Paper to be presented at the DRUID Academy 2012

on

January 19-21
at
University of Cambridge /The Moeller Centre

Cognitive distance in research collaborations

Margherita Balconi
Università degli studi di Pavia, Pavia, Italia
Dipartimento di Economia Politica e Metodi Quantitativi
balconi@unipv.it

Valeria Lorenzi
Università degli studi di Milano Bicocca
Dipartimento di Scienze Economico Aziendali
v.lorenzi2@campus.unimib.it

Pier Paolo Saviotti
Univérsité Pierre Mendés-France, Grenoble, France
INRA GAEL
pier-paolo.saviotti@wanadoo.fr

Antonella Zucchella
Università degli studi di Pavia, Pavia, Italia
Dipartimento di Ricerche Aziendali R.Argenziano
antonella.zucchella@unipv.it

Abstract
This paper addresses the cognitive dimension of proximity/distance in research collaborations. While the theory of optimal cognitive distance assumes that learning is the motive of collaborations, we suggest a new interpretative framework, whereby specialised firms collaborate with the purpose of accessing (rather than acquiring) the different specific fields of knowledge needed to solve problems emerging during the research and innovation process. We apply it to interpret the research collaborations of a sample of Italian biotech firms. In order to measure cognitive distance, we introduce an index originally developed by ecologists to measure distance between species. Our results show that most partners have a high cognitive distance (79,4/100 on average) and that even small firms realize extensive networks of collaborations. This can be easily explained by our framework, since the loose constraint of absorptive capacity implied by knowledge accessing does not limit the number and variety of collaborators between specialized and diverse actors.

Jelcodes:O32,-

Cognitive distance in research collaborations

M. Balconi°, V.Lorenzi^, P.P. Saviotti*, A. Zucchella#

° Dipartimento di Economia Politica e Metodi Quantitativi, Università di Pavia, Italy;

address: via Ferrata 1, I27100 Italy;

tel. +39 0382 985916;

fax: +39 0382 422583

e-mail: balconi@unipv.it

^ Dipartimento di Scienze Economico Aziendali, Università di Milano Bicocca, Italy;

e-mail : valeria.lorenzi85@gmail.com

*INRA GAEL, Univérsité Pierre Mendés-France, Grenoble, France;

e-mail: pier-paolo.saviotti@wanadoo.fr

Dipartimento di Ricerche Aziendali R.Argenziano, Università di Pavia, Italy;

e-mail: antonella.zucchella@unipv.it

Abstract

This paper addresses the cognitive dimension of proximity/distance in research collaborations. While the theory of optimal cognitive distance assumes that learning is the motive of collaborations, we suggest a new interpretative framework, whereby specialised firms collaborate with the purpose of accessing (rather than acquiring) the different specific fields of knowledge needed to solve problems emerging during the research and innovation process. We apply it to interpret the research collaborations of a sample of Italian biotech firms. In order to measure cognitive distance, we introduce an index originally developed by ecologists to measure distance between species. Our results show that most partners have a high cognitive distance (79,4/100 on average) and that even small firms realize extensive networks of collaborations. This can be easily explained by our framework, since the loose constraint of absorptive capacity implied by knowledge accessing does not limit the number and variety of collaborators between specialized and diverse actors.

Key words: biotech, cognitive distance, knowledge, learning.

1. Introduction

Firms engaged in complex and uncertain innovation processes depend crucially on interactions and collaborations with external actors, in order to overcome the limitations of specific knowledge bases. This paper addresses the cognitive dimension of proximity/distance in research collaborations, focusing merely on scientific and technological knowledge.

The literature on cognitive proximity is rather limited. The few extant contributions are based on the hypothesis that the main purpose pursued by collaborating firms is mutual learning (Nooteboom, 2000 and 2007; Wuyts et alii, 2005). In this perspective a trade-off exists between the novelty value of the knowledge firms seek to acquire and the difficulty of acquiring it, so that the 'optimal cognitive distance' between the interacting firms should be of a medium level. We note that if learning what others know were the general motive of collaborations, over time firms should become increasingly similar to each other, and enlarge their scope. While this might well occur in many instances, for many years the general trend of the industrial landscape has been in the opposite direction. Without downplaying the importance of learning, we merely make the obvious statement that it is more likely that firms learn in order to hone and build upon their distinctive competences, rather than to increase the number of competences. The question we address is what motivates leading firms to interact frequently and often repeatedly with other organizations, while increasing (not decreasing) their specialization and focalization.

To advance the understanding of this issue, we draw upon a stream of literature on strategic alliances. According to Mowery et alii (1996) and Grant and Baden-Fuller (2004), collaborating firms, rather than learning, might only seek to access the partners' knowledge, without attempting to assimilate, acquire and apply it autonomously for commercial ends. Based on this motivational basis, we are able to come up with a novel explanation of the degree of cognitive distance between partners and to propose an interpretation of the behaviour of small specialised knowledge intensive firms, that need to access external knowledge, without intending to enlarge their focus.

After delineating the conceptual framework, we apply it to an empirical investigation of the research collaborations of a sample of 31 Italian biotechnology firms. Since these firms are mostly small, privately-owned and have few (if any) patents, which have indeed typically a single applicant, it is difficult to identify the alliances in which they are involved and the fields of knowledge they master. However, since they publish much more than they patent, we could rely on their co-publications in order to discover the organizations with which they collaborate in research. By examining the 511 scientific papers which they co-authored with individuals belonging to other organizations, we found 1.244 research collaborations and 845 collaborating entities (other firms, hospitals, universities and research organizations). Since many of the collaborating organizations are big and have a wide scope, in order to measure cognitive distance we consider only the subunits directly involved in the collaborations examined. Obviously, the collaborations revealed by co-publications are not representative of all firms' connections. However, given the central role which the generation of new knowledge plays in knowledge intensive industries, the investigation of research collaborations is interesting per se.

To measure cognitive distance we resorted to a methodology elaborated by ecologists to measure the distance of species based on the number of traits they share as a ratio of the total potential number. Thus we do not assess the extent to which the fields of knowledge possessed by the actors are similar or different, but rather we identify the domains of specialization of the various agents and count the number of fields they have in common, each field being a trait characterizing the agents. The higher the number of fields they share, the more similar they are. Put differently, what we measure is the degree of cognitive overlap of the agents.

The paper is organised as follows. In section 2 we discuss the literature and propose a new theoretical conceptualization of cognitive distance, from which we derive the interpretative framework guiding the empirical investigation. In section 3 we explain the methodology used to measure cognitive distance. Section 4 describes the data and the classifications adopted. Section 5 illustrates the structure of collaborations among the various types of actors involved (firms, universities, hospitals, research organizations) while section 6 shows and discusses the results obtained with regard to cognitive distance. Section 7 draws some conclusions and proposes suggestions for future research.

