

A cognitive theory of the firm

Bart Nooteboom
January 2006

Abstract

This paper gives an overview of the main elements of a cognitive theory of the firm, and organization more widely, with a focus on innovation. The theory claims that the main purpose of an organization is to align cognition, in a wide sense of perceptions, meanings, interpretations, values and norms of behaviour, thus limiting ‘cognitive distance’ between people within an organization, in order to achieve coordination. This is done on the basis of an organizational ‘cognitive focus’ that has a ‘competence side’ related to capabilities and knowledge, and ‘governance side’ related to norms and values of conduct. The paper starts with a review of related theories of the firm: the competence view from Penrose, the transaction cost view from Williamson, and the evolutionary view from, in particular, McKelvey and Aldrich. Next comes a note on the relation between the notions of firms and organizations. The paper proceeds with an outline of the underlying theory of ‘embodied’ cognition, the notions of cognitive distance and organizational focus, inter- and intra industry differences between organizations, organizational boundaries, organization between organizations, and learning and invention.

Introduction

This paper offers a cognitive theory of the firm with a focus on innovation and learning. In this introduction, a survey is given of its similarities and differences with other theories of the firm that are closest: those of Penrose and Williamson, March’s (1991) problem of combining ‘exploitation and exploration’, and evolutionary theory of organization, in particular the work of McKelvey (1982) and Aldrich (1999), with some reference to the work of Hannan and Freeman.

Penrose (1959) proposed that firms achieve competitive advantage on the basis of organization-specific resources. That stands in contrast with the ‘market positioning view’ attributed to Porter (1980, 1985), derived from industrial organization economics, according to which organizational capabilities are easily imitated and thus do not constitute a basis for durable competitive advantage. With her view, Penrose inspired a stream of ‘resource’, ‘competence’ and ‘capability’ based theory of the firm, in the management and organization literature. That literature is too large to survey here, and while the present paper builds on that stream, I limit attention to Penrose as its fountainhead, and to the ‘dynamic capability’ view that developed later (Teece et al. 1997; Dosi et al. 2000), since in this paper also the emphasis lies on dynamics. Teece et al. (2000: 339) proposed that ‘Dynamic capabilities .. reflect an organization’s ability to achieve new and innovative forms of competitive advantage despite path dependencies and core rigidities in the firm’s organizational and technical processes’. So far, the literature has offered limited insight in how that is done, and this forms the focus of the

present paper. Holland (1975) and March (1991) indicated the problem of combining exploitation and exploration. To survive in the short term, firms must efficiently exploit existing resources, including knowledge, which requires a certain clarity and stability of goals, standards, meanings, roles, tasks, and skills, and to survive in the long term they must also engage in exploration, which entails ambiguity of meanings, and break-up of existing standards, roles, tasks, and skills. The combination of the two is a paradoxical and arguably the most important task of management. Here, use is made of Nooteboom's (2000) proposal of a path along which exploitation can yield exploration.

Penrose also proposed that expansion of the firm is constrained not by limits to economy of scale, or diseconomies of scale, related to products or by the size of their markets, since firms can expand by adding new products to their portfolio, but by the scope of managerial resources, in particular the ability of existing management to select and introduce additional management and the rate at which such incoming management can adapt to existing plans, procedures etc. I accept that, and a similar idea lies behind the present paper, but I have two questions. What, more precisely, does that managerial ability entail, and what is it that takes time for incoming staff to adapt to? Penrose (op. cit: 206) proposed that the growth of total supply of management services is faster than growth of the firm, up to a point, and then possibly declines. Why would that be? The view given in the present paper is similar to that of Penrose, but more specific, in that the firm, and organizations more widely, are seen as limited by the ability to coordinate *cognition* in the firm, and new entrants to the firm, whether managers or other staff, need to adapt to what I call the 'cognitive focus' of the firm. Or, in other words, managerial resources are seen as lying primarily in guiding and coordinating cognition in the firm. In this sense the present paper can be seen as an extension of the view of Penrose. Here, cognition is a wide notion, including intellect and rationality as well as perceptions and understandings more widely, including values, value judgements, feelings and emotions. That notion is based on a branch of cognitive science known as 'embodied cognition' (Lakoff & Johnson 1999), which will be discussed. A detailed inventory of similarities and differences between Penrose's 1959 view and the view of this paper is given in Appendix A.

Transaction Cost Economics (TCE; Williamson 1975, 1985) proposed that, given bounded rationality, organization serves to manage risks of opportunism by means of hierarchical monitoring and control. The analysis presented in this paper includes this issue of governance, in management of relational risk. However, here, next to *dealing with* opportunism, to the extent that it exists, with instruments of control, allowance is made for alignment of goals, values and motives that may *reduce* opportunism, substituting loyalty and intrinsic motivation, at least in part, for coercion and material incentives. In other words, in contrast with Williamson (1993), the present view allows for trust that goes beyond calculative self-interest (Nooteboom 2002). This is related to the idea of Hannan & Freeman (1977) of 'normative order' as part of 'organizational form'.

An important difference with TCE also is that here the focus is on (radical) innovation, which was neglected in TCE. Williamson (1999: 1103) admitted that TCE 'makes only limited contact with the subject of learning'. The present analysis does retain the notion of specific assets from TCE, as one cause of dependence that may be

opportunistically exploited. Indeed, from the present perspective of cognition, the building of mutual absorptive capacity and trust yield new forms of specific assets.

Evolutionary theories of the firm (Nelson and Winter, 1982; McKelvey 1982; Baum and Singh 1994; Aldrich, 1999; Hodgson and Knudsen, 2004), analyse the dynamics of firms and industries in markets on the basis of the evolutionary explanatory principles of *variety generation, selection and replication*. While the content of these processes in economics and organization deviate widely from that of biology, one can maintain the principles in generalized form, abstracted from all specifically biological content, in 'Universal Darwinism' (Hodgson 2002a).

The great merit of evolutionary theory is that it makes us look at what McKelvey (1982) called both 'autogenic' causes of change, from actors in organizations, in variety generation and replication, and 'allogenic' causes, in the environment of markets and institutions, in selection. Thus, it keeps us from the error of treating change as resulting from design of clairvoyant managers, neglecting bounded rationality and error, as well as the opposite error of environmental or technological determinism, neglecting actors. In a universal Darwinism, emptied of specifically biological content, the evolutionary processes need to be specified in terms of learning and innovation (variety generation), competition in markets and institutional environments (selection) and imitation, growth, teaching, training (replication). Here variety generation in the form of learning is fundamentally a matter of cognition, and replication is fundamentally a matter of communication, so that the use of a cognitive theory is elementary for an evolutionary theory of organizations and economies. However, the connection between evolution and the cognitive theory of organization is not elaborated in the present paper, but the two do fit together.

Hannan and Freeman (1977) specified 'organizational form' as consisting, next to the 'normative order' already mentioned, of 'formal structure' and 'patterns of activity'. While I accept the normative order, which will play an important role in my conception of organization, and I accept that at any moment a organization has formal structure and a pattern of activities, I claim that organizations can maintain their identity while changing formal structure and patterns of activity, and that, counter to Hannan and Freeman (1984) they may escape from organizational 'inertia'.

Part of McKelvey's (1982: 115) definition of organizations is that they are 'myopically purposeful (boundary maintaining) activity systems'. In his view, 'myopic purposefulness' entails boundary maintenance. I accept that organizations are activity systems, and that due to bounded rationality they are 'myopically' purposeful, but I do not accept that this means, as McKelvey (ibid.) specified, that, in line with the evolutionary perspective, variety generation is 'mostly blind in terms of what will work in the long term in an uncertain, changing environment'. To what extent 'variety generation', i.e. invention, is blind is something to be investigated, not assumed simply because that is what the evolutionary perspective prescribes. McKelvey (op cit: 170) saw organizational identity as formed by a set of 'dominant competencies' associated with 'primary tasks'. I sympathise with that, but organizational competence needs to be further unraveled. If it is tied to 'primary tasks', a problem arises concerning the identity of organizations with multiple primary tasks, such as conglomerates, as McKelvey recognized (op. cit: 211).

