

Search Regimes and the Industrial Dynamics of Science

Andrea Bonaccorsi

Published online: 22 August 2008
© Springer Science+Business Media B.V. 2008

Abstract The article addresses the issue of dynamics of science, in particular of new sciences born in twentieth century and developed after the Second World War (information science, materials science, life science). The article develops the notion of search regime as an abstract characterization of dynamic patterns, based on three dimensions: the rate of growth, the degree of internal diversity of science and the associated dynamics (convergent vs. proliferating), and the nature of complementarity. The article offers a conceptual discussion for the argument that new sciences follow a different pattern than established sciences and presents preliminary evidence drawn from original data in particle physics, computer science and nanoscience.

Keywords Dynamics of science · Information science · Life science · Materials science · Reductionism · Knowledge dynamics · Diversity · Complementarity

Introduction

According to many observers, the current acceleration in scientific progress is driven by three broad scientific areas that try to explain previously unexplored or poorly understood phenomena: life, information and materials. What is interesting is that large parts of these scientific disciplines, literally, were not in existence half a century ago.

While they may originate from earlier discoveries, it is fair to say that computer or information science, life sciences based on molecular biology, and materials science did not exist at the beginning of the twentieth century, but were fully

A. Bonaccorsi (✉)
School of Engineering, University of Pisa, Via Diotisalvi 2, 56125 Pisa, Italy
e-mail: a.bonaccorsi@gmail.com

developed only after the Second World War, and received a spectacular impetus in the last three decades of the previous century. In this sense, they may be considered as *new sciences*. This is in sharp contrast to old scientific disciplines (in the modern sense) that originated after the seventeenth century scientific revolution, namely physics, astronomy, mathematics, chemistry and their subdisciplines.

This does not mean that new sciences have no relation with these scientific disciplines. In a certain sense, the distinction between new sciences and old sciences is blurred. In addition, the labels “old” and “new” may involuntarily introduce value judgments, which are clearly meaningless. Nevertheless, taking into account historical antecedents and the continuity in scientific progress, it is appropriate to keep the simple labelling and to explore this distinction.

Due to the large impact of new sciences, it is interesting to ask whether they are substantively *different* from consolidated disciplines, and whether differences matter for policy making and society at large.

In this article, we provide preliminary evidence that these new sciences indeed have characteristics that differentiate them from earlier science. In order to do so, we address some of the classical questions regarding the rate and direction of the growth of knowledge, but reformulate them, suggesting new types of measures and developing prototypes of new indicators. In Section “Are New Sciences Different?”, we do this by referring to the scientific literature and developing empirical propositions. In Sections “Rate of Growth”, “Degree of Diversity” and “Level and Type of Complementarity”, we translate the theoretical discussion from the scientific literature into concepts that are more familiar in social sciences, particularly in economics and sociology of science, and propose some operationalizations to validate, although in a crude way, the empirical propositions. We also offer preliminary evidence using these new indicators. After this empirical reconstruction, in Section “Conclusion: Search Regime Agenda”, we propose a generalization, building up the notion of changes in *search regimes* and calling for further research.

Are New Sciences Different?

Modes of Scientific Production

The proposition that scientific activity is undergoing deep changes at the end of the twentieth century is not new. The most influential reconstruction has been proposed as a transition between Mode 1 and Mode 2 science. According to Gibbons et al. (1994), we witness a shift from disciplinary, university-based or large government laboratory-based, investigator-driven type of science, to one which is multi-disciplinary, based on networks of distributed knowledge, and oriented towards problem solving and societal challenges. In the authors’ words “we are now seeing fundamental changes in the ways in which scientific, social and cultural knowledge is produced. (This) trend marks a distinct shift towards a new mode of knowledge production which is replacing or reforming established institutions, disciplines, practices and policies” (Gibbons et al. 1994, cover page).

These propositions are not without critics, however. The claim that science is moving towards more multi- or inter-disciplinarity is, from an empirical point of view, controversial (Metzger and Zare 1999; van Leeuwen and Tijssen 2000). Even in fields in which it seems a basic condition for science, such as nanoscience, a closer look reveals that disciplinary bases of knowledge are not at all eroded (Meyer and Persson 1998; Schummer 2004; Rafols and Meyer 2006). A more accurate representation is the one, in which there is no transition, but rather “Mode 1 and Mode 2 knowledge production are coexisting, coevolving and interconnected” (Llerena and Mayer-Krahmer 2003, p. 74).

One of the difficulties from these streams of literature is that they emphasize changes that take place *outside* science, that is, in the institutional, political, financial and social environment surrounding science. Even the claim that there is intrinsic tendency towards increasing inter-disciplinarity is predicated more on grounds of societal demand than of internal logic of modern science. This makes many arguments fragile. In fact, several authors have correctly pointed out that there are historical examples of Mode 2 types of knowledge even in nineteenth century science (Pestre 1997, 2003), and that in the last part of the twentieth century, there is no transition but rather “a shift of the balance between the already existing forms of Mode 1 and Mode 2” (Martin 2003, p. 13).

It is therefore important to examine the differences between new and old sciences in greater detail, and try to draw implications for policy making and society. In doing so, we will mainly use materials developed in the scientific literature, such as the debate on impossibility and the limits of science, the difference between particle physics and matter physics as well as the explanation of life. In addition, we will refer to selected contributions from philosophy of science dealing with the issue of the rate and direction of growth in scientific knowledge and the problem of reductionism and complexity in modern science.

Reductionism and the Limits of Science

A useful starting point is to consider modern science as methodologically driven by reductionist search strategies. The qualification “methodologically” implies that there is no need for commitment to a particular ontology or metaphysics, insofar as scientists use reductionism in practice. In the words of Piet Hut: “Science would be inconceivable without an initial phase of reductionism. (...) Traditionally, the particular approach to reductionism in science has been one of taking things apart: of an analysis of the whole in terms of its parts. The hope, here, was that a sufficient understanding of the parts, and their relations to each other, would enable us to see exactly in what way the whole is more than the sum of the parts; how exactly it can function in ways that none of the parts, by themselves, can” (Hut 1996, p. 167).