2. Conceptual framework

In the course of research and innovation processes firms may need to reach for disciplines and skills they do not master and are only occasionally useful. Within the context of a high division of intellectual labour, this leads to collaborations with the actors commanding them (Feldman 1994). Interactions with complementary sources of knowledge are particularly critical to small specialised firms.

The literature on networks is enormous (Malerba and Breschi, 2005) and the biotech industry is widely recognised as one of the main cases of distributed innovation, where research and development take place through collaborations among organisations belonging to different scientific and business areas (Powell and Brantley, 1992; Powell et alii, 1996, Pisano, 2007, Stuart et alii, 2007). Depending on the desired balance between exploration and exploitation, two main kinds of network are found in biotechnology (March, 1991; Gilsing and Nooteboom, 2005; Rothermael and Deeds, 2004): partnerships between universities, research centres and dedicated biotechnology firms, based mostly on the exploration and creation of knowledge; and partnerships between dedicated biotechnology firms (DBF) and large pharmaceutical companies, based primarily on the exploitation of knowledge.

This literature devotes great attention to the issue of geographic proximity. There is wide consensus that geographic proximity matters especially for the communication of tacit knowledge which necessarily involves face to face contacts, while codified knowledge may also be accessed at distance (Nonaka, 1994; Nonaka and Takeuki, 1995; Storper, 1997). Face to face contacts also help create trust, while the sharing of experiences and technologies as well as the mobility of employees enhance learning processes. The new economic geography based on spillovers and agglomeration economies as well as the literature on regional innovation systems (Feldman, 2000; Cooke,1998 and 2002; Asheim and Gertler, 2005) are also extremely wide, and to give even a cursory account of them goes well beyond the scope of this article.

Much less studied are other dimensions of proximity. The concept of cognitive distance (CD) - or the converse, cognitive proximity - has been proposed by Nooteboom and defined as a way “to interpret resource heterogeneity between the firms that hold these different resources” (2007, p.1017). We define CD as the distance between the specialised knowledge bases of the actors involved in a collaboration (Nesta and Saviotti, 2005; Pyka and Saviotti, 2005).

Nooteboom (2000 and 2007) develops a theory of the “optimal” degree of CD, which is based on the assumption that *mutual learning* is the objective driving different actors to interact. In interfirm relationships learning entails a trade-off between the advantage of high cognitive distance in terms of novelty value of the partner’s knowledge, and the disadvantage of low mutual understanding (due to the problem of absorptive capacity[1]). Thus, also considering the problems linked to cognitive lock-in phenomena (Lambooy, 2003), a firm can expect a greater extent of learning from an organization having the knowledge it wants to acquire than its same knowledge. However, the costs and difficulties of communication and assimilation tend to rise with a growing cognitive distance, till mutual understanding is precluded. "If effectiveness of learning by interaction is the mathematical product of novelty value and understandability, the result is an inverse U-shaped relation with cognitive distance" (Wuyts et alii , 2005, p. 279; Nooteboom et alii, 2007; Gilsing et alii, 2008). Thus the optimal CD is the degree of CD which produces the most effective learning.

However, relationships among actors may serve other purposes than learning. Amongst various contributions on alliances and the role of firm-specific knowledge in firm strategy, Hamel (1991, p.84) pointed out that a distinction between the objective of *gaining access* to a partner's skills (e.g. relying on a partner's employees for some critical operation) and *internalizing* them still needed to be clearly drawn. Later Mowery et alii (1996) very clearly stated that firms, rather than using alliances to acquire capabilities, may use collaboration to gain access to the knowledge capabilities of partnering firms, without aiming at internalizing or acquiring them. They argued that if collaborations led to acquire the partner's technological capabilities, then over time alliances would produce increased similarity in their knowledge bases. But if collaborations led only to access partner's competences, partners would maintain, and possibly increase, their knowledge specialization. By examining 792 alliances established in the years 1985-876 and including at least one US firm[2], they find that the

---

[1] Cohen and Levinthal (1990 :128) define absorptive capacity as the sum of the abilities to recognize the value of new information, assimilate it, and apply it to commercial ends.

[2] The empirical approach used by Mowery et alii focuses on the citation patterns in a firm's patent portfolio, which allows to observe the changes in the relationship of one firm's technology portfolio to that of a partner firm occurring before and after alliances

capabilities of partner firms become more divergent in a substantial subset of alliances, which is consistent with the 'knowledge accessing' view of collaborations.

Developing this idea, Grant and Baden-Fuller (2004, p. 63) contend that 'knowledge accessing provides the predominant motive for alliance formation, especially within the knowledge-based sectors where alliance activity has been especially prevalent (e.g. pharmaceuticals, semiconductors, aerospace, telecommunications, and consumer electronics)'[3]. They argue that the trend since the early 1980s towards the emergence of increasingly focused companies 'is not obviously consistent with the firms using alliances to continually broaden their knowledge bases as they acquire their partners' knowledge' (2004, p.65). Moreover, the two approaches (respectively the *learning view* and the *knowledge accessing view* of collaborations) have differing predictions about a firm's potential for managing multiple alliances. If alliances are about acquiring knowledge, then each firm's number of alliances will be limited by its absorptive capacity. Conversely, knowledge accessing alliances would permit a firm to engage in extensive networks of alliances.

Similarly, Nielsen and Nielsen (2009) underline that alliance partners might intend to create value through combining their separate knowledge bases, without seeking to learn from each other, but rather willing to maintain their distinctive competences. If innovation is enabled by combining different knowledge specializations, we suggest that collaborations between the same specialized knowledge carriers should tend to be repeated any time a combination of their types of knowledge is required.

[3] To make clear the difference between learning and knowledge accessing, Grant and Baden-Fuller make the following examples (2004: 65) : 'Daimler-Benz's initial collaboration with Swatch in designing its "Smart car" was motivated not by Mercedes' desire to acquire Swatch's precision engineering and microdesign capabilities and Swatch's desire to acquire Daimler's automotive know-how, but by both parties' desire to create value through combining their separate knowledge bases. Similarly, Luciano Pavarotti's collaboration with the Spice Girls was for the purposes of combining their different styles and capabilities in a music album, not about Pavarotti learning how to be a girl band, or the Spice Girls acquiring operatic skills'.

In short, according to the *knowledge accessing view* the main purpose of collaborations is likely to be overcoming the cognitive limits set by specialization and division of (intellectual) labour. By teaming with actors commanding different specializations, firms avoid spending efforts to acquire new capabilities and enlarging their cognitive scope (i.e., *they avoid* the costs of learning). They only need to be able to integrate the contributions of the various actors. As put by Grant (1996, p. 115–16) 'To integrate separate areas of specialist knowledge, there needs to be some intersection of the separate knowledge sets in the form of 'common knowledge'.