Aldrich (1999: 2) defined organizations as ‘Goal-directed, boundary maintaining, and socially constructed systems of human activity’. I accept all of this except the stable boundaries, to the extent that they refer to boundaries of activities. I claim that organizations can retain their organizational identity (not to be confused with legal identity, cf. Hodgson (2002b)), in spite of boundaries that are fuzzy, in the sense that activities are shared with other organizations, and variable, in the sense that activities are outsourced or integrated. This corresponds with one of Penrose’s ideas concerning the ‘metamorphosis of the firm’ (Pitelis 2001; Peteraf, Pitelis and Zollo, 2005).

I accept that on the whole the stability of organizations is an empirical phenomenon, and I recognize that, as many have pointed out, for evolution to work organizations should have a reasonably stable, cohesive identity (see e.g. Hodgson and Knudsen 2004). However, this can be conceived other than by means of a fixed set of skills, competencies, or primary activities.

In my cognitive theory of the firm, the identity of a firm, and organization more generally, as well as its boundaries, are seen as determined by a culturally constituted organizational cognitive ‘focus’, which limits ‘cognitive distance’ between people sufficiently to allow for efficient mutual understanding and agreement, for the sake of efficient coordination (Nootboom 1992, 1999). This entails coordination on the ‘competence side’ of knowledge, skills, or competencies, as well as on the ‘governance side’ of goals, interests, motives and conflict resolution. The latter corresponds with Hannan & Freeman’s (1977) notion of the ‘normative order’.

This cognitive view of an organization as a ‘focusing device’ is similar, to a greater or lesser extent, to views of organization as ‘a system for sense-making’ (Weick 1995), ‘collective mind’ (Weick and Roberts 1993), system of ‘shared meanings’ (Smircich 1983), and a ‘knowing organization’ (Choo 1998).

The present theory includes an explanation of the boundaries of the firm and of inter-organizational collaboration. Most fundamentally, while cognitive focus gives organizations an advantage over ‘the market’ (whatever that precisely is), its disadvantage is that by its nature it also entails a myopia that needs to be repaired with complementary cognition from outside the organization, at a greater cognitive distance, in ‘external economy of cognitive scope’ (Nootboom 1992).

The paper is structured as follows. First comes a discussion of the definition of firms and the relation between firms and organizations. Second comes a summary of the ‘embodied cognition’ theory of cognition used, and the notion of ‘cognitive distance’ that it yields. Third, the notion of organization as a ‘focusing device’. Fourth, inter- and intra industry differences between organizations. Fifth, organizational boundaries, sixth, organization between organizations, and seventh, transformation, learning and invention.

Firms and Organizations

Here, organizations are defined as:

Myopically goal-directed, socially constructed, cognitively focused systems of coordinated activities.

This definition adopts several elements from definitions offered by McKelvey (1982) and Aldrich (1999), but also deviates from them. The ‘system of activities’ includes actions, knowledge and technology, and competencies in which the three come together. As noted by McKelvey, goal directed systems of activity generally entail a certain focus on core or distinctive competencies. A certain stability of the system of activities is needed for an organization to function, compete, build absorptive capacity (Cohen & Levinthal 1990), build and retain competencies, attract and train people, and to build internal and external relationships. However, in contrast with McKelvey and Aldrich, the definition does not stipulate stability of the organization’s boundary of activities. This allows for an organization to outsource activities without becoming a different organization.

Note that the definition of organization given above does not require organizations to have multiple persons, and includes the single-person, owner/manager firm. That would also *need* to be the case, since it would be odd if a firm were not an organization until an owner/manager engaged a partner or employed another person, even if it were a part-timer for only, say, one hour a month, or if it stopped being an organization if it hived off its last employee. Excluding single person firms would also entail that roughly 50 % of all firms would be excluded from organizational analysis. Single owner/manager firms satisfy the definition of an organization in that they entail coordination of actions, knowledge and technology. They need to coordinate the search and handling of material and informational inputs, technology, related knowledge and skills, search, knowledge and supply of customers, and knowledge of institutional conditions (legal, financial, etc.) and how to deal with them.

For example, a single person shoe repair shop typically repairs shoes as well as other leather goods (purses), may manufacture simple leather goods, such as belts, may add the cutting of keys, and sells related goods, such as polish, laces, purses, and belts. It requires the requisite skills and machinery. The craft activities should have a short cycle, allowing for frequent and sudden interruption to serve incoming customers, and thereby utilize time in between customer arrivals that would otherwise be idle. The choice and coordination of space, machinery, portfolio of activities, purchasing, finance, administration and requisite knowledge and skills requires organization.

Nowadays, in the Internet economy, in many cases one can have an idea for a product, outsource its artistic and technical design, outsource its production, probably in a low-wage country, search for and contact suppliers via the Internet, have materials shipped directly to the producer, outsource distribution of the product to customers, search for customers, contact them and maintain communication with them on product quality via the Internet, send their data to an outside agency for invoicing, and have the data sent to an outside accountant/tax consultant who also files tax returns. This entails much organization but can be done by a single person.

While traditionally production has been seen as the physical transformation of physical inputs into physical outputs, this applies too narrowly to manufacturing while the majority of employment now lies in a wide variety of services. Therefore, the notion of production has to be extended to other ways of adding value. While manufacturing yields a utility of form and function of some physical good, in services one can produce utilities of time and place, in distribution and transport; of financial means and security, in banking and insurance; of a wide variety of repair and installation services; of physical

and mental well-being, in health services; of entertainment and enjoyment, in the hospitality and sports business; of knowledge and information, in the IT and media business, teaching and consultancy; of justice, in legal services; of a wide variety of public services and amenities. Here inputs and outputs may be abstract, in the form of information or feelings, and transformation may concern reduction, abstraction, expansion, combination or shift of information, mental or psychic states, etc.

Embodied Cognition

For my cognitive theory of organization I adopt a view of knowledge and learning that is much used in the organization literature, and is known as the ‘activity theory’ of knowledge (Blackler 1995), according to which mental models or categories or schema’s of knowledge are developed from experience in interaction with the (physical and social) world (Kolb 1984; Levitt and March 1988). Though not often recognized, this goes back to the developmental psychology of Piaget (1970, 1974) and Vygotsky (1962), according to which ‘intelligence is internalized action’. For example, in child development groping and prodding develop into pointing, which forms the basis for reference that is the basis for language. Playing with blocks provides a basis for learning to count, add and subtract. Experience in dealing with physical objects yields metaphors by which we construct abstract concepts (Lakoff and Johnson 1980).

In sociology this view is related to symbolic interactionism (Mead 1934, 1982). In philosophy, it is closely related to American pragmatism (since Peirce 1957) and the ‘organicism’ of Whitehead (1929), according to which any element in the system is an outcome of relations with other entities, and in which individuals both constitute and are constituted by society (Hodgson 1993: 11).

This organic, interactionist view is crucial, since it provides a perspective from which we may transcend the otherwise seemingly irreconcilable gap between economics, with its methodological individualism, and sociology with, in some branches, its tendency towards methodological collectivism. The individual is social in that one derives one’s individuality in interaction with others, but what one makes of the interaction is not the same as what others make of it. Individuality is a function of inherited endowments of mental constructive potential and interactions along individual courses of life that yield the experience for construction. Hence there is ‘cognitive distance’ between people to the extent that they have developed along different paths, in different environments. This distance is both a problem, for mutual understanding and collaboration, and an opportunity, to learn something new from people who have constructed their cognition differently (Nooteboom 1992, 1999).

While the perspective of cognition adopted here is connected with interpretive views of knowledge and meaning (Berger and Luckmann 1967; Weick 1979, 1995), it is less subjectivist than some of them. I maintain that even though we cannot claim to know the world objectively, since we cannot ‘descend from our mind’ to test the claim, it is reasonable to assume that an external reality does exist, somehow, and that if indeed our mental structures are constructed in interaction with it those structures in some sense represent that reality, in what Lakoff and Johnson (1999) called ‘embodied realism’. This presumption of some external touchstone of reality allows for evolutionary selection of

knowledge. A mental 'neural Darwinism' also seems to obtain, in which neural structures emerge, in selection, reinforcement and reproduction, according to their performative success (Edelman 1987, 1992).