Methodological reductionism drives the search for causal explanations in modern science. The behaviour of phenomena at larger scales is explained in terms of the behaviour of constituent elements that are located at more elementary levels of organization. In turn, this implies that scientific disciplines are linked to each other in terms of a hierarchy of explanatory power: “Taken to extremes, it would see human psychology reduced to biochemistry, biochemistry to molecular structure,

molecular structure to atomic physics, atomic physics to nuclear physics, nuclear physics to elementary particle physics, and elementary particles physics to quantum field or superstrings” (Barrow 1998, p. 70).

This explanatory strategy has been particularly prominent in life sciences, starting with the famous question *What is life*, from the physicist Schrödinger (1944). The molecular biology revolution, supported in the United States in the 1930s by the Rockefeller Foundation and fuelled after the Second World War by the discovery of the DNA structure, explicitly aimed at explaining higher level phenomena, such as cell behaviour or human diseases in terms of elementary constituents at molecular level (Lewontin 1991). As classically stated by another physicist, Richard Feynman, the single most important thing to understand life is that all things are done by atoms, and all actions done by living things can be understood in terms of movements and oscillations of atoms (Feynman 1963). This approach is dominant in neurosciences as well as in cognitive sciences. The combination of biochemical and biophysical explanations with newly invented manipulation techniques and powerful computer computation has been driving life-sciences to astonishing results in the last three decades.

Reductionism is commonly associated to a tendency of theoretical unification. If phenomena at different levels of organization can be explained by means of elementary constituents, this will *reduce* the number of different theories needed to understand reality.

We propose this is not necessarily true. If a reductionist strategy is applied to systems of increasing complexity, then the results are a unique and unprecedented combination of unified explanations on the one hand, and proliferation of specialized sub-theories on the other. The unification of explanation does not produce unified theories, but higher level theories covering a proliferation of specialized sub-theories. This dynamic has far reaching effects on the practice and policy of science.

In order to understand this counter-intuitive effect, it is useful to refer to the scientific literature that has dealt with the problem of *limits* to scientific knowledge, or the existence of *impossibility* of explanation (Casti and Karlqvist 1996; Maddox 1998; Barrow 1998).

The theoretical physicist John Barrow has proposed a useful distinction between the search for scientific *laws* and the analysis of the *outcomes* of these laws.¹ Following this distinction, it can be said that sciences born in the late twentieth century, as opposed to old sciences, follow reductionism in the attempt to use causal explanation, but then have to give an account of a variety of phenomena for which the reductionist explanation is not enough: “There may be no limit to the number of

¹ “This division of the scientific perspective into laws and outcomes helps us to appreciate why some of the disciplines of science are so different in outlook. Ask the elementary particle physicist what the world is like and they may well tell you that it is very simple—if only you look at it in the ‘right’ way. Everything is governed by a small number of fundamental forces. But ask the same question of biologists or condensed-state physicists, they will tell you that the world is very complicated, asymmetrical and haphazard. The particle physicist studies the fundamental forces with their symmetry and simplicity; by contrast, the biologist is looking at the complicated world of the asymmetrical outcomes of the laws of Nature, where broken symmetries and intricate combinations of simple ingredients are the rule” (Barrow 1998, p. 66).

different complex structures that can be generated by combinations of matter and energy. Many of the most complicated examples we know (brains, living things, computers, nervous systems) have structures which are not illuminated by the possession of a Theory of Everything. They are, of course, permitted to exist by such a theory: But they are able to display the complex behaviours they do because of the way in which their subcomponents are organized” (Barrow 1998, p. 67; emphasis added).

While these systems have been the object of investigation for a long time, it is only through late twentieth century new sciences that the search for causal explanation has been made possible.

An interesting representation of this situation is offered by Fig. 1, developed by Barrow upon initial unpublished suggestion of the mathematician David Ruelle. Here various disciplines are mapped in a space defined by the uncertainty about fundamental mathematical laws and the level of complexity of phenomena. Using this representation, classical sciences deal with simple phenomena (chemical reactions, particle physics, quantum gravity) of increasingly difficult mathematical representation, or with complex phenomena (dynamics of solar system, meteorology, turbulence), whose fundamental laws are well known. On the contrary, new sciences (living systems, climate) are located in a region in which both uncertainty and complexity are high. It is clear from the discussion of the author, that, although not present in the figure by Ruelle, computer science and materials science would also be located in the same area. Interestingly, social sciences are located in the extreme north-east region of the map.

A similar argument has been proposed by the quantum cosmologist James Hartle. He starts with a classical reductionist proposition based on the astonishing success of quantum theory in terms of explanatory power: “at a fundamental level, every prediction in science may be viewed as the prediction of a conditional probability for alternatives in quantum cosmology” (Hartle 1996, p. 136). However, computing these probabilities requires at least a coarse-grained description of the values the variables can take (alternatives) and of circumstances under which the probabilities are conditioned (conditions). Furthermore, computing these probabilities may require different periods of time, up to extremely long periods. By combining these three dimensions, Hartle is able to map scientific disciplines in increasing order of complexity, as in Table 1.

This reconstruction² confirms the point made by Barrow and Hut. New sciences extend the search for causal explanation, which is the fundamental thrust of modern science, to new and more complex phenomena. In doing so, they require “longer

² His discussion is illuminating: “...as we move through the list we are moving in the direction of the study of regularities of increasingly specific subsystems of the universe. Specific subsystems can exhibit more regularities that are implied generally by the laws of dynamics and the initial condition. The explanation of these regularities lies in the origin and evolution of the specific subsystems in question. Naturally, these regularities are more sensitive to this specific history than they are to the form of the initial condition and dynamics. This is especially clear in a science like biology. Of course, living systems conform to the laws of physics and chemistry, but their detailed form and behaviour depend much more on the frozen accidents of several billion years of evolutionary history on a particular planet moving around a particular star than they do on the details of superstring theory or the ‘no-boundary’ initial condition of the universe” (Hartle 1996, pp. 133, 134).