According to us, this common knowledge is present if the collaborating units are specialized in subfields of the same scientific/technological area. In this case the researchers affiliated to the collaborating entities belong to the same epistemic community, even if they are specialized in different disciplines. Sharing the same language, the same fundamental scientific principles, the same system of interpretation of experiments etc. enables them to combine and integrate their different specialist competences.

This stream of literature neither refers explicitly to the concept of CD nor makes any suggestion regarding the more likely degree of CD between partners willing to access each other's knowledge. We believe that this connection is important, since it opens the way to a novel interpretation of the degrees of cognitive distance between partners. Very simply, on the basis of the knowledge accessing motive, we infer a high degree of CD between collaborating organizations or, more precisely, between the teams of researchers or technologists that directly interact. In other words, the more likely interactions should be those between teams belonging to the same epistemic community but specialized in different fields. It is the common knowledge background, the fact of belonging to the same epistemic community that also explains how the different teams find each other. They probably participate to the same meetings, conferences, events and have common personal acquaintances that facilitate their encounter and mutual trust.

Thus we have two different predictions and interpretations of the levels of CD between interacting organizations. According to the learning view, the likelihood of alliance formation is highest for firms that have medium levels of technological cognitive distance. According to the knowledge accessing view, the likelihood is highest for firms that have high levels of technological cognitive distance.

In the next sections we shall examine which of the two theoretical frameworks is more suited to interpret the data on the research interactions realized by a sample of Italian biotech firms. A crucial prerequisite to address this issue is to be able to measure cognitive distance.

3. A methodology to measure cognitive distance

Both Wuyts et alii (2005) and Gilsing et alii (2008) use the technological classes of patents to indicate the specialization of firms and measure CD on the basis of CRTA, which is the Pearson correlation index of the distribution across technological classes of the revealed technological advantages (RTA) of each firm relative to the collaborating firms. The RTA of a firm in a particular technological field is given by the firm's share in that field, relative to its overall share of patents granted to all collaborating companies. Positive values of CRTA indicate similarity of the pattern of technological specialization of firms. Gilsing et alii (2008) apply this method to measure CD between the 116 largest companies in three industries (chemical, automotive and pharmaceutical industries) that were also establishing technology based strategic alliances and consider 400 technological classes at two-digit level. They find that their hypothesis of an inverted U-shaped effect of CD on innovation performance of firms is confirmed. Wuyts et alii (2005) apply the same index to 67 among the largest firms operating in ICT industries, considering 31 technological classes related to ICT, but do not find any relation between CD and the likelihood of alliance formation.

The first observation we make is that this methodology cannot be applied to small firms specialized in one or, at most, few fields. The second and more important observation is that this methodology does not seem to be adequate to test a theory which gives a fundamental weight to the novelty value of knowledge and to absorptive capability. The correlation of the distribution across technological classes of RTA of two big size firms tells nothing about the specialization of the subunits that directly collaborate. But it is at this more micro level that the purported learning or knowledge access should take place, an insufficient absorptive capability would be an impediment to collaboration and the novelty value could be assessed. For this reason a methodology is required that allows to compare the specialization of the sub-units that are directly involved in an alliance. In addition, the knowledge space must be subdivided into fields of knowledge sufficiently disaggregated to permit to identify the specialized competences that differentiate firms operating within the same aggregated area.

Since formally the problem of measuring CD between firms (or other entities, like public research organizations or university departments) is identical to that faced by ecologists attempting to measure the distance of different species, once we substitute areas of competences for biological species traits, we decided to resort to measures developed by ecologists (Pielou, 1984). They have built a number of indexes of similarity or of their converse, dissimilarity or distance. While some of them can only be used for continuous or for highly variable data, others take into account only the presence or absence of certain characteristics. Furthermore, some measures are inappropriate for cases in which most of the data points are zero, as happens when one analyses small specialized firms or subunits of big organizations that typically possess only one or few competences (characteristics), among many potential ones.

We chose the index called Percentage Remoteness (PR), which is the complement of Ruzicka's similarity index (RI). According to Pielou (1984, pp. 43-44 and 55-57) this measure has the advantages of (i) being usable for presence, absence data and (ii) not being adversely affected by the presence of few ones and many zeros in the data. The PR measure is calculated by first computing Ruzicka's similarity index and then its complement to 100. To calculate Ruzicka's similarity index we need to compute the minimum (MIN) and maximum (MAX) for each component of the technology vectors representing the knowledge bases of the collaborating partners (Fig.1 and equations 1 and 2).

Figure 1. Example of steps in the calculation of Ružička's similarity index (RI) and of percentage remoteness (PR)

	KB_1	KB_2	MIN	MAX
T_1	0	1	0	1
T_2	1	0	0	1
T_3	0	0	0	0
T_4	0	0	0	0
T_5	1	1	1	1
			ΣMIN = 1	ΣMAX = 3

In the examples of figure 1, the technology vectors representing the knowledge bases of two firms, KB_1 and KB_2, contain five component technologies (T_1-T_5). In the KB vectors the number one indicates the presence of a technology in the KB of the firm and zero its absence.

Equation (1) is Ružička's index of similarity RI.

$$RI = 100 \times \frac{\sum_{i=1}^{s} \min(x_{i1}, x_{i2})}{\sum_{i=1}^{s} \max(x_{i1}, x_{i2})} \qquad (1)$$

Equation (2) is the calculation of PR, percentage remoteness.

$$PR = 100 - RI \qquad (2)$$

In this paper we shall use PR to measure cognitive distance among pairs of collaborating entities. Obviously, in order to apply it, we need to know the vector of the fields of competence that characterize each interacting actor.

4. Research design and data

To study cognitive distance in research collaborations, we develop an exploratory analysis of 1.244 collaborations realized by a sample of 31 Italian biotech firms with other organizations during the period 1992-2008. The data we use to trace collaborations are the scientific publications co-authored by these firms and the other entities. Even though we are aware that what we find is only a part of the overall innovation network involving biotech firms, we think joint publications reveal an important part of the network directly linked with the R&D activity.

Identifying the main cognitive specializations, or fields of competence, of the various actors has been particularly demanding, since it required a thorough examination and comparison of various data sources. More precisely, we examined the Directory of Italian biotech firms (www.italianbiotech.com), the web sites of the firms, press articles, various internal documents, patent descriptions and the content of scientific articles. After excluding many cases for insufficient information, a sample of 31 firms which published at least one article was left. These 31 firms belong for the most part to the red biotech sector, with only 3 specialised in the green area; they are mainly independent small or medium sized companies, but 5 are affiliates of multinational companies; all of them have at least one patent registered at USPTO or WIPTO.

With regard to collaborators, many of them are big, such as hospitals, universities and multinational companies. However, in the collaborations only a small part of these organizations is involved (Rosenkopf and Nerkar, 2001). Starting from the name of the individual co-authors, we were able to find the relevant subunits and to identify their specific knowledge bases.

