 1997. Appropriability hazards and governance in strategic alliances: A transaction cost approach. *Journal of Law Economics and Organization* 13: 387–409.

Penrose, E. 1959. *The theory of the growth of the firm*. Oxford: Basil Blackwell.

Shaver, J.M. 1998. Accounting for endogeneity when assessing strategy performance: Does entry mode choice affect FDI survival? *Management Science* 44: 571–585.

Silverman, B. 1999. Technological resources and the direction of corporate diversification: Toward an integration of the resource-based view and transaction cost economics. *Management Science* 45: 1109–1124.

Silverman, B.S., J. Nickerson, and J. Freeman. 1997. Profitability, transactional alignment, and organizational mortality in the US trucking industry. *Strategic Management Journal* 18(Special issue): 31–52.

Teece, D. 1982. Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization* 3: 39–63.

Teece, D. 1986. Profiting from technological innovation. *Research Policy* 15: 286–305.

Williamson, O.E. 1975. *Markets and hierarchies: Analysis and antitrust implications*. New York: Free Press.

Williamson, O.E. 1985. *The economic institutions of capitalism*. New York: Free Press.

Williamson, O.E. 1991a. Strategizing, economizing, and economic organization. *Strategic Management Journal* 12(special issue): 75–94.

Williamson, O.E. 1991b. Comparative economic organization: The analysis of discrete structural alternatives. *Administrative Science Quarterly* 36: 269–296.
Williamson, O.E. 2005. The economics of governance. *American Economic Review* 95: 1–18.

Transnational Strategy

Peter Williamson
University of Cambridge, Judge Business School, Cambridge, UK

Abstract

A 'transnational strategy' refers to the idea that companies can improve their competitiveness by designing a strategy that optimizes the interactions between national subsidiaries (making it a 'trans' national network). Bartlett and Goshal (Bartlett, C. A. and Goshal, S. *Managing across borders: The transnational solution*. Boston: Harvard Business School Press, 1989) proposed such an organization as 'the transnational solution' to the age-old problem of managing across borders: the issue of how an international company can improve the trade-off between the pursuit of global efficiency and responsiveness to differences between national markets. The ideal of a transnational strategy has, however, proven difficult to achieve in practice.

Definition A transnational strategy is designed to gain competitive advantage by coordinating specialized capabilities and resources distributed between organizational units located in different nations.

The common metaphor of international expansion as a process of establishing a series of 'sister subsidiaries' around the world is at least tinged with the idea, if not of cloning, of planting a series of home-nation flags on foreign soil. After all, the most widely accepted rationale for why multinational companies exist is that they have proprietary advantages (such as technologies, product designs or brands) that are valuable and scarce in foreign markets and are most efficiently replicated abroad through transfer internally within the firm rather than through trade in an external market (Dunning 1993; Caves 2007). So an international firm has to recreate some of itself in a new market if it is to be successful.

For many years, therefore, much of the research into multinational organizations focused on the question of to what extent firms should build overseas subsidiaries that were clones of their home-country organization rather than building subsidiaries that were heavily adapted to differences in the local environment prevailing in the host country (Paterson and Brock 2002). These included issues of what activities a local subsidiary should perform. Should it, for example, be restricted to sales and distribution or did it need to manufacture locally? Or should it perform most of the value-adding activities required to deliver the offering to the customer? How much should a subsidiary be allowed to deviate from corporate processes and policies to meet local market needs? How much autonomy should the local subsidiary managers be given in making these kinds of decisions?

There are powerful arguments for standardization, including the potential to reap global economies of scale, the emergence of multinational customers who demand common service levels wherever they were supplied, greater ease of benchmarking and controlling operations that adopt common processes. On the other hand, different customer tastes, needs and distribution channels, strong local competitors with offerings better adapted to local conditions, differences in culture and work practices, and idiosyncratic host government policies all argue for greater local adaptation. These choices came to be presented in the 'integration-responsiveness grid' (Prahalad and Doz 1987). Firms could choose to be global businesses – with high levels of standardization across their subsidiaries and heavy central control – or portfolios of locally responsive/multi-domestic subsidiaries that differed substantially from each other and from their parent and where local management had high levels of decision-making autonomy. 'Multi-focal' firms would choose to be somewhere in-between: standardized for some activities and locally adapted for others.