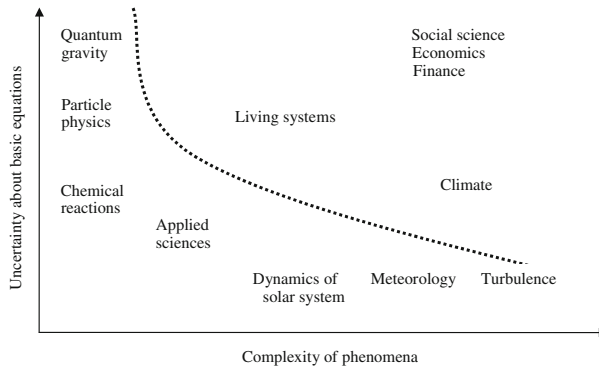


Fig. 1 A schematic representation of the degree of uncertainty that exists in the underlying mathematical equations describing various phenomena relative to the intrinsic complexity of the phenomena. Source: Barrow (1998, p. 68), after David Ruelle

Table 1 Some differences between the sciences

Scientific discipline	Length of coarse-grained		Computation of conditional probabilities
	(a) description of alternatives	(b) description of conditions	
Classical physics	Very short	Short	Very short
Astronomy	Short	Short	Short–Long
Fluid mechanics	Short–Long	Short	Short–Long
Chemistry	Short–Long	Short–Long	Long–Very long
Geology	Long	Long	Long
Biology	Long–Very long	Long–Very long	Long–Very long
Psychology	Very long	Very long	Very, very long (?)

Source: Hartle (1996, p. 135)

descriptions”, i.e. many specific sub-theories that account for “detailed form and behaviour” of the very “different complex structures that can be generated by combinations of matter and energy” or “the way in which subcomponents are organised”. In this sense new sciences share with old sciences the same methodological foundations, but generate invariably more local theories, rather than unified grand theories.

Complexity of Scientific Objects

Therefore there is no difference between old and new sciences in the adoption of methodological reductionism as the dominant search strategy. The important difference is another one. New sciences have a crucial element in common: they deal with objects (i.e. systems) that are far more complex than physical or chemical systems explained by old sciences.

Without entering into terminological discussion, by complexity, we may refer to one or more of the following aspects, largely discussed in the scientific literature: (a) complex systems have a large number of variables, parameters and feedback loops (Kline 1995); (b) complex systems have a number of interdependent hierarchical layers (Simon 1981); (c) the dynamics of complex systems is dependent on initial conditions (Ruelle 1991).

What is the overall effect of a reductionist strategy applied to complex systems? These systems are in existence because of the same laws discovered by physics and chemistry, but their behaviour requires *additional* levels of explanation, exactly because, as mentioned, “they are able to display the complex behaviours they do because of *the way in which their subcomponents are organized*” (Barrow 1998, p. 67, emphasis added). Therefore the explanation of complex systems requires information on constituent elements, but also information on architecture, or organization of elements.

Contrary to a superficial interpretation, information on architecture or organization is *not* reducible to information on elements. At each level of a complex system, causal information on physical or chemical mechanisms must be complemented by *additional* information. This idea was clearly formulated by Michael Polanyi (1962) and further developed by a few system theorists, particularly by Stephen Kline (1995). According to the former: “Lower levels do not lack a bearing on higher levels; they define the conditions of their success and account for their failures, but they cannot account for their success, for they cannot even define it” (Polanyi 1962, p. 382). This means that explanations in terms of constituent elements located at lower levels of organization are necessary but not sufficient for explanations at higher levels. They define the conditions for higher level phenomena, but are not enough to explain them.

This idea has been formalized by Kline (1995) as Polanyi’s principle, as follows: “In many hierarchically structured systems, adjacent levels mutually constrain, but not determine, each other. (...) We cannot add up (aggregate over) the principles of the lower level (...) and obtain the next higher level (...) and we cannot merely disaggregate the higher level to derive the principles of the next lower level. (...) In order to move either upward or downward in level of aggregation within a given system with interfaces of mutual constraint, we must supply added information to what is inherent in each level for the particular system of interest” (Kline 1995, pp. 115–116).

Let us give an example of how this principle has worked in one of the new sciences, life sciences. An interpretation of molecular biology largely held in life sciences has been summarized in the notion of “one gene, one enzyme”, i.e. the belief that there is an almost exclusive relation between a gene sequence and the coding of a protein. This principle implies that the working of proteins (for example their role in producing diseases) can be traced back to biochemical constituents. The so called time ‘central dogma’ of molecular biology postulated a strictly unidirectional flow of information from DNA to RNA to protein. This belief became established quite early in the evolution of molecular biology. It realizes the most ambitious reductionist programme in modern science, clearly formulated by Max Delbrück in 1935 and then Erwin Schrödinger in 1944: explaining the

phenomena of life by reducing them to the elementary level of chemistry and physics. The development of the scientific programme of molecular biology was deeply rooted in a physical and mathematical view of living matter (Kay 1993; Morange 1998). According to this perspective, it would be possible to explain phenomena at higher levels of organization (human body, organism, tissue, cell, proteins) by means of elements at a lower level of organization (genes, biochemical constituents of DNA).

Further development in life sciences, particularly with the emergence of genomics and proteomics, has revealed that the tridimensional structure of proteins is not uniquely determined by the sequence of amino acids that constitute them. In addition, we now know that whether a sequence is transcribed will depend on where it ends up in the genome, on its spatial relations to other genes, and even to other structures in the cell, or, as proposed by epigenetic research, on particular events occurring in the adult individual. Furthermore, the role of RNA is not passive but active, breaking the supposed uni-directionality of the process of transcription. Information on chemical constituents is not enough, but additional layers are needed—for example architectural information on the geometry of protein folding (Corbellini 1999; Dupré 2005), or information on the interaction between gene sequences and the environment (Lewontin 1994; Richerson and Boyd 2005; Amato 2002): more generally, information on *organization* of constituent elements. Therefore, as predicted by Polanyi's principle, information on chemical constituents of genes must be integrated with information coming from *higher levels of organization*, such as the cell or the organism, or even the environment. This information is sometimes called the symbolic order by molecular biologists (Danchin 1998).