Overall, we distinguished 25 fields of competence (see their list in the appendix), so that, for each of our firms and their collaborators we could construct a vector, constituted by these 25 components, where the presence of a field of competence is denoted by a one and its absence by a zero. Given the small size and the high level of specialization of most of the firms or of their collaborators (considering only the relevant subunits of big entities), the fields of competence vectors contained few ones and many zeros. The nature of the data we used is important, because it constrained the measure of distance we could use.

5. The structure of collaborations

F published 511 articles with 845 collaborating institutions (table I). Since various firms have the same collaborators, the number of relationships (links) realised by F (898) is higher than the total number of partners. The number of collaborations, where a collaboration is a co-publication of an F firm with any co-author, is even higher (1.244), since some relations are repeated (1,4 times on average).

Table I. Main data on the network of collaborations

Total # of articles published by F	511
Total # of collaborators C (# of nodes)	845
Total # of relationships (# of links)	898
Total # of collaborations (# of links*value of each link)	1.244
# of articles per firm : Average	16,3
Median	8
Modal value	2
Min. Value	1
Max. Value	59
# of collaborators (C) per firm : Average	29
Median	18
Modal value	3
Min. Value	1
Max. Value	148
Value of links : Average	1,4
Median	2
Modal value	2
Min. Value	2
Max. Value	31

The collaborating institutions (C from now on) are different kinds of organisations worldwide (34 countries in total): universities, hospitals, research institutions (including science parks, non-profit organizations, government laboratories) and firms, mainly of big or medium size, to the subunits of which we refer in the paper (as said above).

The relative importance of the above mentioned organizations as co-publishing partners of F does not vary significantly if we consider their number, or rather the number of relationships or of collaborations (table II and fig.2). Universities are always the most important partner (with a share of about 44%), followed by hospitals (about 33%). The weight of firms and research organization is much lower and rather similar, ranging between 10% and 14%[4].

Interestingly, the number of collaborators per firm is quite high (the median value is 18). For small firms such an extensive network of alliances is consistent with the knowledge accessing view (Grant and Baden-Fuller, 2004), especially considering that the degree of

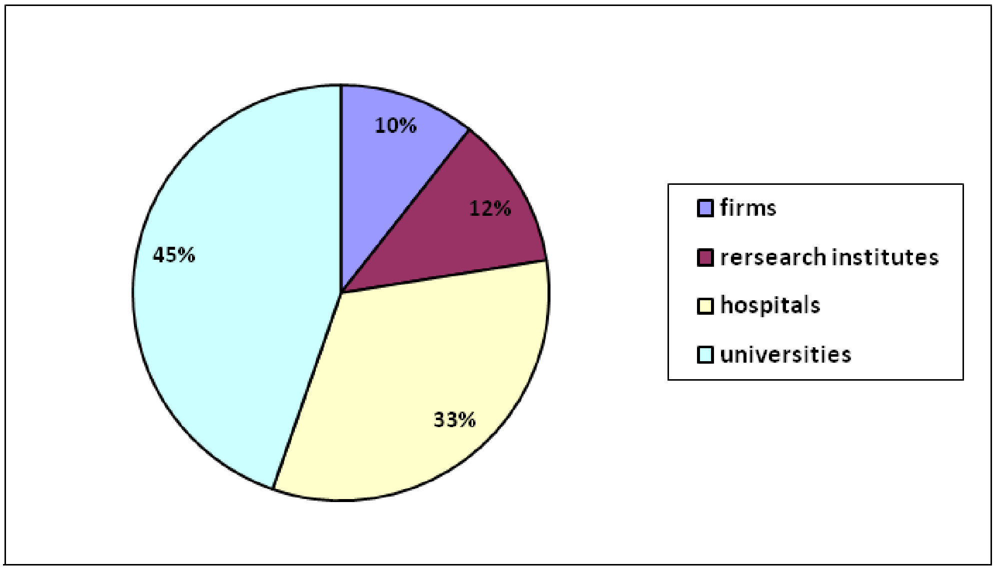
[4] The most important partners are University Federico II of Naples, University Statale of Milan and University La Sapienza of Rome, CNR among research institutes and S.Raffaele of Milan among hospitals.

knowledge overlapping with the various collaborators is on average very low (as will be illustrated in the next section). The fact that small firms seek to learn so many new disciplines does not seem to make much sense.

Table II. Weight of the various types of institutions collaborating with the focal firms

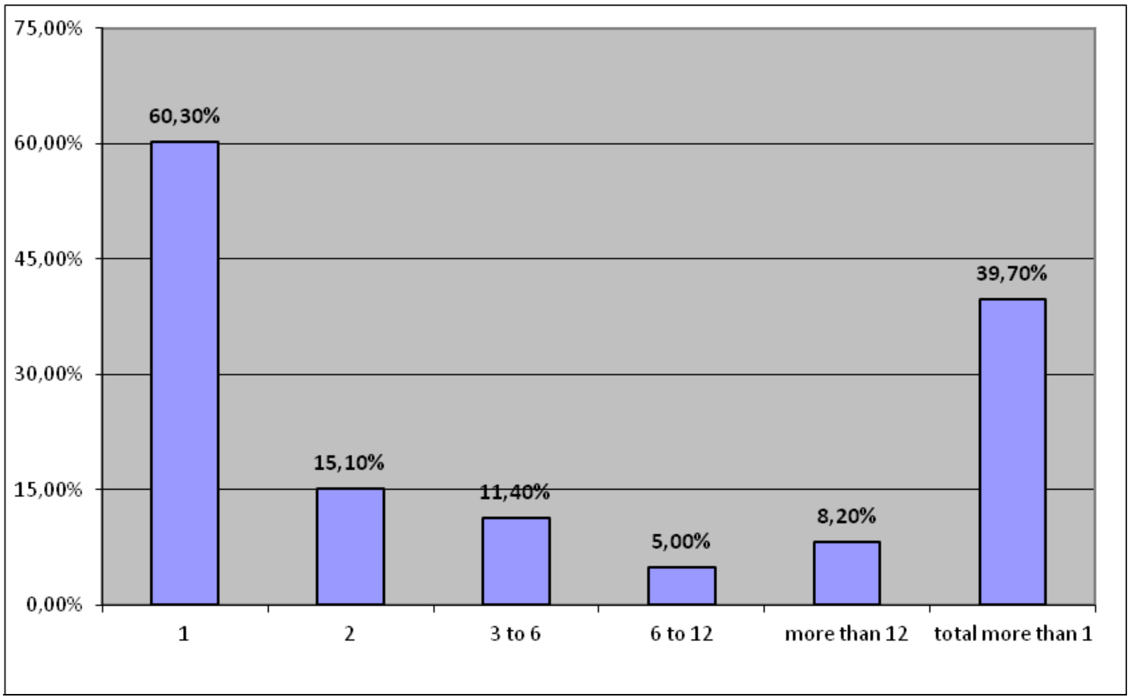
Types of collaborators	# collaborators		# relationships		# collaborations	
	N.	%	N.	%	N.	%
Firms	92	10,9%	95	10,6%	126	10,1%
Research Institutes	97	11,5%	107	11,9%	179	14,4%
Hospitals	281	33,3%	295	32,9%	396	31,8%
Universities	375	44,4%	401	44,7%	543	43,6%
Total	**845**	100,0%	**898**	100,0%	**1.244**	100,0%

Figure 2. Weight of the various types of collaborating institutions (shares of relationships)



Overall, 39,7% of collaborations are repeated (fig. 3) and the share of “strong ties”, that is of collaborations repeated more than 12 times, is relevant (20% of repeated collaborations). This seems to confirm that specialised actors repeatedly need to access complementary knowledge carriers in those areas where a combination of given knowledge fields is required in order to solve problems in the innovation process.