The Transnational as a Re-conception of the Multinational Corporation

The transnational strategy sought to break out of the confines of this integration-responsiveness trade-off. It conceived of the multinational firm as a 'differentiated network' (Ghoshal and Nohria 1989; Ghoshal and Bartlett 1991). The question was no longer how much to allow a subsidiary to adapt to its local environment. Instead, it was the role it could optimally play in maximizing the performance of the firm's international network. And to play this role, what capabilities would it need and how should it best interact with the headquarters and its sister subsidiaries? In this sense the transnational strategy and organization had its roots in the 'heterarchy' model (Hedlund 1986) that allowed for a variety of unique roles for each of a multinational's subsidiaries as well as the units within its corporate headquarters. This conception of the multinational as a differentiated network that transcended national markets also opened up the possibility of a new source of competitive advantage that multinationals might draw upon. Because their innovation and sources of knowledge were distributed around their international networks, new sources of advantage could be created by sharing innovations, knowledge and best practices between subsidiaries within the network as well as with the corporate headquarters. The important knowledge flows could be lateral as well as vertical.

Structures, Roles and Processes Necessary to Implement the Transnational Strategy

Bartlett and Goshal (1989) recognized that designing and implementing an optimal transnational network would be complex. Different functions, such as R&D, manufacturing, and sales and service, would need to be differently configured across countries, with some functions being present in some subsidiaries and not others depending on the relative attractions of that location in terms of variables such as cost, local knowledge, the nature of local institutions, as well as the maturity and depth of capabilities present in the subsidiary at any point in time. Some subsidiaries could be global 'strategic leaders' for certain activities, some could be 'contributors' to the global network of that activity, while others, where the location was less attractive and/or the subsidiary's capabilities weaker, could only be local 'implementers' of existing processes and offerings for that activity. In the event that a particular international location turned out to be strategically important because of its rich stock of local resources and knowledge, but the multinational either had no local subsidiary or a very weak one, that location would become a 'black hole'.

Bartlett and Ghoshal expressed the key challenges in designing and managing this transnational network in terms of continuously balancing three opposing forces: the drive for global efficiency; the need for national responsiveness; and the potential to reap the benefits of innovation and shared learning through interactions between subsidiaries as well as with headquarters units. The transnational solution recommended that these challenges be divided down into three pairs of trade-offs, each handled by a different managerial role. Global business unit managers would seek to optimize the configuration of activities for their business line across all subsidiaries, balancing global efficiency and coordination against the need for local responsiveness. Country managers located in each subsidiary would not only promote the necessary level of national responsiveness to adapt to the needs of their local market and manage local stakeholders, but also capture local learning and innovations that could be shared with sister subsidiaries and headquarters units. Global functional managers located in the headquarters, such as the heads of finance, information technology or human resources, would be not only seeking to implement common systems across all subsidiaries to gain efficiencies where appropriate, but also to promote the sharing of best practices and local innovations in their functions across subsidiaries. Through the interaction between these roles, the differentiated network could be optimized. Performance measures and incentive systems would need to be aligned with each of these differentiated roles.

Challenges of Implementing the Transnational Strategy in Practice

The theoretical ideal of the transnational strategy and structure, however, proved to be very complex to manage in practice. It required high levels of difficult coordination between different subsidiaries and headquarters units that threaten to impose excessive overhead costs and risked the organization becoming so focused on internal communication that it lost sight of external customers. In the highly interdependent transnational, where capabilities were widely dispersed and each unit depended on its sisters for success, it was difficult to hold individual managers accountable for performance and control of the direction of the company. These issues spawned a large body of research examining how subsidiaries in a multinational company behaved in practice, how they interacted with fellow subsidiaries and the headquarters, and how these interactions could be improved (Gupta and Govindarajan 1991; Birkinshaw 1996; Birkinshaw and Hood 1998).