A similar pattern emerges in materials science. Modern materials science is based on the notion that materials can be characterized by a hierarchy of space dimensions (atomic, nanometric, micrometric, millimetric, macroscopic) with strong interactions across levels (*spatial hierarchy*: Smith 1981) and by a hierarchy of time dimensions (differences in typical time scale at which phenomena occurring at various levels reach equilibrium, or get relaxed: Olson 1997). Due to this complexity, it would be impossible to derive new structures directly from desired macroscopic properties, but following the so called *reciprocity principle* (Cohen 1976) one can always decompose the macroscopic property and add architectural information, so that it becomes possible, at each level of the complex system, to determine the unitary processes needed to obtain the property at the upper level. In this sense, materials can really be “designed”, following what is called multi-scale multi-disciplinary integration.

Therefore, while life and materials sciences have been deeply influenced by a reductionist programme, what they have learned is that the study of complex systems does *not* reduce the number of theories, but calls for additional theories instead. After the complete mapping of human genome, it appeared clear that there is no simple relation between gene sequences and protein functions, but rather that the space of these mapping relations is enormously large (Lewontin 1991; Kevles and Hood 1992; Danchin 1998). What is crucial for our discussion is that at higher levels there may be many *competing* theories, each of them fully consistent with a

general theory on constituent elements. Scientists may share a common fundamental theory at the level of constituent elements (e.g. a molecular theory of life or a theory about proteins involved in cancer), but still diverge on more specific theories that link this causal explanation to the various levels of organization of life. The convergence around fundamental explanations does not produce, as in astronomy or particle physics, convergence around one or few experiments needed to validate the theory, but rather spurs dozens of more specialized theories. These theories are mutually compatible at the level of the causal explanation, but mutually incompatible (weakly or strongly) at the level of specific mechanisms across the layers of the complex system.

Due to the combination between reductionism and complexity, the progress of science does not subsume competing sub-theories under progressively more general theories, but works the other way round: each general theory at the level of constituent elements ignites the generation of many specialized sub-theories.

The Changing Boundary Between Natural and Artificial

New sciences are concerned with properties of nature at a very low level of resolution in the hierarchy of matter of the universe. This has far reaching implications on the relation between scientific explanation and technological application, or between natural and artificial.

Classical sciences are oriented towards the explanation of phenomena as they are observable in the real world. The distinction between understanding the properties of nature and making use of these properties for practical purposes is then inscribed into the constitution of modern science.

In the history of science, it is largely recognized that a number of discoveries originated from very practical problems, in engineering as well as daily life, so that the linkage between “the philosophers and the machines” (Rossi 1962), was much closer than expected (Galison 2003; Conner 2005).

At the same time, it is also true that knowledge about the artificial developed in an autonomous way, building upon scientific theories but also producing autonomous knowledge. This body of knowledge, called engineering, proceeds with its own procedures and validity criteria, and cannot be conceptualized as applied science (Kline 1985; Vincenti 1990; Bucciarelli 1994). Historically, there have been numerous occasions in which technology has actually *preceded* science, inventing working devices before a satisfactory explanation of why they worked was available (Kline and Rosenberg 1986; Rosenberg 1994).

The dynamic equilibrium between science and technology, or scientific disciplines and engineering disciplines, is currently under deep change as witnessed by a large literature. We put forward the conjecture that this change is *not* only due to societal or economic pressures, or to a trend towards multi-disciplinarity as proposed by the current literature, but to the internal dynamics of growth of knowledge.

The clear boundary between natural and artificial and between scientific explanation and engineering design, which was firmly established in classical sciences, is made more permeable in new sciences by several new factors.

First, a reduction in the time interval between prediction, observation and manipulation is taking place in new sciences. In the history of modern science, we observe that theoretical prediction has usually preceded the empirical observation, due to limitations in experimental technology. As has been made clear by historians, observation was limited by problems associated to the use of light in optical systems for human vision and subsequently by the level of energy required to observe phenomena at smaller and smaller scale (Galison 1997).

Thus, for example, the theoretical prediction of the structure of large molecules has been advanced much before electron microscopy was in place. In turn, the lower the level of resolution in the hierarchy of matter, the more difficult it is to manipulate the physical reality. In chemistry, for example, materials have long been manipulated in large samples, exploiting statistical properties, rather than molecule by molecule.

In general, observation may go *deeper* in the structure of matter than manipulation. There may be situations in which science is in a predictive stage but its development is made more difficult by obstacles in manipulating the matter at the same level of resolution. A huge effort is then made in designing and manufacturing new experimental facilities. The unbalance between observation and manipulation may be a major reason for slow scientific development.

This state of affairs has deeply changed in the last three decades for new sciences mainly due to the application of scientific breakthroughs to scientific equipment and the rise of an independent instrumentation industry. Milestones in these developments are the invention of monoclonal antibodies, recombinant DNA, and polymerase chain reaction in life sciences, the refinement of electron microscope, and the invention of atomic force microscope (AFM) and scanning tunnelling microscope (STM). All these developments took place between the 1970s and 1980s, contributing immensely to the acceleration of new sciences in the last part of the century.

In a historical perspective, it can be said that the production of scientific instrumentation has gained autonomy with respect to the final users, giving origin to a relatively independent professional community (research technologists) that are able to develop “fundamental instrument theory and the design of generic equipment” (Joerges and Shinn 2001, p. 245). Instrumentation itself has become the object of scientific inquiry, making it easier to get so called Grilichesian breakthroughs, or the invention of methods to produce invention (Zucker and Darby 2003).

These techniques, however, not only open a new world to the observation of lower levels of organization of matter. Indeed, what they offer is, for the first time in the history of science, the possibility to manipulate *and* observe matter at lower levels of resolution *at the same time*.

Perhaps the applications in nanoscience are the most impressive. Richard Feynman firstly considered the very idea of designing and producing objects at the scale of a nanometre in a famous lecture in 1959. Indeed, there was plenty of knowledge on phenomena that occur at a very small magnitude. However, most of the new knowledge was gained only when the feasibility of actually producing objects at that scale was demonstrated. The use of AFM and STM permit scientists to move atoms one by one, building up configurations upon design. This has

changed the nature of experiments in new sciences, bringing together inextricably the observational attitude of scientists (controlled experiment) and the inventive attitude of engineers.