Figure 3. Distribution of 1244 collaborations according to the number of times they are repeated



With regard to the geographic distribution of collaborations[5], 76% are with partners located within Italy (table III). In more detail, 32,3% of collaborations are established within the same Italian region and 6,3% within the same macro area, while 37,5% with partners located in the rest of Italy. Outside Italy, European partners have a slightly higher share than partners located in the rest of the world (ROW). Within ROW, 77% of collaborations occur with partners located in the United States.

Table III. Distribution of collaborations by geographic distance and number of times they are repeated

Geographic distance	**Distribution of collaborations by the number of times they are repeated**									
	1		2 to 6 times		> than 6 times		Total relationships (links)		Total collaborations (links* value of each link)	
	N.	%	N.	%	N.	%	N.	%	N.	%
A) Same italian region	191	25,5	42	30,9	8	66,7	241	26,8	401	32,2
B) Same Italian macroarea, but outside the region	43	5,7	12	8,8	0	0,0	55	6,1	78	6,3
A+B=C	**234**	**31,2**	**54**	**39,7**	**8**	**66,7**	**296**	**33,0**	**479**	**38,5**
Rest of Italy	279	37,2	57	41,9	4	33,3	340	37,9	466	37,5
Total Italy	**513**	**68,4**	**111**	**81,6**	**12**	**100**	**636**	**70,8**	**945**	**76,0**
Europe	122	16,3	17	12,5	0	0	139	15,5	163	13,1
Rest of the World	115	15,3	8	5,9	0	0	123	13,7	136	10,9
Total	**750**	**100%**	**136**	**100%**	**12**	**100%**	**898**	**100%**	**1.244**	**100%**

Thus, while regional embeddedness does not limit the search for a research partner, still the fact that collaborations with Italian partners account for three forth of the total seems to

[5] With regard specifically to biotechnology, a comprehensive review of the literature on the geography of knowledge interactions is provided by Gertler and Levitte (2005). They report a growing appreciation that also non-local (national and global) linkages are essential to successful innovation, even taking for granted the importance of proximity. Moreover, their empirical investigation shows that knowledge networks of Canadian innovative biotech firms are to a great extent global, notwithstanding the value of local networks and specific forms of embedding.

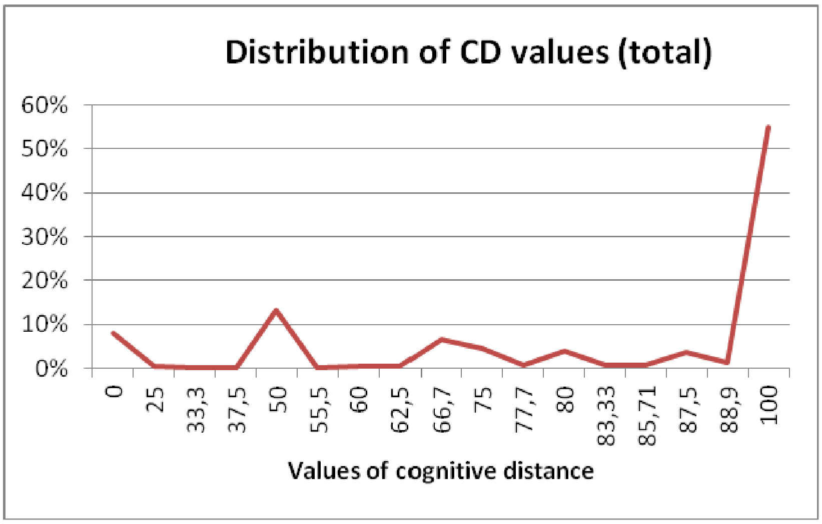
indicate that geographical distance and cultural proximity are important. This appears much more clearly when repeated partnerships are considered, since the average frequency of collaboration rises when the geographic distance of the partners falls.

The existence of partnerships with entities located in USA, Japan, Canada or Australia suggests that another crucial factor inducing collaborations is likely to be the distance with respect to the technological frontier of the time. This is consistent with the argument put forth by Moodysson and Jonsson (2007), namely that the convenience of local collaborations can never replace "the extreme requirements of cutting-edge specialized knowledge", that force DBFs to seek partners in the global arena. In fact in many fields of biotechnology and medical research the frontier is located in the USA (Dosi, Llerena, Sylos Labini, 2006) with other important organizations being located in Canada or Australia. Thus Italian biotechnology firms will opt for local knowledge whenever that is available, but will go anywhere to obtain knowledge which is scarce or unavailable locally. Of course local and international collaboration are not equivalent. The local ones may be aimed at solving recurrent problems which need continuous consultation, as shown by the very high contribution of local partnerships to repeated co-publications. On the other hand, the more expensive collaborations with a very distant partner will be used to acquire very scarce but very important knowledge. Similarly, the attractiveness of particular, 'catalyst', institutions (Aygodan and Lyon's, 2004) could also explain the collaborations with Italian universities and research institutes located outside the same region or macro area. Summarizing, direct and continuous interactions are easier in geographic proximity, but when locally unavailable knowledge becomes crucial, it does not matter how far the partner is located. Thus, even though geographic and cultural distance are likely to be barriers to collaboration, representing a “cost”, they can be compensated by the benefits arising from collaborating with particularly interesting partners.

5. Cognitive distance of collaborating partners: main results

In the great majority of cases, firms and their co-publishing partners have a very high cognitive distance (CD). The mean CD is 79,4, measured on a scale 0-100, while both the mode and the median are 100, meaning that co-publishing partners do not share any specialised cognitive field (fig.4).

Fig.4 Distribution of values of cognitive distance (898 cases)



These values are consistent with *the knowledge accessing view*, which predicts that the likelihood of collaborations is highest between actors with high levels of technological cognitive distance, but not with *the learning view*, that predicts medium levels of technological cognitive distance.