Cognition is thus 'embodied' as well as embedded in the outside world, and its embodiment denies any Cartesian dualism of body and mind (Damasio 1995). There is a continuum of bodily, endocrinal, involuntary neural, emotional, and rational levels of activity that interact (Damasio 2003). Note that in this perspective cognition is a fundamentally social notion, in its construction from interaction. It is also a broad notion, including value judgments, feelings and emotions. As a result, cognitive distance has many dimensions. It has a more substantive side of cognition in a narrower sense of job-related knowledge and skills, and a more intentional, normative, moral side of goals, values, interests and ways of resolving conflicts.

In economics, a connection between rationality and emotions was recognized by Herbert Simon (1983). In view of limited mental resources for rational evaluation, it is rational, in the sense of adaptively efficient, to relegate familiar behaviour to routinized conduct, in what Polanyi (1962) called *subsidiary awareness*, so that scarce capacity for attention can be dedicated to unfamiliar situations, in *focal awareness*. However, a problem with routinised conduct is that it may continue in situations in which it does not apply. A switch is then needed back into focal awareness, and this may require a jolt of emotion, on the occasion of unfamiliar events.

One of the attractions of embodied cognition is that it provides continuity with social psychology, with its insights into decision heuristics that mingle emotions and rationality (Bazerman 1988). According to social psychology (Kahneman and Tversky 1979; Kahneman et al. 1982; Lindenberg 2000, 2003; Tversky and Kahneman 1983), we have multiple, sometimes mutually conflicting, mental frames, as complexes of mental schema's, in terms of which we interpret events, attribute competencies and intentions to people we interact with, and which guide our actions. One frame may be oriented at 'guarding one's interests' (what an economist attributes universally and unconditionally to people), another at 'acting appropriately' (what a sociologist tends to attribute to people; see Lindenberg). Our actions are interpreted by others as signaling underlying mental frames (Lindenberg op. cit; Six 2005). At any moment one frame may be 'salient' or in 'focal awareness' (Polanyi 1962), but relational signals may yield a switch to another frame, leaping from subsidiary awareness, by which our view of what happens and what needs to be done also switches. Thus we may vacillate between self-interest and acting appropriately.

Concerning learning and innovation, a crucial point from the above is that learning takes place on the basis of experience and on the basis of interaction with others at sufficient but not too large cognitive distance. It will be argued that in both cases what is essential is that new ideas arise from applying one's existing knowledge in novel contexts, supplied by new areas of application or new relations.

Organization as a Focusing Device

The notion of cognitive distance forms the basis for a cognitive theory of the firm (Nooteboom 1992, 1999, 2000). One problem for achieving collective goals lies in the

cognitive distance between its members. If it is too large, it is very cumbersome to coordinate competence and governance. More precisely, it would be cumbersome, and inefficient, to achieve mutual understanding on perceptions of the environment, goals and priorities of the firm, relevant technologies, products, markets, actions in jobs and roles, and technical coordination between them, in 'dominant competencies' (McKelvey 1982), on the competence side, and categories, values and rules for alignment of interests and priorities, ways for people to deal with each other, and conflict resolution, on the governance side. To function as a coordinated system of actions, organizations need some more or less specialized shared language or jargon, shared perceptions, understanding and morality, as part of organizational culture (Schein 1985). Without such focus of shared perceptions, meanings, understandings and values, too much effort, time and aggravation would have to be spent on disambiguating meanings, eliminating misunderstanding, setting priorities, establishing directions, coordinating activities, and negotiating the terms of collaboration. This organizational cognitive focus, produced and reproduced by organizational culture, forms the core of organizational identity. It constitutes a cohesive whole of perceptions, meanings and values that define roles, relations and procedures of interaction, and thereby yield cohesion and stability of organizations. It constitutes the advantage of organization over market, in contrast with, but not entirely separate from, the logic of transaction cost economics.

The intentional, moral, institutional side of organizational focus, concerning values and norms of behavior and conflict resolution is needed, among other reasons, because imperfect monitoring and measurement of performance hinder alignment of purpose and conflict resolution by means of incentives. Moral guidance partly replaces, and complements, purely extrinsic motives by more intrinsic ones that require less monitoring. The importance of this has increased since work has become more professional and based on higher levels of knowledge work that are more difficult to monitor and measure.

In addition to the distinction between the competence and governance side of focus, there are three dimensions for both: *width*, i.e. the range of different areas of knowledge, competence and governance, *depth*, i.e. the variety of concepts and rules in each area, and *flexibility*, i.e. allowance for improvised, unforeseen meanings, standards, division of labour, roles, skills, procedures, and the like. The distinction between width and depth is similar to such distinction in product differentiation and portfolio's of products in general. There, width refers to the different kinds of products, and depth to the variety of qualities and brands in each. Here, the contrast between firms and markets appears: firms yield maximum scope, variety and flexibility, in both competence and governance. That would be impossible to coordinate, in central planning, and when tried would yield stagnation: the fixing of ideas, meanings, standards, division of labour to maintain exploitation would disastrously limit the scope for exploration.

Organizational focus cannot be integrally and instantly re-shaped as a function of experience in selection, and this yields organizational inertia (Hannan and Freeman 1985), to some extent. The limits to such change lie in the systemic cohesion of elements of cognitive focus and in the fact that cognitive focus serves as an absorptive capacity that tends to mostly confirm itself in its functioning (imprinting) and thereby is path-dependent. However, and this will be discussed later, there is a process by which absorptive capacity does transform itself in its functioning, so that there is some escape

from inertia, but in a series of conditioned steps that require time. It is an empirical question to what extent the speed of that is sufficient to escape from competitive pressures.

Organizational focus emerges from the imprint from the entrepreneur who started the organization, is subject to some drift due to turnover of staff, and to shifts due to crises, caused, in particular, by shifts in the environment, or by new, challenging interpretations of the environment, and by the weeding out by selection, in population effects. When resources are scarce and competition is tight, selection is likely, in the long run, to yield organizational cognitions and structures that reflect the exigencies of the environment of markets and institutions. Consider, for example, the view that stable environments tend to favour 'mechanistic' environments while turbulent environments tend to favour 'organic' ones (Burns and Stalker 1961), or more specialist vs. more generalist organizations (Hannan and Freeman 1977).

Organizational focus will be narrowest in width and depth, and cognitive distance smallest, in single-person, owner-manager firms, wider in work groups, such as 'communities of practice' (Brown and Duguid 1996), wider yet in larger firms consisting of multiple groups, and widest in multi-divisional firms. On the basis of technological profiles Patel and Pavitt (2000) found that large firms have a wide scope of technologies in their portfolio, ranging between 10 and 20 different areas of technology, including ones that do not belong to their core activities. On the other hand, flexibility tends to be higher in small firms, for several reasons. The system to be coordinated is less complicated and coordination can often take the form of direct supervision (Mitzberg 1983). There is less need to specialize and codify knowledge for division of labour and coordination across diverse units at different locations. In larger firms, especially width and depth of job-related variety is larger, in a wider and deeper division of labor, and flexibility is less to the extent that exploitation is systemic or complex, i.e. there are many component activities, which are densely connected, with narrow constraints on the linkages (Teece 1986; Langlois and Robertson 1995; Levinthal 2000). Flexibility then is limited since it would jeopardize systemic integrity. However, in large organizations distance on the moral side of cognition is not necessarily larger. In fact, since in face-to-face work groups there is more informal, spontaneous social control of free-ridership (Simmel 1950), there the need for more explicit moral focus is less. In larger organizations more attention may be needed to the moral dimension of organizational focus across different work communities (Campbell 1994).