Second, at lower levels of resolution, the properties of matter can be designed. This is a consequence of the transition between classical mechanics to quantum mechanics when moving down to the atomic scale and it is a fundamental feature of the rapid development of nanoscience and nanotechnology. There is no way to predict the properties of a nanotube, a nanoprocessor or a molecular motor than actually designing it molecule by molecule or even atom by atom, either top down from a large structure, or bottom up from chemical reactions, and observing it. As nanoscientists use to say, “for us everything is artificial” (Bonaccorsi and Thoma 2007).

In terms of the quotations from Barrow, Hurtle and Hut above, at the nanoscale there are as many states of matter as they are designed in experiments, and explanations drawn from quantum theory require very long specifications of initial conditions and parameters—indeed, a full description of the design. At this scale the search for fundamental properties of matter and the design of novel forms are intrinsically part of the same intellectual venture.

Third, the increase in computing power has greatly reduced the time and cost of simulation. It is now possible to design and test a large number of objects without the need to produce them physically. This has greatly reduced the distance between scientific investigation and design of new objects. For example, in chemistry the rapid development of computational chemistry has made it possible to produce a proliferation of targets, while in conventional chemistry the cumulateness of knowledge favoured convergence in a few directions. Another example is the emergence of bioinformatics, leading to the possibility to explore computationally, rather than physically, gene sequences or protein networks. Along these lines, achievements in mathematics and algorithms become much closer to physical reality than it was the case in the past (Skinner 1995; Wagner 1998). The increase in computing power is a radically new factor in science.

In these cases, the boundary between knowledge of scientific laws and design of artificial objects becomes blurred, because scientists systematically move back and forth. For all these reasons there are deep differences between eighteenth–nineteenth and late twentieth century sciences in the relation between discovery and invention, or science and technology. These differences do not come from a shift in incentive structures and the social contract (intellectual property, role of industry, government funding) but originate within the fabric of science.

Are New Sciences Really Different?

The discussion above leads us to formulate a few propositions on the characteristics of new sciences. What characterizes late twentieth century science is *not* a process of specialization and creation of new disciplines, is *not* inter- or multi-disciplinarity, and is *not* the orientation towards problem solving induced by shifts in the incentive structure and the social contract. All these aspects are not new or empirically questionable. Specialization and creation of new disciplines has been typical of all modern science, the disciplinary bases of knowledge are not at all eroded, and

orientation towards problem solving has historical antecedents and does not seem to be uniquely associated to changes in the institutional context.

What characterizes new sciences is, instead, the extension of the scientific method to new and highly complex fields, supported by an unprecedented growth in experimental instrumentation and computing power as well as leading to new relations between natural and artificial, or between explanation and manipulation. In a few words, new sciences are reductionist sciences that address new complex phenomena by breaking the boundary between natural and artificial. These changes go deeply into the way in which scientific practice is carried out.

In order to support the claim that new sciences are really something new, however, we have the burden of proof of showing intrinsic and relevant differences. If not, then our claim would only be another way of labelling existing phenomena.

In the rest of the article, we try to demonstrate that new sciences indeed exhibit a dynamic that is qualitatively different from established ones. More generally, we propose that there is a set of abstract and consistent dynamic properties of the search process that transcend specific disciplines and that can be identified, and, to a certain extent, operationalized. We propose three dimensions of analysis for these dynamic properties and formulate propositions to be expanded and validated in the following sections. These propositions translate the theoretical discussion above into dimensions (and then, hopefully, variables) that are closer to the perspective of social science, particularly economics and sociology of science. These are: rate of growth, degree of diversity, type of complementarity.

First of all, new sciences *grow very rapidly*. Not only do new sciences grow rapidly when in their infant stage, but they continue to create new fields of research at high pace even in their maturity. Higher growth comes mainly from the flow of entry of new fields and is reflected in published output. Second, new sciences *grow more diverse*. The acceptance of unifying theories does not lead to a small number of research directions, but rather to an explosion or proliferation of several competing sub-theories. Therefore, we expect not only large diversity, but also growing diversity (divergent search).

Third, new sciences *make use of new forms of complementarity*. While established sciences make use of physical infrastructures (technical complementarity), new sciences require additional forms of complementary inputs. These refer to the integration of heterogeneous competences (cognitive complementarity) and to the need to interact with actors that are institutionally located at various levels of complexity of the systems under investigation (institutional complementarity).

If we find support for these propositions then they might be generalized into a new notion, which we call *search regime*, as the set of abstract dynamic properties of search. For each of these propositions, we first briefly review the relevant literature, then propose one or more operationalizations and examine the preliminary evidence.

Units of Analysis, Boundaries, Classifications

In order to empirically validate the above propositions a number of difficult methodological issues must be addressed. It is correct to admit that our theorizing is more at the level of constructs than of observable variables. In fact, our main

interest is the dynamics of entry of new fields within new and established disciplines, as generated by laboratories working in competition. Unfortunately, several empirical strategies are precluded.

First of all, we cannot observe laboratories. Although, following the literature in the sociology of science, we recognize that laboratories would be the appropriate unit of analysis (Latour and Woolgar 1979; Knorr-Cetina and Mulkey 1983; Latour 1987; Pickering 1995), observing laboratories in a variety of fields and for many years would clearly require a collective long term research strategy (Joly 1997; Laredo and Mustar 2000; Laredo 2001). Our only object of observation is the output produced by scientists in the literature.

Second, we cannot offer an exhaustive classification of disciplines based on careful bibliometric techniques. Our definition of established and new sciences is broad and intuitive. Clearly all our arguments may stay or fall depending on the classification adopted, particularly at the boundaries of disciplines, because, from a theoretical point of view, any classification is likely to introduce value judgments (Bryant 2000). A long tradition in scientometrics, after the seminal work of Michel Callon and co-authors (Callon et al. 1983; Callon 1986) has used co-word analysis and clustering to build up profiles of emerging fields and to examine new disciplines (Callon et al. 1986; Rip 1988; Noyons 2004).