The Long-Term Influence of the Transnational Ideal

Other researchers began to examine the knowledge flows and learning processes within multinational corporations, an area that Bartlett and Ghoshal's transnational solution had flagged as critical to a transnational network, but had not explored in detail (Frost 2001; Hansen et al. 2005). This line of thinking was taken one step further with the proposition that a modern multinational should not only be a differentiated network of subsidiaries but that firms could adopt a 'metanational' strategy which sought to build competitive advantage by mobilizing knowledge from both existing subsidiaries and external sources (such as customers, suppliers and partners) around the world to fuel innovation (Doz et al. 2001).

While it proved difficult to fully implement in practice, therefore, the transnational strategy did provide an influential step forward in thinking, encouraging researchers and managers to regard the multinational as a differentiated network, rather than a portfolio of country subsidiaries. As a result, many of today's multinationals have echoes of the transnational strategy in their structures and processes, which facilitate joint projects and knowledge exchange between subsidiaries.

See Also

▶ Multinational Subsidiary Mandates
▶ Strategy and Structure of the Multinational Enterprise (MNE)

References

Bartlett, C.A., and S. Goshal. 1989. *Managing across borders: The transnational solution*. Boston: Harvard Business School Press.

Birkinshaw, J.M. 1996. How subsidiary mandates are gained and lost. *Journal of International Business Studies* 27: 467–496.

Birkinshaw, J.M., and N. Hood. 1998. Multinational subsidiary evolution: Capability and charter change in foreign-owned subsidiary companies. *Academy of Management Review* 23: 773–795.

Caves, R.E. 2007. *The multinational enterprise and economic analysis*, 3rd ed. Cambridge: Cambridge University Press.

Doz, Y.L., J. Santos, and P.J. Williamson. 2001. *From global to metanational: How companies win in the global knowledge economy*. Boston: Harvard Business School Press.

Dunning, J.H. 1993. *The globalization of business*. London: Routledge.

Frost, T.S. 2001. The geographic sources of foreign subsidiaries' innovations. *Strategic Management Journal* 22: 101–123.

Ghoshal, S., and C.A. Bartlett. 1991. The multinational corporation as an interorganizational network. *Academy of Management Review* 15: 603–625.

Ghoshal, S., and N. Nohria. 1989. Internal differentiation within multinational corporations. *Strategic Management Journal* 10: 323–337.

Gupta, A.K., and V. Govindarajan. 1991. Knowledge flows and the structure of control within multinational corporations. *Academy of Management Review* 16: 768–792.
Hansen, M., M. Mors, and B. Lovas. 2005. Knowledge sharing in organizations: Multiple networks, multiple phases. *Academy of Management Journal* 48: 776–793.
Hedlund, G. 1986. The hypermodern MNC: A heterarchy? *Human Resource Management* 25: 9–35.
Paterson, S.L., and D.M. Brock. 2002. The development of subsidiary-management research: Review and theoretical analysis. *International Business Review* 11: 139–163.
Prahalad, C.K., and Y. Doz. 1987. *The multi-national mission: Balancing local demands and global vision*. New York: Free Press.

Tushman, Michael L. (Born 1948)

Charles A. O'Reilly III
Stanford University, Graduate School of Business, Stanford, CA, USA

Michael L. Tushman is the Paul R. Lawrence Class of 1942 Professor of Business Administration at the Harvard Business School. He has published 12 books, more than 80 articles and chapters, and won numerous awards for his contributions to our understanding of technological innovation and organizational change and evolution. He is also an extraordinary teacher who is able to translate his scholarly research into highly pragmatic applications.

While pursuing his undergraduate degree in Electrical Engineering at Northeastern University, Mike was a work-study student at General Radio, a local high-tech engineering company that specialized in high-end customized electronic instruments. During this period Mike watched as a formerly great company slowly deteriorated before being resuscitated by a new senior leadership team to emerge as GenRad, a more focused and market-oriented company. Witnessing this transformation led him away from electrical engineering and into the study of organizational change and evolution – themes that have characterized his research for the past 30 years. He first received an MS degree from Cornell, where he authored his first book, *Organizational Change: An Exploratory Study* (1974). He then entered the Ph.D. programme at the Sloan School at MIT and graduated in 1976.