Intra- and Inter-Industry Differentiation

Cognitive distance also applies to the higher aggregation level of organizations, in differences in organizational focus, i.e. differences in shared language, meanings, perceptions, understandings and values and norms of behaviour (Schein 1985). In empirical work, measures of the cognitive distance between firms have been constructed on the basis of indicators from organizational data and technological profiles derived from patent data (Wuyts et al. 2005; Nooteboom et al. 2005).

Within industries, there is often considerable cognitive distance, thus difference in organizational focus. This is in line with the resource/competence based view that firms

vary, even within industries, in their distinctive competencies. Yet, these differences are smaller than inter-industry differences, particularly concerning the substantive side of technologies and skills, due to shared technologies, market demand, market structures, technical and professional standards, etc., within industries, yielding what may be seen as a common pool of competencies (McKelvey 1982). On the basis of technological profiles of a set of large firms, constructed from US patent data, Patel and Pavitt (2000) confirmed that is considerable variation in such profiles within industries, and (even) greater variation between industries.

As a result of within-industry similarities of skills and technologies, to some extent staff exchange between organizations is feasible and can create and confirm the identity of an industry (McKelvey 1982: 197), yielding 'industry recipes' (Spender 1989). This is also enhanced by pressures towards conformity from needs of social, political or financial legitimation (Dimaggio and Powell 1983). However, while firms may share component skills and technologies, particularly within an industry, their overall composition, in a coherent system, may differ substantially. To the extent that the system is complex, or systemic, as defined earlier, piecemeal, local variation, of single elements, is problematic, since it soon affects the systemic integrity of the system as a whole (Teece 1986; Langlois and Robertson 1995). Then, when there is one difference between firms there tend to be many, in an alternative systemic whole (Levinthal 2000). In sum, when operational structure is complex, with similar components integrated systems may differ greatly. Between industries, also the components differ, to a greater extent.

Systemic effects of complementary activities are not only technological and epistemic but also behavioural. Even within industries organizational focus is more varied, and organizational cognitive distance is correspondingly greater, on the moral, intentional, institutional side, in different styles or cultures of management. While firms within an industry may use the same technology, in the production of similar products for similar markets, their organizational cultures and management styles may be quite different. Deep differences in fundamental perceptions, views and (largely tacit) assumptions concerning man, his knowledge (e.g. objective or constructed), his relation with his environment (passive or active), his morality (basically good or bad), and relations with other people (egotistic or altruistic) (Schein 1985), yield differences in risk perception and acceptance, pro-activeness ('locus of control'), formality or informality, rivalry or cooperation, intrinsic or extrinsic motivation, instruments and styles of governance and conflict resolution. From an evolutionary perspective, the persistence of such differences, in spite of selection pressures, suggests that on the moral, intentional side there are different ways to be successful, within an industry.

As indicated earlier, a central issue is how to make the combination of exploitation and exploration (March 1991). While earlier some industries were relatively stable, allowing for a focus on exploitation, and others were in a state of flux, yielding a focus on exploration, now the combination of the two is needed in most if not all industries. Combination of the two is particularly difficult when exploitation is highly systemic, as defined earlier. Then, by definition, units within the system hardly have any room for the experimentation and deviation needed for exploration, since they would jeopardize systemic integrity. In that situation, exploration needs to be relegated to a different time or place. The classic case is the division between departments for production and for R&D. This yields the classic problem of divergent mentalities and priorities between

them, with resulting misunderstandings, conflicts and recriminations. It is difficult to find an organizational focus that accommodates both. One method is to engage in cross-functional teams, and another is frequent staff rotation, with an organizational focus to support that. Another would be to create more flexibility by decomposing the exploitation system into more autonomous parts, as long known from systems theory. Such choices may be made differently by different firms even within an industry.

How stable are organizational identity and hence cognitive focus, and differentiation between organizations? In their study of large firms Patel and Pavitt (2000: 317) found that: '90% of firms have profiles of technological competence that are statistically similar between 1969-74 and 1985-90' which 'remains true even after taking account of acquisitions and divestments', and '.. of 41 of the largest firms only one had a technological profile statistically different'. One argument for such stability derives from the earlier analysis of systemic or complex structure of complementary activities. When piecemeal, local change of a single element entails systemic breakdown, in a tightly coupled system, there is a formidable obstacle to change. However, there are also cognitive reasons and reasons of cultural reproduction. Existing forms of thought that yield absorptive capacity tend to consolidate and confirm themselves in the process of absorption, thus yielding cognitive path-dependency. However, according to Piaget there is a process by which absorption can also lead to a shift of cognitive structures, and this formed the inspiration for Nooteboom's (2000) 'cycle of discovery', that will be used, in a later section, to deduce a dynamic capability that can enable a firm to escape from its path-dependencies. Cultural differentiation between organizations is maintained, in spite of turnover and exchange of staff, because in the entry into an organization there is self-selection according to expected fit to organizational culture, as well as adaptation by socialization into organizational culture. Furthermore, according to the idea of intelligence as internalized action the further development of cognition reflects the environment, in this case the organization, in which it takes place. This confirms the problem of limited opportunity and speed of integrating new staff, identified by Penrose.

As noted before, between firms in different industries there are greater differences also on the substantive side of cognitive focus. There are, for example, deep differences in professional skill. As McKelvey (1982: 202) phrased it '.. Would you fly on an airplane that had recently been staffed with non-airline employees? Would you enter a coal mine operated by hotel employees? Would you eat in a restaurant staffed by truckers?'. However, even between industries isolation is far from complete, and replication across industries does take place. McKelvey (1982: 206) suggested that to the extent that organizations are simpler characteristics are more easily exchanged, also between industries.

For an illustration, consider the emergence of self-service. It emerged in retailing, largely outside large firms, initiated by independents but swiftly adopted by large chain store firms after it proved a success. Self-service retailing constitutes a distinct 'species' from service retailing, with a different structural logic, in that a fundamental reversal of roles occurred between shop attendant and customer. In service, the attendant moves about to collect items for a shopping basket, while the customer remains stationary at a counter, and in self-service these roles are switched, with the customer picking out its own goods, and a stationary attendant at a check-out, in a different lay-out of the shop. This eliminated an obstacle to shop size. In a large shop, with many products, under

service the attendant would have to move about too far, with an unacceptable increase of waiting time for the customer. The emergence of self-service, with its attendant opportunity for larger shop size, was favored by a shift of the selection environment towards knowledgeable customers who no longer needed advice from shop attendants, an increased demand for less frequent, 'one-stop' bulk shopping, due to greater scarcity of time, enabled by transport capacity from car ownership and by refrigerated home storage capacity. In its turn, self-service affected selection conditions, in co-evolution, in that it enabled economies of scale that pushed out small shops. With its demand for pre-packaged goods, it also had wide repercussions for the selection conditions in packaging and food industries. In replication, however, isolation was very limited. The principle of self-service was quickly and widely adopted in other industries, such as restaurants.

It is doubtful whether organizational focus could survive a merger or acquisition, and this contributes to their frequent failure. In view of greater difference in focus between than within industries, mergers and acquisitions are more likely to succeed within industries than between industries (Nooteboom 1999), and this is confirmed empirically (Bleeke and Ernst 1991). Patel and Pavitt (2000: 318) also found, on the basis of technological profiles derived from patent data, that in acquisitions 'Only in very few cases involving substantial technological activities were the technological profiles of the acquired firm different from the acquiring firm'.

Organizational Boundaries

One reason for integration of activities in a single organization, recognized also by transaction costs economics, is technological inseparability (Williamson 1975). Skinning an animal for leather cannot be separated from skinning the same animal for meat. It takes two hands to clap and two people to move furniture. However, as discussed earlier, there are also wider, systemic effects of complementarity that yield a complex structure of many, sometimes densely and tightly connected, component skills, roles and technologies, which appear to be neglected by transaction cost economics. Beyond that, however, according to transaction cost economics there are reasons of governance, in the control of opportunism by hierarchical control, under behavioral and environmental uncertainty. In the present paper the argument is that a culturally determined cognitive focus may be necessary for coordination, not only to deal with opportunism but also to limit it. As indicated this includes both an epistemic order for perception, interpretation and mutual understanding, and a normative order to limit opportunism. It follows that the advantage of organizational focus, relative to fragmentation among independent producers, is greatest to the extent that trustworthiness cannot be assured by outside institutions. Like cognitive focus, trust(worthiness) has two dimensions, of competence and intentions (Nooteboom 2002).