We have not pursued this strategy for this article, with the exception of nanotechnology (see *infra*), where we exploited a newly constructed dataset built upon an iterative lexical-citation algorithm developed by Michel Zitt. More generally, the literature on new fields based on the structure of citations, co-authorships, and co-occurrence of words has shown how variable the boundaries can be in a relatively short time frame (Leydesdorff 1987; Rip 1988) and also how sensitive classifications are with respect to the window of observation (Zitt and Bassecoulard 1994; van Raan 1997; 2004). Trying to pursue a full profile operationalization would not be appropriate for our purposes.

Therefore, we do not claim any validity at the level of classification, profiling and boundaries of disciplines and fields. Rather, we will use *new words* in science as being part of disciplines in a self-evident way, without any effort to define boundaries, with the exception of nanotechnology. The arguments proposed in this article are abstract and exploratory, and hopefully do not require the solution of all practical and empirical problems to demonstrate validity.

So what can we observe to corroborate the propositions sketched above? First, we assume that the process of entry in science is indirectly witnessed by the appearance of a *new word* in the scientific literature. We have used full words in keywords (whenever available) and in titles, eliminating words appearing only once or twice in titles and keeping all words appearing in keywords in the dataset. Taking into account the literature quoted above, it is clear that a new field would produce not single new words, but *clusters* of new words. But our goal is not to examine the full profile of new fields (which would require clustering and mapping). Observing *single new words* is a promising starting point to examine the dynamics of science, because they are associated to novelty. This novelty may well come from intentional strategies of scientists (trying to get recognition and visibility by labelling things in different ways), but it is hard to believe that this effect shows systematic differences

across disciplines or that the same labelling words are used repeatedly by several scientists in an independent way. Therefore we focus on some statistical properties of new words: rate of growth after entry, composition effect (ratio between new and old words within disciplines), concentration index, rank correlation over time.

We are also fully aware of the many limitations of word analysis in science, and more generally of single indicators (Martin 1996; van Raan 2004). However, the use of *single* new words also offers some advantages. They can be used to validate two crucial dimensions: the count of occurrences of new words after their first entry may approximate the rate of growth of new fields in a discipline, and the distribution of new words may approximate the diversity and divergent/convergent dynamics.

Second, more classical indicators of co-authorship and co-invention will be used to approximate new types of complementarity. Taken together, these measures call the attention to new dimensions of the dynamics of science.

Rate of Growth

Why do Growth Rates (in Science) Differ?

Do new sciences grow more than established ones, and if so, why? The question of what determines the long run aggregate growth of science inspired the founders of scientometrics and has led to suggestions of general laws of growth of science, either exponential (Price 1951, 1961; Granovsky 2001) or linear (Rescher 1978, 1996).

We refer to this tradition by suggesting that the question of growth should be interpreted in economic terms, i.e. by addressing the causes and consequences of *different rates of growth of scientific output*, not at the aggregate level, but *across fields*. While the question of aggregate scientific production is interesting (science as such or the total economy), for our purpose it is more useful to refer to production in given scientific fields. This different question has been recently raised by Richard Nelson, who in a certain sense asked “why growth rates differ” among scientific disciplines or, more generally, areas of human know how (Nelson 2005). This intriguing question articulates a rich tradition in history of science and philosophy of science, addressing the determinants of discovery (for a general introduction see Wightman 1951; Hanson 1958; Holton 1986; Ziman 1978, 2000).

When we see a sudden change in the rate of total production in a field, we may think that many young scientists entered the field, or incumbent scientists or teams greatly improved their productivity, or scientists from different fields turned to a new one. When we see a steady state rate of growth in the long run, we may think of totally different dynamics. While the economics of science until now (Stephan 1996; Wible 1998; Shi 2001) has been mostly influenced by the economics of information (Dasgupta and David 1987, 1994), we propose to shift the emphasis to the *industrial dynamics of science*, by examining the pattern of entry and growth of new fields.

Our empirical question can be addressed by examining: (a) whether new sciences grow more than established ones in the aggregate; (b) whether the difference in the rate of growth can be explained by differences in the dynamics of entry.

Aggregate Growth: Preliminary Evidence

Answering the first question is not easy, since after the pioneering contributions quoted above there are no studies on differences between rates of growth. An official analysis, based on a short time window (1995–1999), is offered in the authoritative *Third Report on Science & Technology Indicators* of the European Commission. According to this source (European Commission 2005) on the basis of ISI data, some of the new sciences are indeed among the fastest growing in the last decade (average rate of growth of 10% for computer science, 35% for materials science).

Another partial evidence comes from the Observatoire des Sciences et Techniques, which estimates that the annual average growth of WOS publications, all fields considered, stands at between 1% and 2%, while some fields like human genetics or nanotechnology have grown since their inception between 8% and 14% (OST 2004).

In the case of nanotechnology, specific studies confirm that the growth pattern in the first decade has been exponential with a constant rate of growth with this order of magnitude (Meyer 2001; Zucker and Darby 2003).

Growth from Entry of New Fields

Although all these elements are difficult to access for external observers, an imperfect but reasonable approximation can be reached by using *new words* in published output.

As preliminary empirical evidence, we investigated the structure of two broad scientific disciplines, namely high energy physics and computer science. They largely differ in scientific age and maturity, size of the scientific community, nature of knowledge, institutional setting and the relation between theoretical and experimental activity. Due to these differences, we believe that we are covering sufficient variability to test the validity of our concepts in a crude and preliminary way: high energy physics is the paradigm of established science, while computer science is a new science.

We examine data on publications of the top 1,000 scientists worldwide in terms of citations received, using commonly accepted (and freely accessible) data sources in the respective communities (see Appendix 1 for details). Focussing on top scientists has, of course, limitations, but they are better positioned to identify promising research fields and introduce new concepts (and therefore words) in the literature. For reasons of comparability and coverage of data, we focus our analysis on the period from 1990 to 2001. Due to the exploratory nature of our work, this limitation may be considered acceptable.