Clearly our results could have been affected by (i) the method used to measure CD, (ii) the way in which the cognitive fields constituting the knowledge bases of co-publishing partners are classified, (iii) the fact that the expected CD for co-publications is not necessarily the same as for other types of collaboration, which may have also different purposes.

Although we cannot be certain that the CDs we measure are the 'true' ones, we can still expect that the high values we generally find are not an artefact of our method: a simple visual inspection of the data matrix displaying the competences possessed by all the co-publishing partners shows that, in the vast majority of cases, they don't have any competence in common. Thus, we consider the result obtained a realistic representation of the studied sample. Moreover, it makes sense that collaborations that yield papers are done with partners endowed with complementary specializations, and that each party is not driven by the purpose of learning each other's competence, but rather by the usefulness of accessing it.

With regard to the system used to classify fields of competence, any classification system is by definition hierarchical, since within each field of competence we can usually identify several subfields at a lower level of aggregation. Cognitive distances and costs of communicating specific knowledge depend strongly on the level of aggregation used. Real cognitive distances within a group of technological fields at a given level of aggregation (intra-group distances) should be generally smaller than the distances between two groups of technological fields at a higher level of aggregation (inter-group distances). For example, if two potential partners having competencies in biotechnology and in electronics attempt to collaborate they are likely to face much higher barriers than two partners having competencies in two different classes of biotechnology. In the first case researchers belong to the same epistemic community, so that they have the common knowledge that makes communication and integration of competences easy, while in the second case they do not. We can observe that all competencies included in our sample are medical ones, except one "green" competence, sharing a non negligible part of concepts and theories. Furthermore, most of the co-publishing firms in our sample are highly specialized and their KB contains a very small number of competencies. Even in the case of large or very large co-publishing organizations - such as universities or hospitals - the collaboration occurs with a very small subset of the organization (department, laboratory, unit, etc.) having very specialized competencies. Thus, the very high cognitive distances we observe depend on the relatively low level of aggregation we have used. Our co-publishing partners can share a lot of knowledge even if their specialization is different.

We could say that, the lower the level of aggregation at which we measure cognitive distance, the more *local* this measure is, in the sense that it indicates the relative values of the cognitive distances *across* a group of fields of knowledge at a low level of aggregation. If we wanted to find an absolute measure of cognitive distance encompassing all levels of aggregation, we would need to calibrate it with respect to the maximum possible cognitive distance between any pair of cognitive fields or subsets of knowledge. Such a measure is for the moment impossible to carry out. The local measure of cognitive distance we propose is still useful since many technological alliances occur by combination of different but not too different fields of specialization.

The third factor potentially affecting the CD values is the type of collaboration. The collaborations we are examining do not represent the full network of partnerships established by our sample of firms. It comprises only those relations among actors who conduct research together, and are thus co-authors of the resulting publication. Co-publications are but one codified result of searching and problem solving previous to the realization of a project involving some marketable outcome - a by-product of a preliminary exploratory phase - and we can expect the average cognitive distance involved in co-publishing to be different than for the joint creation of a new drug or a new plant variety. While testing such hypothesis is beyond the scope of this paper, this is an interesting topic for future research.

In summary, the dominance of large CDs in our sample of co-publications is likely to reflect the high degree of specialization of co-publishing partners and the high degree of 'local' differentiation of their knowledge, which is compatible with a very large extent of shared knowledge which allows them to communicate across the cognitive distance observed.

Coming back to the analysis of the data, we also observe that CDs vary only very slightly with the frequency of co-publications (fig.5-6).

Fig.5 Distribution of CDs values in repeated collaborations (148 cases)

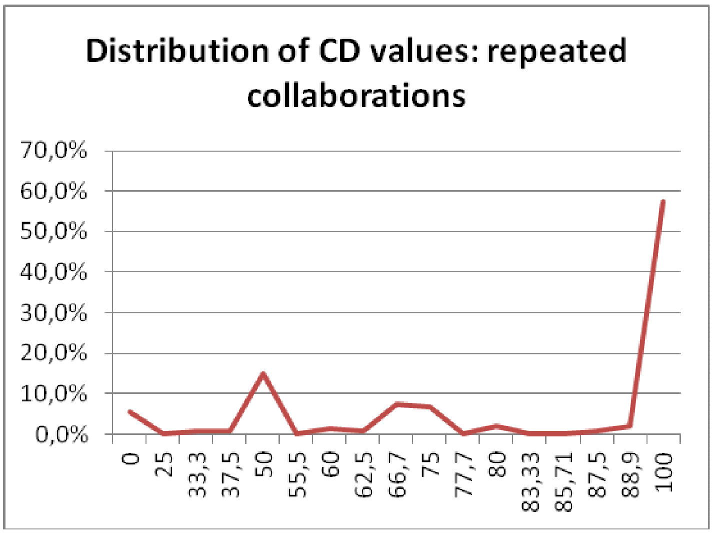
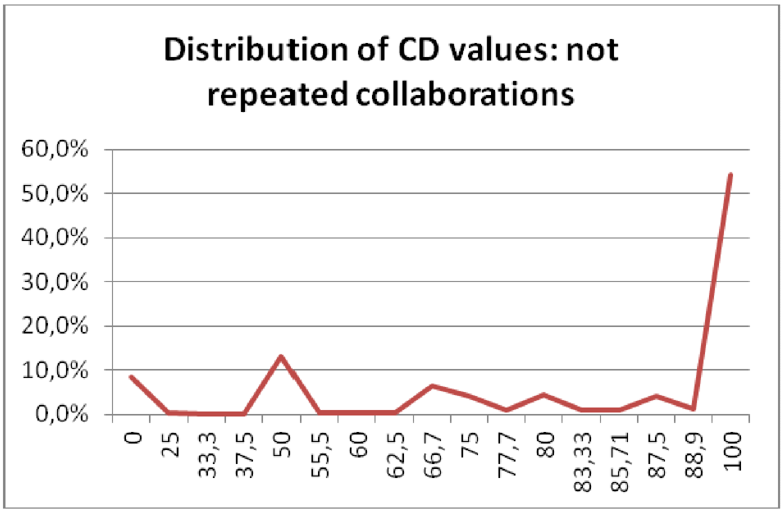


Fig.6 Distribution of CDs values in not repeated collaborations



Tab IV. Average cognitive distance between focal firms and different kinds of collaborators

	Average CD
Universities	81,37
Hospitals	79,93
Research Insititutes	74, 84
Firms	74,36
Not repeated collaborations	79,14
Repeated collaborations	80,61

Notes: *The only differences statistically significant are those between universities and research institutes, and universities and firms (t-test: null hypothesis of equality of means rejected at 5%). **The mode and the median values are always 100.

Only by type of partner (tab.IV) does one find some difference of observed cognitive distances. The average CDs is highest for universities, followed by hospitals and it is lowest for firms and research institutes. But the differences are too small to suggest some meaning.