On the competence side, integration is needed more to the extent that knowledge is tacit and difficult to codify, requiring close and frequent interaction, or is systemic and inseparable, requiring close-knit communities of practice, or is so new that no widespread absorptive capacity for it has yet emerged. In particular, organizational focus is needed when there is a lack of reliable judgement and control of trustworthiness of competence by outside institutions, such as technical standards and certification, and adequate

monitoring and control, offered by professional or industry associations or governmental agencies. On the governance side, organizational focus is needed to the extent that it is difficult to align and secure interests and priorities by means of formal contracting, supported by adequate monitoring and control, the industrial and institutional environment does not support trustworthiness in the intentional sense of benevolence (limitation of opportunism), there is no adequate ethic of decency in business dealings, and there is no effective reputation mechanism. Also, the degree of opportunism to be expected depends on outside pressures of survival (Pettit 1995; Nooteboom 2002). Sharper price competition will reduce the ability of firms to 'give and take' and to surrender some advantage for the sake of loyalty to partners.

However, the advantages of organizational focus are to be traded off against its disadvantages. As noted in TCE, the (legal) independence of an organization, with its own responsibility for survival, yields a more powerful incentive for efficiency and innovative ingenuity and enterprise than protection within an integrated firm. From a cognitive perspective I would add that the advantage of 'the market' is that it offers more width, depth and flexibility of competence and knowledge than could effectively be coordinated inside an organization. This will presently be analysed in more detail.

Earlier, I indicated that, in contrast with earlier scholars of organization (McKelvey, Aldrich), I do not accept stable boundaries of activity as necessarily part of organizational identity. Organizations can hive off or integrate activities without thereby having to substantially alter their cognitive focus. Note, however, so that while organizations may not have clear boundaries of activities, in sharing activities with other organizations, and may have shifting boundaries, in outsourcing and integrating activities, firms do and should have a clear legal identity, as pointed out by Hodgson (2002a). Unclear boundaries of legal ownership and liability would create institutional havoc. Note also the condition, familiar from the alliance literature, that when organizations outsource activities they must often still retain absorptive capacity with respect to those activities, to properly collaborate and coordinate with outside sources (Granstrand et al. 1997) in what Patel and Pavitt (2000) called 'background' next to 'core' competencies. In other words, some of the related competence remains part of organizational focus.

Now the question becomes: what determines the boundary of cognitive focus? This, I propose, basically entails a trade-off between diversity, with larger cognitive distance, for the sake of exploration, and cohesion, with smaller cognitive distance, for the sake of exploitation. As indicated earlier, firms require exploitation for survival in the short term, and exploration for survival in the long term, avoiding 'inertia'. Combination of the two is problematic to the extent that exploitation requires narrow focus while exploration requires wide focus. As indicated earlier this is the case, in particular, if exploitation is systemic. Then, if exploration needs to proceed simultaneously with exploitation, it needs to be located elsewhere, and a problem arises in maintaining a focus that accommodates both. Then, there is the option to do one and seek the other in collaboration outside the firm. The classic example is biotechnology, for example in the drug industry, where, in many cases, small technology based firms provide the exploration and large pharmaceutical firms provide the exploitation. The literature on biotechnology is too large to survey here, but for a recent study of this division of labour for exploration and exploitation, see Gilsing and Nooteboom (2006).

However, from the present cognitive perspective there is a more fundamental reason for outside collaboration. Organizational focus creates organizational myopia, and in addition to all the other motives for inter-firm alliances, familiar from the extensive alliance literature, this gives an additional, cognitive reason, to prevent myopia by means of complementary outside cognition from alliance partners (Nooteboom 1999), in what Nooteboom (1992) called 'external economy of cognitive scope'. Questions then arise concerning organization between organizations. That is the subject of the next section.

Organization between Organizations

If there are cognitive and other reasons for inter-organizational collaboration, how does the organization of that compare to the notion of organization discussed so far? Basically, it entails organization between different cognitive foci, corresponding with different organizations, at a greater or lesser cognitive distance between them. This requires some shared but less detailed focus, i.e. with limited width, relating only to few activities, sufficient for mutual absorptive capacity to cross cognitive distance, in 'forms of organization between market and organizations'.

Here, as in the collaboration between people, but now on the level of organizations, the challenge is to find partners at optimal cognitive distance. As analyzed in Nooteboom (1999), mutual understanding decreases with distance, but novelty value, i.e. the potential for Schumpeterian 'novel combinations', increases. In collaboration for exploitation, the novelty value of distance is less relevant, or even absent, while in collaboration for exploration it is crucial. If performance is the mathematical product of novelty value and mutual understanding, then an inverse U-relationship between performance and distance ensues, implying some optimal distance, large enough to yield novelty, but not so large as to preclude mutual understanding and willingness and ability to collaborate. This is illustrated in Figure 1.

This optimum is not fixed. In particular, absorptive capacity, as a function of organizational focus, depends on the accumulation of knowledge and competence from past R&D (Cohen and Levinthal 1990), production, marketing, organization, and in particular, experience in collaborating with others at sufficient cognitive distance. In other words, experience in dealing with others who think differently yields competitive advantage, in a firm's own markets as well as in the 'market for alliances'. Larger absorptive capacity yields an upward shift of optimal cognitive distance, as illustrated in Figure 1. The hypothesis of optimal cognitive distance, and its dependence on cumulative R&D, was empirically confirmed by Nooteboom et al. (2005), in an econometric study of the productivity in terms of exploratory patents from 994 alliances between 116 firms in several industries in the period 1986-1996.

Collaboration between organizations with different foci requires the development of some shared focus, in the sense of sufficient common understanding, needed for coordination and collaboration, on both the competence and governance side. However, whereas integration instead of collaboration would require further merging of cognitive foci, for the sake of coordination within a firm, maintenance of independence, and minimization of shared focus needed for mutual understanding, allows for greater cognitive distance, and hence for more cognitive variety for the sake of innovation. From

this perspective, alliances are dynamically more efficient than mergers and acquisitions (Nooteboom 1999). This, I think, is a feature that yields a cognitive limitation of firm size not taken into account by Penrose (1959).

There is much more to be said about details of the organization of inter-organizational relations, but that goes beyond the scope of the present paper. Very concisely, on the governance side there are instruments derived from TCE, such as contracting plus requisite monitoring, in so far as feasible in view of uncertainties of environment and behaviour, a balance of ownership of relation-specific assets, balance of mutual dependence more generally, and posting of hostages, typically in the form of competitively sensitive information. Other instruments are reputation mechanisms and trust building by cultural alignment of values, personal empathy and identification, and routinization of conduct (Nooteboom 1999, 2002).

Also, mutual understanding and agreement, and opportunities for creation and access to novelty and absorptive capacity, depend on structure of networks, in terms of density and 'structural holes' (Burt 1992), positions in such structure, such as (different kinds of) centrality vs. peripherality, and the content and strength of ties. The latter are to a greater or lesser extent amenable to strategic choice and action. Much research on such effects of network structure and ties has been conducted in social network research. That literature also is too large to survey here, but for a recent study see Gilsing & Nooteboom (2005).

Learning and Invention

A crucial remaining issue concerns dynamic capabilities, i.e. the ability of organizations to learn, make inventions and transform themselves and avoid inertia.

According to most accounts, the main trigger of radical innovation is a shock in the form of a break or shift of the environment, which may increase competition for scarce resources, disadvantage incumbent firms, and create new opportunities for novelty. Such a shift or shock may be due to natural disaster, political upheaval and war, a shift in demand, a shift in institutions (e.g. regulations for protecting the environment), or a shift due to developments in related industries or markets. An example was given earlier, in the emergence of self-service in retailing, stimulated by an emerging demand for one-stop, bulk, low-service shopping, and its implications for conditions in other industries (packaging and food). However, this tells us only of new opportunities, of how radical innovation is enabled or challenged, not of how it is generated.