After graduating, he joined the faculty at the Graduate School of Business at Columbia University. His earliest research focused on the management R&D laboratories (e.g., Katz and Tushman 1979; Tushman 1977, 1979) and how communication within the lab affected performance. As he worked to integrate his research and teaching interests, he, along with his colleague David Nadler, developed a congruence framework for organizational diagnosis (e.g., Nadler and Tushman 1980) that has become a widely used framework for both teaching and consulting.

Based on these early studies of innovation and change and in work with a series of doctoral students (notably Elaine Romanelli, Phil Anderson, Lori Rosenkopf and Mary Benner), he extended this interest to how organizations evolve over time. This has resulted in a series of groundbreaking papers. His 1985 paper proposed the idea that organizations evolve through a series of punctuated equilibria characterized by periods of stability interrupted by periods of foment. This was followed by an equally important paper in 1986 with Phil Anderson that showed how these disruptions varied depending on whether the nature of the change was competency-enhancing or competency-destroying. For example, competency-destroying changes often led to the failure of established firms and the entry of new companies. Competency-enhancing changes advantaged established firms. These two papers opened up new lines of research for understanding the success and failures of existing companies. In 2003, he and Mary Benner then showed how technological change can either accelerate or reduce the ability of organizations to adapt. This paper was chosen as the winner of the Academy of Management Review 2004 Best Paper Award.

In pursuing his research, Mike has also explored the role of senior leaders and their teams as agents of organizational change and adaptation. For example, he has shown, in an important 1996 paper with Lori Rosenkopf, that in turbulent contexts it is the combination of CEO and top management team change that is more likely to be successful in helping a firm adapt.

More recently, he has crystallized many of these early insights into a comprehensive framework, showing how organizational alignment, technological change and executive leadership can predict organizational adaptation in the face of change. This more recent research focuses on how organizations can be designed and led to simultaneously explore and exploit – to be ambidextrous or able to compete in mature as well as emerging markets and technologies. This research stream shows how ambidextrous ► organizational design can enable a firm to adapt to disruptive change (see, for example, his 1996 and 2008 papers).

Professor Tushman continues to have a major impact not only on scholarly research but through his teaching and writing, which has affected how companies and managers understand and are able to adapt to changes in technologies and markets.

See Also

- ► Innovation Strategies
- ► Management of Technology
- ► Organizational Ambidexterity
- ► Organizational Design

References

Katz, R., and M. Tushman. 1979. Communication patterns, project performance, and task characteristics: An empirical evaluation and integration in an R&D setting. *Organizational Behavior and Human Performance* 23: 139–162.

Nadler, D., and M. Tushman. 1980. A model for diagnosing organizational behavior. *Organizational Dynamics* 9: 35–51.

Selected Works

Tushman, M. 1974. *Organizational change: An exploratory study and case history*. Ithaca: New York State School of Industrial and Labor Relations, Cornell University.

Tushman, M. 1977. Special boundary roles in the innovation process. *Administrative Science Quarterly* 22: 587–605.

Tushman, M. 1979. Determinants of subunit communication structure. *Administrative Science Quarterly* 24: 82–98.

Tushman, M. 1996. Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review* 38: 8–30.

Tushman, M. 2008. Ambidexterity as a dynamic capability: Resolving the innovator's dilemma. *Research in Organizational Behavior* 28: 185–206.

Tushman, M., and Phil Anderson. 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly* 31: 439–465.

Tushman, M., and Mary Benner. 2003. Exploitation, exploration, and process management: The productivity dilemma revisited. *The Academy of Management Review* 28: 238–256.

Tushman, M., and Elaine Romanelli. 1985. Organizational evolution: A metamorphosis model of convergence and reorientation. *Research in Organizational Behavior* 7: 171–222.

Tushman, M., and Lori Rosenkopf. 1996. Executive succession, strategic reorientation and performance growth: A longitudinal study in the US cement industry in stable environments. *Management Science* 42: 939–953.