A final point we want to make regards the relationship between cognitive and geographic proximity. Boschma (2005) suggests that the communication difficulties due to distance in space could be surmounted by proximate cognitive capabilities between collaborating actors. More explicitly, Freel (2003) states that cognitive and spatial proximity have an inverse relationship. According to this Author, if the requisite knowledge is cognitively distant from the firm's internal knowledge base, spatial proximity becomes important in assisting effective knowledge transfer, while in the presence of cognitively proximate knowledge spatial proximity is less important.

Table V. Geographic and cognitive distance between focal firms and the various types of collaborators

Geographic location of partners	**N. Partners**	**Cognitive distance**		
		Low (0-33)	Medium (33-66)	High (66-100)
Firms				
Total Italy	63	64,71%	63,16%	48,78%
Europe	29	11,76%	21,05%	28,05%
Rest of the World	26	23,53%	15,79%	23,17%
Research institutes				
Total Italy	174	71,43%	88,29%	75,96%
Europe	24	17,86%	4,50%	15,03%
Rest of the World	14	10,71%	7,21%	9,02%
Hospitals				
Total Italy	396	93,75%	51,85%	85,80%
Europe	65	0,00%	33,33%	8,88%
Rest of the World	44	6,25%	14,81%	5,33%
Universities				
Total Italy	312	64,00%	60,81%	80,97%
Europe	43	24,00%	13,51%	8,71%
Rest of the World	54	12,00%	25,68%	10,32%

In contrast with this literature, we find no inverse relationship between geographic and cognitive distance (table V). This can be easily explained based on our previous discussion. In particular, Italian firms need to search partners globally in order to access the specialized knowledge they lack and is not present nearby. Therefore the types of knowledge found locally are not substitutes for those found in distant places. Even if communication at distance is more costly, there is no local, cheaper source to resort to. In conclusion, the need to reach cutting-edge knowledge, which is different in terms of specialization but understandable by actors possessing the same background knowledge, loosens also the constraint of geographical proximity.

7. Conclusions

In this paper we introduced a novel approach to understand the degree of cognitive distance between organizations collaborating in research. Instead of assuming mutual learning as the main motive of collaborations, we argued that knowledge accessing seems to be a motivation more consistent with an industrial landscape characterised by firm specialization and focalization. In other words, specialised firms collaborate with other actors with the purpose of accessing (rather than acquiring) the different specific fields of knowledge which are needed to solve the problems emerging during the research and innovation process. Moreover, since high levels of specialization mean that a general field of knowledge/competence is divided into many different areas – the areas of specialisation - in order to capture and measure the degrees of cognitive distance which really matter it is necessary to consider these very disaggregated fields. Aggregated classifications of fields of knowledge might be misleading, since firms specialised in different knowledge niches might seem identical.

In addition, it is very unlikely that collaborations involving a knowledge exchange involve entire organizations. Especially if big organizations are considered, be they universities, hospitals or firms, in order to understand the logic of the exchange it is necessary to focus only on the subunits directly involved, bringing to the surface the competences of the team of researchers which are the real subjects of the interaction.

We argued that knowledge accessing instead of knowledge assimilating involves a lower need of absorptive capacity, so that the level of CD among collaborating entities is likely to be higher than predicted by the theory of optimal cognitive distance. In particular this holds if the interacting knowledge carriers are members of the same epistemic community, endowed with different specialist competences, but sharing the same language, basic scientific principles and paradigms, so that they can understand each other quite well. Put differently, what is required for a successful combination of the different specialised competences is that the actors involved share the common background knowledge that underlies their different specializations.

A crucial problem we had to face has been to come up with a metrics useful to measure cognitive distance at a very disaggregated level. We borrowed a measured developed by ecologists to compare different species on the basis of the number of traits they have in common, as a percentage of the total potential number of traits. In our case the traits are the knowledge fields possessed by the interacting actors.

The results of our calculations show that most co-publishing partners have a high cognitive distance, the average for the whole set being 79,4 out of a maximum of 100. This finding is clearly consistent with the knowledge accessing motive. Moreover, since accessing knowledge is much less demanding than learning it, even small firms may be able to engage in an extensive network of collaborations. The (looser) constraint of absorptive capacity does not limit the number and variety of collaborators, as was found in our investigation. Firms thus 'can reconcile the benefits of knowledge specialization with those of flexible integration.' (Grant and Baden-Fuller 2004: 62).

However, we acknowledge that the microcosm we investigated is rather narrow. Even though the picture it provides is consistent with well known larger trends, clearly an investigation focusing on bigger firms could provide a quite different picture. In particular, the learning motive might be overwhelming, since in order to enlarge the scope of their activities big firms might intend to assimilate the knowledge of partners. This is an issue open for future investigations. Moreover, since the measure of cognitive distances we propose in this paper is not the only possible one, other measures should be tested and compared to the one we used. Finally, the results obtained for co-publications should be compared to those obtained for different types of technological collaborations, for example those aimed at the joint creation of a new drug.

Acknowledgments

The authors acknowledge the financial support of the Italian Ministry for Education, Universities and Research (FIRB, Project RISC - RBNE039XKA: 'Research and entrepreneurship in the knowledge-based economy: the effects on the competitiveness of Italy in the European Union') and of EC (FP6) Project, NEST-2006-PATH-Cul, CID - Cultural and Innovation Dynamics, Contract n.: FP6 – 043345). We are grateful to Roger Coronini and Giada Palamara for their help in identifying and classifying the co-publications of our Italian biotechnology firms.

Appendix

List of knowledge fields

Cardiology, Dermatology, Diabetes, Endocrinology & Metabolism, Epathology, Gastroenterology, Hemathology, Hereditary diseases, Immunology, Nephrology, Neuroscience/neurology, Oncology, Ophtalmolic diseases, Osteo-articular diseases, Otorhinolaryngology, Pain, Respiratory diseases, Rheumatology, Skeletal muscle diseases, Urology/ gynecology , Vaccines, Veterinary science/green, Virology/infectious diseases, Diagnostic supportive applications, ICT Supportive applications.

References

Asheim, B. and Gertler, M.S. (2005). 'The Geography of Innovation: Regional Innovation Systems' in Fagerberg J., Mowery, D.C., Nelson, R.R. (Eds). *The Oxford Handbook of Innovation*. Oxford: Oxford University Press.

Audretsch, D. and Stephan, P. (1996). 'Company-scientists locational links: the case of biotechnology'. *American Economic Review*, **86**, 641-652.

Blossom Associati (2008). *Biotecnologie in Italia 2008. Analisi strategica e finanziaria*. Milano: Blossom Associati.

Boschma, R. A. (2005). 'Proximity and innovation: a critical assessment'. *Regional Studies* **39**, 61–74.

Breschi, S., and Malerba, F. (Eds) (2005). *Clusters, Networks and Innovation*. Oxford: Oxford University Press.