In evolutionary theory, generation of new variety, in new ideas, is generally ascribed to errors in replication, mistakes, random, uninformed trials as steps into the dark ('mutations'), or to variations guided from higher level 'search routines' (Nelson & Winter, 1982). Other literature also suggests that there are higher level 'dynamic capabilities' that direct the change of lower level capabilities (Teece et al. 1997). Dynamic capabilities include rational inference of cause-effect relations, rules for experimentation, ability to utilize organizational memory, and exchange of codified knowledge with others, in what (Nonaka and Takeuchi 1995) called 'knowledge combination' and Zollo and Winter (2002) later called 'deliberate' as opposed to experiential learning'.

As indicated earlier, I concur that for learning interaction with other people having different knowledge is a crucial source, to profit from ‘external economy of cognitive scope’ (Nooteboom 1992). I add that for this one should seek optimal cognitive distance. Thus, part of dynamic capability is the ability to find partners, at optimal distance, and to effectively understand and collaborate with them, in the governance of ‘relational risk’. This dynamic capability in the form of alliance capability can be developed by building absorptive capacity and experience in communicating and collaborating with partners who think differently.

Using a simulation model, March (1991) showed how the generation of new ideas, in exploration, can follow from novel combinations that result from personnel turnover, where people from outside an organization carry fresh ideas into the organization that may disturb the efficiency of exploitation but contribute to exploration. That is certainly part of the process of generating novel combinations, but it is also limited due to the isolating mechanisms indicated before, especially between industries, and self-selection of entrants to fit organizational focus, and socialization into that focus.

In addition to such learning and innovation by novel combinations of heterogeneous sources of existing knowledge, there is also a more experience-based form, in a path of invention where exploration emerges from exploitation (Nooteboom 2000).

Innovation may arise as a splitting off from existing dominant designs. Again, the earlier example of self-service applies. However, the more radical or competence destroying the innovation is (Tushman and Anderson 1986; Anderson and Tushman 1990), and in more dimensions (technology, product, distribution, consumption, production process, supplier industries, organization, human resources, education and training, legal and financial resources needed), the more difficulty it will meet to emerge and survive from within. Even if wider industry conditions allow for an innovation, on a lower level, within existing firms, there may be obstacles that force the innovator to exit and start on his own, in the classical phenomenon of the entrepreneurial spin-off.

This applies to the case of self-service. Prevailing industry conditions, in availability and production of goods, location of shops, distribution of goods to shops, etc. did not block its emergence, though they did inhibit the full realization of its potential until the innovation affected industry conditions due to its success. Dreesmann (1963) recounted how as early as 1912 the principle of self-service was introduced in the US, in California, on the initiative of a few independents. Supermarkets first broke through in 1930, in the state of New York, on the initiative of M. Cullen, who is a paradigm case of the entrepreneurial spin-off. He was employed by the third largest food company in the world, submitted his ideas to the board, did not obtain support, and set out for himself. Two years later he had eight supermarkets, and large firms did not follow suit until 1937.

On some level, depending on how competence destroying an innovation is, and in how many dimensions it innovates, the entrepreneur has to escape from the grip of the existing selection environment, within a job, profession, organization, industry, country or wider economic system.

The first reason for such exit or escape is to gain the *opportunity* of being different, without thereby succumbing to the grip of selection (Baum & Singh 1994: 13). However, there are other, cognitive reasons to exit to a different environment (Nooteboom 2000). A second reason is to gain the *motive* to change a dominant design, a third reason is cognitive: to gain new *insight* into where the limitations of existing

dominant designs lie, and a fourth, also cognitive, is to gain novel experience and insight into *elements of novelty*, in experimentation with *hybrids* (Holland 1975; Hannan & Freeman 1989; Powell 1991) that leads on to insight into what more wide-reaching, architectural change is needed, and how it might be tackled, to allow for the full realization of an innovation's potential.

It is only when existing dominant designs are subjected to new challenges, threatening survival, that one is *willing* to make the sacrifices of modifying or replacing proven and efficient assets for exploitation. The idea that failure to achieve objectives is an important motivational condition for change is an old one (Cyert and March 1963; March and Simon 1958). What is new here is that one need not passively wait for a new environment to arise as a new challenge but can actively choose an environment that yields interesting new challenges.

Next to motivation for change, it is only under novel conditions that one gets new *evidence* of where limitations of validity and hence priorities for change lie. At first, one will seek differentiation of dominant designs, staying close to them, in proximate or 'problemistic search' (Cyert and March 1963), to maintain exploitation as much as possible. Here, one may tap from earlier experience, going back in organizational memory to what was tried but failed in earlier exploration, at the time that current dominant designs had not yet emerged.

Next, when that does not meet challenges of survival, looking further afield by comparing one's own failures with apparent successes in newly encountered practices of others in the novel environment one can obtain hints as to how inadequate performance may be repaired by adopting elements from such local practice. And, finally, it is by experimentation with hybridization, incorporating foreign elements into one's own practice, that one has an opportunity of experimenting with novelty before surrendering the basic architecture or logic of the dominant design, and one is willing to make more drastic, architectural or fundamental change, with new elements in new architectures (Henderson & Clark 1990), because now one has a clue how to do it, in eliminating barriers to the full realization of the proven potential of novel elements or principles, currently constrained by the dominant design. It is in this last stage, of experimenting with novel architectures or basic principles of new configurations of old elements from the old dominant design and new elements from dominant designs found in the new environment, that radical innovation emerges.

Note that this process yields a path of exploitation that leads on to exploration, where one sticks as tenaciously as possible to the dominant design while yet being prepared to deviate from it, in increasingly drastic departures, in a process of discovery based on shifting practice. In this account, the shift of environment is undertaken voluntarily and by design. It may also be imposed unexpectedly from the outside, when an invading competence destroying innovation forces one to adapt.

An empirical illustration is the path of innovation or exploration that emerges, in the past mostly unintended, from multinational companies (MNC's) transferring products to new markets, where they run up against novel demands and constraints that require adaptation, and give new insights from previously unfamiliar local practice for doing things differently, from local competitors, suppliers and customers. While in the past MNC's stumbled on such invention or discovery from other motives for expanding into foreign markets, such as maintenance of growth, or escape from intense price competition

in saturated home markets, they now discover its potential for discovery, and use it as a deliberate policy for that purpose (Bartlett & Ghosal 1989; Nooteboom 2000). In culture studies also, it is a familiar principle that novel culture tends to arise at, and penetrate from, the periphery of existing culture (Lotman 1990).

In sum, this heuristic or ‘cycle of discovery’ (Nooteboom 2000) yields a path by which exploitation can yield exploration, by profiting from experience in novel contexts of application. While this goes beyond learning by making combinations of knowledge held between different agents, there are connections. First, in the stage of hybridization, in a novel context of action, part of the process is adoption of elements of practice from others. However, one first had to identify which others, and the need and opportunity for fruitful exchange was inferred from experience in the novel context. Second, learning by interaction in general, in whatever context, may include the logic of the heuristic, in that by collaborating with others one has to apply one’s knowledge in the novel context presented by the knowledge and practice of the partner. Thus, empathy, i.e. the ability to place oneself in the position of the partner is as important for learning as it is for building trust (Nooteboom 2002).

Conclusion

A coherent theory of the firm should include a number of features: an account of why organizations exist, how they coordinate activities, differences between organizations within and between industries, boundaries of the firm, collaboration between organizations, problems and solutions in the combination of exploitation and exploration, and processes of transformation and invention. With such a range of subjects, the treatment of each must be limited. Further requisite detail requires a full book, which is in hand.