A difficult methodological question is: Do observed differences in rates of growth after entry depend on intrinsic dynamics of science, or do they rather reflect institutional differences and external influences (e.g. funding, government priorities, industry demand)? How endogenous is observed scientific growth? In general, it is recognized in the literature that a pattern of sudden growth can be generated by

external factors, such as an increase in funding (see the classical example of anthrax research after terrorist attack in 2001–2002). We must clearly be careful here. Although the question is, in general, a difficult one (Rosenberg 1994), we believe that observed differences in rates of growth are too systematic to be entirely explained by external factors only. In particular, since we are observing large disciplines, it is not credible that all observed patterns depend on external factors.

We take the first appearance of a word in the literature as an event of *entry*. We start from the simple observation that the rate of growth of a scientific field *after* an event of entry is subject to largely different patterns. In particular, there are important cases in which the introduction of a new word is followed by an extremely rapid process of growth, in some cases an exponential growth. What produced such acceleration? Which was the main obstacle to the increase in the rate of production of new knowledge?

We next move to the level of disciplines. A fascinating question, which we can only explore here, is whether differences in the long run rate of growth of disciplines depend on the dynamics of entry of new fields. In order to examine this question, one should have access to several decades or even centuries of records in the scientific literature.

The only “quasi-natural” experiment available is nanotechnology. Complete data on nanotechnology have been produced by Hullmann dated from 1981 and reported in the *Third European Report on S&T Indicators*; further studies have been published in a special issue of *Research Policy* in 2007 (edited by Bozeman and Mangematin). In this case it is clear from data that not only individual fields (such as carbon nanotubes, nanocoatings or nanobiotechnology), but the whole discipline had an impressive growth. In less than 10 years, almost 100,000 scientists worldwide were mobilized around the new discipline. Worldwide, several thousand new institutions entered the field. Indeed, large part of this dynamic is accounted for by an impressive process of entry of new words, a process that surprisingly *increases over time* (Bonaccorsi and Vargas 2007).

It is much more difficult to examine the rate of growth of large established disciplines because of lack of data. In this case, however, we may examine the dynamics of more recent years.

Are there systematic differences in the internal dynamics of disciplines? In other words, can we observe different patterns in the rate of entry of new fields in a discipline and in the ratio between new and established fields? We compute the turnover ratio, defined as the ratio between newly appearing words and the total number of words in articles published by the top 1,000 scientists in high energy physics and computer science. The results are striking (Fig. 2).

In high energy physics the total number of words is much larger than in computer science, ranging from 10,000 to 30,000 per decade. In each year, the proportion of newly appearing words, i.e. of words that do not appear in any of the preceding years in the time frame, is rather low. In computer science, on the contrary, the total number of words does not exceed 5,000, but the proportion of newly appearing ones is much larger. Figure 2 summarizes this finding: between 40% and 60% of words appearing in computer science are new, while the same ratio is between 20% and 25% in most years for high energy physics.

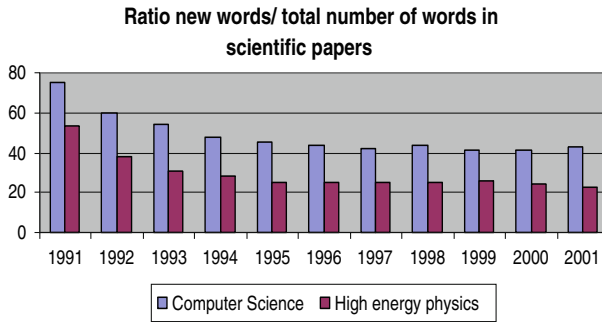


Fig. 2 Ratio between newly appearing words and total number of words in high energy physics and computer science

We therefore establish the following stylized findings:

- i) there is limited but consistent evidence that *new sciences grow more* than the average at the aggregate level;
- ii) there are fields that grow extremely rapid and fields characterized by slow growth after entry. Fast growing fields are distributed across all disciplines;
- iii) *disciplines largely differ in the composition of fields characterized by different rates of growth*. In some disciplines, it seems that new fields are generated continuously, so that the turnover ratio is extremely high, while in other disciplines the turnover is much lower.

Degree of Diversity

Diversity Before Paradigmatic Change versus Diversity Within Normal Science

For one reason or another, scientific fields may see highly differentiated rates of growth and disciplines seem to differ deeply in their industrial dynamics.

After having taken these differences into account, a further question deals with the *direction* of growth. Obviously, we do not deal here with predicting the direction of research. We put forward a more modest question: *how many different directions* of research are pursued at any given time? Following a tradition of research on laboratory practices and the relation between discovery and representation (Collins 1985; Galison 1987; Lynch and Woolgar 1990) Ian Hacking has offered a classical articulation of laboratory sciences, which we follow here for descriptive purposes (Hacking 1992). According to him, laboratory science can be characterized with respect to as many as 15 different dimensions, such as ideas (questions, background knowledge, theory, specific sub-hypothesis, theoretical model of equipment), things (experimental target, source of experimental modification, revealing apparatus, tools, data generators) and symbols (data, evaluation of data, reduction of data, analysis of data, interpretation). In principle, each of these dimensions might be labelled and entered into the title or the keyword of a scientific article, reflecting the

way in which laboratories converge among themselves across all these dimensions, or rather follow very differentiated and diverse patterns of research.

There are two aspects of the problem of direction. The first one is a static problem: how many different directions of research are there in a given field at any point in time, or which is the *degree of diversity*? The second aspect is dynamic: given a certain degree of diversity, however defined, is diversity increasing or decreasing over time? Do scientists search in the same (multidimensional) directions or do they depart from each other? Or, put in other terms, do we observe a *pattern of divergence* or a *pattern of convergence*?

Why there may be diversity in science? A ready-to-made answer is found in Kuhn's classical distinction between pre-paradigmatic science and normal science (Kuhn 1962). Before the emergence of a new paradigm, science is in a fluid stage, in which there are many competing directions of research that in turn may arise as the result of dissatisfaction of scientists for anomalies in the previous paradigm. But after the emergence of a new paradigm, all scientists share the same view of what the key research questions are, they look at fundamentally the same objects, and employ basically the same methods. There is not much diversity within normal science.