Cohen, W. M. and Levinthal, D. A. (1990). 'Absorptive capacity: a new perspective on learning and innovation'. *Administrative Science Quarterly*, **35**, 128–52.

Cooke, P. (1998). 'Introductions: origins of the concept'. In Braczly, H., Cooke, P. and Heidenreich, M. (Eds.) *Regional Innovation Systems*. London: UCL Press, 2-25.

Cooke, P. (2002). 'Biotechnology clusters as regional, sectoral innovation systems'. *International Regional Science Review,* **25**, 8-37.

Dosi, G., Llerena, P. and Sylos Labini, M. (2006). 'The relationships between science, technologies and their industrial exploitation: An illustration through the myths and realities of the so-called European Paradox'. *Research Policy,* **35**, 1450-64.

Feldman, M. (1994). 'Knowledge complementarity and innovation'. *Small Business Economics,* **6**, 363-372.

Feldman, M. (2000). 'Location and innovation: the new economic geography of innovation, spillovers and agglomeration'. In Clark, G., Feldman, M. P. and Gertler, M. (Eds). *Oxford handbook of economic geography*. Oxford: Oxford University Press, 373-394.

Freel, M.S. (2003). 'Sectoral patterns of small firm innovation, networking and proximity'. *Research Policy,* **32**, 751–770.

Gallaud, D. and Rherrad, I. (2002). *Knowledge production and patterns of proximity.* Industrial Dynamics of the New and Old Economy - Who is embracing whom?, Copenhagen/Elsinore, DRUID Conference.

Gertler, M. S. and Levitte , Y. M. (2005). 'Local nodes in global networks: the geography of knowledge flows in biotechnology innovation'. *Industry and Innovation,* **12**, 487–507.

Gilsing, V., and Nooteboom, B. (2005). 'Density and strength of ties in innovation networks: an analysis of multimedia and biotechnology'. *European Management Review,* **2**, 179-197.

Gilsing, V., Nooteboom, B., Vanhaverbeke, W., Duysters, G. and van den Oord A. (2008). 'Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density'. *Research Policy,* **37**, 1717–31.

Grant, R. M. (1996). 'Prospering in dynamically-competitive environments: organisational capability as knowledge integration'. *Organisation Science*, **7**, 375–87.

Grant, R.M. and Baden-Fuller, C. (2004). 'A Knowledge Accessing Theory of Strategic Alliances'. *Journal of Management Studies,* **41**, 61-82.

Hamel, G. (1991). ‘Competition for competence and inter-partner learning within international strategic alliances’. *Strategic Management Journal,* **12**, 83-103.

Jaffe, A.B. (1986). ‘Technological opportunity and spillovers of R&D: evidence from firms’ patents, profits, and market value’. *American Economic Review,* **76**, 984–1001.

Jaffe, A. B. (1989). ‘Real effects of academic research’. *American Economic Review,* **79**, 957-970.

Jaffe, A. B., Trajtenberg, M. L. and Henderson, R. (1993). ‘Geographic localisation of knowledge spillovers as evidenced by patent citations’. *Quarterly Journal of Economics,* **63**, 577-598.

Lambooy, J. G. (2003). *The role of intermediate structures and regional context for the evolution of knowledge networks and structural change*. Working Paper no. 0309, Jena, Max Planck Institute.

March, J. G. (1991). ‘Exploration and exploitation in organisational learning’. *Organization Science,* **2**, 71-86.

Moodysson, J. and Jonsson, O. (2007). ‘Knowledge collaboration and proximity. The spatial organization of biotech innovation projects’. *European Urban and Regional Studies*, **14**, 115–131.

Mowery, D.C., Oxley, J. E. and Silverman, B. S. (1996). ‘Strategic Alliances and Interfirm Knowledge Transfer’. *Strategic Management Journal,* **17**, 77-91.

Nielsen, B.B. and Nielsen, S. (2009). ‘Learning and innovation in international strategic alliances: An empirical test of the role of trust and tacitness’, *Journal of Management Studies*, **46**, 1031-1056.

Nesta, L., and Saviotti, P.P. (2005). 'Coherence of the knowledge base and the firm's innovative performance: evidence from the U.S Pharmaceutical industry'. *Journal of Industrial Economics,* **53**, 123-142.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science,* **5**, 14-37.

Nonaka, I., and Takeuki, H. (1995). *The Knowledge-Creating Company*, New York and Oxford: Oxford University Press.

Nooteboom, B. (1998). 'Innovation and Inter-firm Linkages: New implications for policy'. *Research Policy,* **28**, 793-805.

Nooteboom, B. (2000). *Learning and innovation in organizations and economies*. Oxford: Oxford University Press.

Nooteboom, B., Van Haverbeke, W. , Duysters, G., Gilsing, V. and van den Oord, A. (2007). 'Optimal cognitive distance and absorptive capacity'. *Research Policy,* **36**, 1016-1034.

Pielou, E.C. (1984). *The Interpretation of Ecological Data*, New York: Wiley.

Pisano, G.P. (2007). *Science Business.* Boston: Harvard Business School Press.

Powell, W.W., Koput, W.K. and Smith-Doerr, L. (1996). 'Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology'. *Administrative Science Quarterly,* **41**, 116-145.

Powell, W. W. and Brantley, P.(1992). 'Competitive cooperation in biotechnology: learning through networks?' In Eccles, R. (Ed), *Networks and organisations: structure, form and Action*. Boston: Harvard Business School Press.

Pyka, A., and Saviotti, P. (2005). The evolution of R&D networking in the biotech industries. *International Journal of Entrepreneur and Innovation Management,* **5**, 49-68.

Rosenkopf, L., Nerkar, A . (2001). 'Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry'. *Strategic Management Journal,* **22**, 287–306.

Rothermael, F.T., Deeds, D.L. (2004). 'Exploration and Exploitation Alliances in Biotechnology: A System of New Product Development'. *Strategic Management Journal,* **25**, 201-221.

Storper, M. (1997). *The regional world: territorial development in a global economy. perspectives on economic change*. New York: Guilford Press.

Stuart, T., Ozdemir, S. and Ding, W. (2007). 'Vertical alliance networks: The case of university–biotechnology– pharmaceutical alliance chains'. *Research Policy,* **36**, 477–498.

Vedovello, C. (1997). Science parks and university-industry interaction: geographical proximity between the agents as a driving force. *Technovation,* **17**, 491-502.

Wuyts, S., Colombo, M.G., Dutta, S. and Nooteboom, B. (2006). 'Empirical tests of optimal cognitive distance'. *Journal of Economic Behaviour and Organisation,* **58**, 277-302.

Zucker, L.G., Darby, M.R. and Brewer, M.B. (1998). Intellectual human capital and the birth of US biotechnology enterprises. *American Economic Review,* **88**, 290-306.