The cognitive theory of the firm outlined in this paper can be seen as partly a transformation and partly an extension of Penrose’s 1959 book, with elements, also transformed, from transaction cost theory and evolutionary theory of organization. It is based on an ‘embodied cognition’ view, derived from cognitive science. This yields the notion of ‘cognitive distance and the need to limit such distance, in an organizational ‘cognitive focus’, for the sake of coordination needed to achieve a collective goal. Differences in such foci yield cognitive distance between organizations. An elaboration of organizational focus allows for an explanation of inter- and intra-industry differences between organizations, the boundaries of organizations and their foci, and the reasons, requirements and instruments for inter-organizational collaboration. The boundary of cognitive focus is determined, first of all, by a trade-off between narrow focus for exploitation and wide focus for exploration. Organizational focus also yields organizational myopia, which necessitates collaboration with outside organizations in order to repair myopia by complementary outside cognition. The analysis yields a preference, from the perspective of exploration, for alliances over mergers and acquisitions.

Concerning organizational learning, the analysis yields two kinds of ‘dynamic capability’. One is the ability to achieve novel combinations of existing knowledge by personnel turnover and by collaboration with outside partners, selecting partners at

optimal cognitive distance, and building, for that purpose, absorptive capacity and the ability of governing relational risk. Second is the ability to seek novel contexts of application as a source of novel insight and motives for change, along a path of differentiation and hybridization of knowledge and competence.

References

- Aldrich, Howard
1999 *Organizations evolving*. London: Sage.
- Anderson, Philip and Michael Tushman
1990 'Technological discontinuities and dominant designs: A cyclical model of technological change'. *Administrative Science Quarterly*, 35/4: 604-633.
- Bartlett, Christopher A. and Sumantra Goshal
1989 *Managing across borders - Transnational solution*. Boston MA: Harvard Business School Press.
- Baum, Joel A.C. and Jitendra V. Singh
1994 'Organizational hierarchies and evolutionary processes: Some reflections on a theory of organizational evolution' in *Evolutionary dynamics of organizations*, J.A.C. Baum and J.V. Singh (Eds.), 3-20. Oxford: Oxford University Press:
- Bazerman, Max H.
1998 *Judgement in managerial decision making*. New York: Wiley.
- Berger, Peter L. and Thomas Luckmann
1967 *The social construction of reality*. New York: Doubleday.
- Blackler, Frank
1995 'Knowledge, knowledge work and organizations: An overview and interpretation'. *Organization Studies*, 16: 1021-1046.
- Bleeke, Joel and David Ernst
1991 'The way to win in cross-border alliances', *Harvard Business Review*, November/December: 127-35.
- Brown, John Seely and Paul Duguid
1996 'Organizational learning and communities of practice' in *Organizational learning*, Michael D. Cohen and L. S. Sproull (Eds.), 58-82. London: Sage.
- Campbell, Donald T.
1994 'How individual and face-to-face group selection undermine firm selection in organizational evolution' in *Evolutionary dynamics of organizations*, Joel A.C. Baum and J.V. Singh (Eds.), 23-38. Oxford: Oxford University Press.
- Choo, Chun W.
1998 *The knowing organization*. Oxford: Oxford University Press.
- Cohen, Michael D. and Dan A. Levinthal

- 1990 'Absorptive capacity: A new perspective on learning and innovation'. *Administrative Science Quarterly*, 35: 128-52.
- Cyert, Richard M. and James G. March
1963 *The behavioral theory of the firm*. Englewood Cliffs NJ: Prentice-Hall.
- Damasio, Antonio R.
1995 *Descartes' error: Emotion, reason and the human brain*. London: Picador.
- Damasio, Antonio R.
2003 *Looking for Spinoza*. Orlando: Harcourt.
- DiMaggio, Paul J. and Walter W. Powell
1983 'The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields'. *American Sociological Review*, 48/2: 147-160.
- Dreesmann, Anton C.R.
1963 *Evolution and expansion* (in Dutch). Leiden: Stenfert Kroese.
- Edelman, Gerald M.
1992 *Bright air, brilliant fire; On the matter of mind*. London: Penguin.
- Edelman, Gerald M.
1987 *Neural Darwinism; the theory of neuronal group selection*. New York: Basic Books.
- Gilsing, Victor A. and Bart Nooteboom
2006 Exploration and exploitation in biotechnology, *Research Policy*, forthcoming.
- Gilsing, Victor A. and Bart Nooteboom
2005 Density and strength of ties in innovation networks: An analysis of multimedia and biotechnology, *European Management Review*, 2: 179-197.
- Granstrand, O., Patel, P. and Keith Pavitt
1997 'Multi-technology corporations: Why they have distributed rather than distinctive core competencies'. *California Management Review*, 39/4: 8-25.
- Hannan, Michael and J. Freeman
1977 The population ecology of organizations. *American Journal of Sociology*, 88: 929-964.
- Hannan, Michael and John Freeman
1984 Structural inertia and organizational change. *American Sociological Review*, 49: 149-164.
- Hannan, Michael and John Freeman
1989 *Organizational ecology*. Cambridge MA: Harvard University Press.

- Henderson, Rebecca M. and Kim B. Clark
 1990 'Architectural innovation: The reconstruction of existing product technologies and the failure of established firms'. *Administrative Science Quarterly*, 35: 9- 30.
- Hodgson, Geoffrey M.
 2002a 'The legal nature of the firm and the myth of the firm-market hybrid', *International Journal of the Economics of Business*, 9/1: 37-60.
- Hodgson, Geoffrey M
 2002b 'Darwinism in economics: from analogy to ontology'. *Journal of Evolutionary Economics* ,12: 259-281.
- Hodgson, Geoffrey M
 1993 *Economics and evolution: Bringing life back into economics*. Cambridge: Polity Press.
- Hodgson, Geoffrey M. and Thorbjørn Knudsen
 2005 'Why we need a generalized Darwinism, and why generalized Darwinism is not enough'. *Journal of Economic Behavior and Organization*, forthcoming.
- Hodgson, Geoffrey M. and Thorbjørn Knudsen
 2004 'The firm as interactor'. *Journal of Evolutionary Economics*, 14: 281-307.
- Holland, John H.
 1975 *Adaptation in natural and artificial systems*. Ann Arbor, MI: University of Michigan Press.
- Johnson-Laird, Philip N.
 1983 *Mental models*. Cambridge: Cambridge University Press
- Kolb, David
 1984 *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.
- Lakoff, George and Mark Johnson
 1999 *Philosophy in the flesh*. New York: Basic Books.
- Lakoff, George and Mark Johnson
 1980 *Metaphors we live by*. Chicago: University of Chicago Press.
- Levitt, Barbara and James G. March
 1988 Organizational learning, *Annual Review of Sociology*, 14: 319-340.
- Lotman, Yury
 1990 *Universe of the Mind; A semiotic theory of culture*. London: I. B. Tauris.
- March, James G.

- 1991 'Exploration and exploitation in organizational learning'. *Organization Science*, 2/1: 101-123.
- March, James G. and Herbert A. Simon
1958 *Organization*. New York: Wiley.
- McKelvey, William
1982 *Organizational systematics: Taxonomy, evolution, classification*. Berkeley: University of California Press.
- Mead, George H.
1934 *Mind, self and society; from the standpoint of a social behaviorist*. Chicago: Chicago University Press.
- Mead, George H.
1982 *The individual and the social self*. Unpublished work of G.H. Mead, D.L. Miller (Ed.). Chicago: University of Chicago Press.
- Mintzberg, Henry
1983 *Structure in fives: Designing effective organizations*. Englewood Cliffs NJ: Prentice-Hall.
- Nelson Richard R. and Sidney Winter
1982 *An evolutionary theory of economic change*. Cambridge: Cambridge University Press.
- Nonaka, Ikujiro and Hirotaka Takeuchi
1995 *The knowledge creating company*, Oxford: Oxford University Press.
- Nooteboom, Bart
2002 *Trust: Forms, foundations, functions, failures and figures*. Cheltenham UK: Edward Elgar.
- Nooteboom, Bart
2000 *Learning and innovation in organizations and economies*. Oxford: Oxford University Press.
- Nooteboom, Bart
1999 *Inter-firm alliances: Analysis and design*. London: Routledge.
- Nooteboom, Bart
1992 'Towards a dynamic theory of transactions'. *Journal of Evolutionary Economics*, 2: 281-299.
- Nooteboom, Bart, Wim P.M. van Haverbeke, Geert M. Duysters, Victor A. Gilsing and Ad van den Oord
2005 *Optimal cognitive distance and absorptive capacity*, under review.

- Penrose, Edith
1959 *The theory of the growth of the firm*, New York: Wiley
- Peteraf, Margaret, Christos Pitelis and Maurizio Zollo
2005 Call for papers special issue on 'Metamorphosis of (theory of) the firm'.
Organization Studies, 26/4: 655-656.
- Pitelis, C. (Ed.)
2001 *The growth of the firm, the legacy of Edith Penrose*. Oxford: Oxford University Press.
- Pettit, Philip
1995 'The virtual reality of homo economicus'. *The Monist*, 78/3: 308-329.
- Piaget, Jean
1970 *Psychologie et épistémologie*. Paris: Denoël.
- Piaget, Jean
1974 *Introduction a l'épistémologie génétique*. Paris: Presses Universitaires de France.
- Polanyi, Michael
1962 *Personal knowledge*. London: Routledge.
- Powell, William W.
1991 'Expanding the scope of institutional analysis' in *The new institutionalism in organizational analysis*, Woody W. Powell and Paul J. Dimaggio (Eds.), 183-203.
Chicago: Chicago University Press.
- Schein, Edgar H.
1985 *Organizational culture and leadership*. San Francisco: Jossey-Bass
- Shackle, George
1961 *Decision, order and time in human affair*. Cambridge: Cambridge University Press.
- Simmel, Georg
1950 *The sociology of Georg Simmel*, Translation Kurt Wolff, Glencoe Ill.: The Free Press.
- Simon, Herbert A.
1983 *Reason in human affairs*, Oxford: Basil Blackwell.
- Smircich, Linda
1983 'Organization as shared meaning' in *Organizational symbolism*, L.R. Pondy, P.J. Frost, G. Morgan and T.C. Dandridge (Eds.), 55-65. Greenwich Conn: JAI Press.
- Spender, J. C.
1989 *Industry recipes*. Oxford: Basil Blackwell.

- Teece, David, Gary Pisano and Amy Shuen
 1997 'Dynamic capabilities and strategic management', *Strategic Management Journal*, 18: 509-533.
- Tushman, M.L. and P. Anderson
 1986 'Technological discontinuities and organizational environments'. *Administrative Science Quarterly*, 31: 439-465.
- Vygotsky Lev
 1962 *Thought and language*, edited and translated by E. Hanfmann & G. Varkar. Cambridge MA: MIT Press.
- Weick, Karl F.
 1979 *The social psychology of organizing*. Reading, MA: Addison-Wesley.
- Weick, Karl F.
 1995 *Sensemaking in organisations*. Thousand Oaks CA: Sage.
- Weick, Karl F. and K.H. Roberts
 1993 'Collective mind in organizations'. *Administrative Science Quarterly*, 39, reprinted in *Organizational learning*, Michael D. Cohen and L. S. Sproull (Eds.). 1996, 330-358. London: Sage.
- Williamson. Oliver E.
 1975 *Markets and hierarchies: Analysis and ant trust implications*. New York: The Free Press.
- Williamson. Oliver E.
 1985 *The economic institutions of capitalism; Firms, markets, relational contracting*, New York: The Free Press.
- Williamson. Oliver E.
 2003 'Calculativeness, trust and economic organization', *Journal of Law and Economics*, 36, 453 - 86.
- Williamson. Oliver E.
 1999 'Strategy research: Governance and competence perspectives', *Strategic Management Journal*, 20: 1087-1108.
- Wuyts, Stefan, Massimo Colombo, Shantanu Dutta, and Bart Nooteboom
 2005 'Empirical tests of optimal cognitive distance'. *Journal of Economic Behaviour and Organization*, 58/2: 277-302.
- Zollo, Maurizio and Sidney G. Winter
 2002 'Deliberate learning and the evolution of dynamic capabilities', *Organization Science*, 13/3: 339-351.

Appendix A Similarities and differences with Penrose's theory

Table 1 Points of agreement with Penrose (1959), with page references

Basic view of the firm

- growth, development of the firm as driver (p1, 7, 29)
- theory of the firm needs to include an entrepreneurial perspective (32)
- entrepreneurship entails seeing opportunities (imaginative effort, instinct for what will work(37)), willingness (strength of character, perseverance) and ability to use them (e.g. timing, persuasion, creating confidence (38)), committing effort and resources to speculative activity (33)
- the environment of the firm is not a given, in the same form for all, and its perception and creation are central (41)
- firm resources are bundles of potential services (25), and it is the services that yield production (and a free resource can render services that are not free (78))
- firm resources are largely firm-specific (24), and their services are heterogeneous across firms (199)
- there is a distinction between entrepreneurial and managerial resources (183)
- firms have multiple activities, and growth takes place, in particular, by new activities, in diversification (70)

Incentives for diversification

- one incentive for diversification may lie in the need for a portfolio of products, in product line diversification (135)
- another may lie in the evening out of fluctuations (138)
- another may lie in the spread of risks in a portfolio (177)
- next to internal expansion, firms may expand externally, by merger and acquisition

Limits to expansion

- resources and the services they offer are to some extent an outcome of experience, and then cannot be purchased or hired in the form needed, and acquired services must be made to fit in the existing constellation of the firm (47)
 - managerial resources are taken up, in part, in guiding and training new management (47)
 - economies available only in expansion may disappear after expansion has been completed, and may then be made independent (101, 262)
 - threat of entry reduces take-over price of incumbents (167)
 - the scope for diversification is restricted by the need to invest for the sake of competing in existing activities (134) and by the need to maintain necessary integration with the rest of the firm, and avoid bureaucracy (208)
 - thus there is a crucial trade-off between speed of expansion and maintenance of control (189), 'fundamental ratio of managerial resources available for expansion
 - new firms need disproportionate attention to current operations (205)
-

Table 2 Points of difference with respect to Penrose

Topic Penrose (with page references)	My cognitive theory of the firm
<i>Basic view of the firm</i>	
<ul style="list-style-type: none"> - basic driver: growth & long run profits (29) - firm as administrative org. (15, 24) - boundary of the firm as area of administrative coordination and authoritative communication (intro, 20) - key resource is managerial - central issue: planning and administration - structure is the creation of men who run the firm (31) - central challenge: combine administration of existing activities and expansion by diversification - large firms have absolute comparative advantage in all respects (225), no clear diseconomies of scale 	<p>underlying that is will to expression, creation, power</p> <p>firm as focusing device</p> <p>boundary as trade-off between variety and cohesiveness</p> <p>is cognitive and communicative human resources, conditions for cognition, action, collaboration, motivation</p> <p>emergent from internal and external interaction</p> <p>combining exploitation and exploration</p> <p>diseconomies of scale in lack of focus or problem in cognitive coordination</p>
<i>Incentives for diversification</i>	
<ul style="list-style-type: none"> - increase of knowledge = increase of range of services from a resource (76) - means of expansion: internal investment, merger/acquisition - ratio needed for administrative work need not increase, due to increased division of labour and substitution of capital for labour (202) - disadvantages of small firms (225, 238) 	<p>also development of new resources</p> <p>also by external collaboration in alliances and networks</p> <p>possible relations with n people is $n(n-1)/2$, yields noise, requires bureaucratization, reduced flexibility</p> <p>compensation of such weaknesses in networks</p>
<i>Mergers and acquisitions</i>	
<ul style="list-style-type: none"> - apparently few problems in mergers and acquisitions; acquisition reduces managerial effort needed for expansion (209) - acquisition easier the smaller relative size of acquired firm (210) 	<p>problems of integration due to difference in focus, reduced flexibility, reduced variety for innovation</p> <p>focus and value of small firm likely to be destroyed</p>