Our notion of divergence suggests that, in new sciences, there is room for much diversity and divergence even *within* established paradigms or accepted theories. Even though scientists share the same scientific paradigm, they still may exhibit considerable diversity in the specific hypotheses or sub-hypotheses they want to test, the precise object they are looking for and the detailed experimental techniques and infrastructure they use. As discussed above, this divergent dynamic is fundamentally dictated by the application of methodological reductionism to objects of increasing complexity and by the blurring of boundaries between natural and artificial.

Convergent versus Divergent Pattern of Search

By convergent search regime, we mean a dynamic pattern in which given one or more common premises (e.g. an accepted theory and an agreed research question or general hypothesis); each conclusion (i.e. experimental evidence or theoretical advancement) is a premise for further conclusions. In addition, all intermediate conclusions add support to a general conclusion.³

By divergent search regime, we mean a dynamic pattern in which given one or more common premises each conclusion gives origin to many other sub-hypotheses and then (possibly) new research programmes. Let us distinguish between *strong* and *weak* divergence. In the former case, hypotheses and then research programmes

³ As Maddox popularizes this point in high energy physics: "At present, these four forces of Nature (electromagnetism, gravity, weak and strong) are the only ones known. Remarkably, we need only these four basic forces to explain every physical interaction and structure that we can see or create in the Universe. Physicists believe that these forces are not as distinct as many of their familiar manifestations would seduce us to believe. Rather, they will be found to manifest different aspects of a single force of Nature. At first, this possibility seems unlikely because the four forces have very different strengths. But in the 1970s, it was discovered that the effective strengths of these forces can change with the temperature of the ambient environment in which they act" (Maddox 1998, p. 126). See also the notion of Holy Graal of unification in physics (Weinberg 1992; 't Hooft 1997; Klein and Lachièze-Rey 1999; Randall 2004; Penrose 2005).

are *competing* in the sense that they cannot be true together, although they are all consistent with the general theory or with hypotheses of higher generality. In the latter they are *different*, in the sense that they are located at a large distance in the multidimensional space, for every reason. It is important to note that there is no way to discriminate between sub-hypotheses using the general theory or hypotheses of higher generality level. This means that it is not possible to select among research programmes starting from the more general point of view. Thus, although science is always based on procedures that ensure what logicians and philosophers of science label “convergence to the truth” (Gustason 1994; Kelly 1996; Martin and Osherson 1998), this may happen following different patterns of search.

In Section “Are New Sciences Different”, we hypothesized that new sciences follow a divergent pattern. While a full theoretical discussion of the underlying rationale for a convergent or divergent pattern of search is offered in Bonaccorsi and Vargas (2007), let us check the conjecture against available evidence. We first produce qualitative evidence, drawn from informal discussions with scientists in several fields.⁴ Then we propose some operationalizations.

Degree of Diversity: Preliminary Evidence

How can we approximate this dynamic from a quantitative point of view? Our reasoning is as follows. If a discipline is subject to a divergent search regime, not only will there be many new words appearing per unit of time, as we saw in the previous section, but also new research programmes will not bring the “old” words with them. Over the course of time, we would observe a few highly used words, perhaps those that originated in the fields, and many words with similar intensity of use. In short, we expect words to be less and less concentrated in divergent regimes. This expectation holds for computer science. On the contrary, in a convergent search regime, such as high energy physics, the same higher level words are used and used again, meaning that many research programmes try to contribute to a common core of hypotheses and research questions. In short, we expect convergent regimes to exhibit a strong *concentration* of words. Interestingly, this is what is found in our data.

In high energy physics, the total number of publications of top 1,000 scientists is much larger than in computer science, partly because of seniority of scientists and partly because of different publication patterns. The number of different words is also much larger ($n = 50,952$ vs. 18,031). This is what one would expect from an established and old discipline, dating back in the history of science. The top 250 most largely used words in high energy physics absorb 29.3% of all used words while the value in computer science is 26.5%. More significantly, the corresponding Herfindahl index is 43.4 in physics, 8.3 in computer science.

⁴ Some of these arguments are the result of intense discussions with scientists in a variety of fields: Laura Redivo and Lorenzo Zanella (Glaxo Smith Kline) for HIV, Antonio Cattaneo (SISSA and Lay Line Genomics) for Alzheimer, Claude Mawas (INSERM, Marseille) for cancer, Giovanni Punzi (INFN) for high energy physics, Bruno Codenotti (IIT-CNR) and Gianfranco Bilardi (University of Padua) for computer science, Paolo Dario (SSSUP) for bioengineering, Fabio Beltram (NEST) for nanotechnology and materials. I apologize to all of them for any misunderstanding. Discussions and joint work with Fabio Pammolli (University of Florence and IMT) have taken place for a long time.

Table 2 Concentration of words in publications of top 1000 scientists in computer science and high energy physics

	Computer science	High energy physics
Number of publications of top 1,000 scientists	9,062	41,770
Number of different words in titles and keywords	18,031	50,952
Average number of words per author	5.35	5.44
Concentration ratio (C_{250})*	26.5%	29.3%
Herfindahl index**	8.3	43.4

* Cumulative market share of top 250 words (Number of occurrences of the top 250 words/ total number of occurrences in all publications)

** Computed only on top 250 words

Source: our elaboration. See Appendix 1 for details

If words are an acceptable approximation, then we find evidence of sharp differences in the degree of diversity in search regimes across disciplines (Table 2). As it is visible in Fig. 3, although the top 250 words account approximately for the same share of the total, a few words in physics absorb a larger share of occurrences, while in computer science the distribution is much more dispersed.

While this evidence tells us something about the degree of diversity, the notion of divergence implies that diversity *increases over time*. In order to test this effect in our data, we compute a simple indicator of turnover, i.e. the proportion of words that are among the top in the first 5 years of the period that are also in the same high position in the last 5-year period. The rationale for this measure is simple: highly

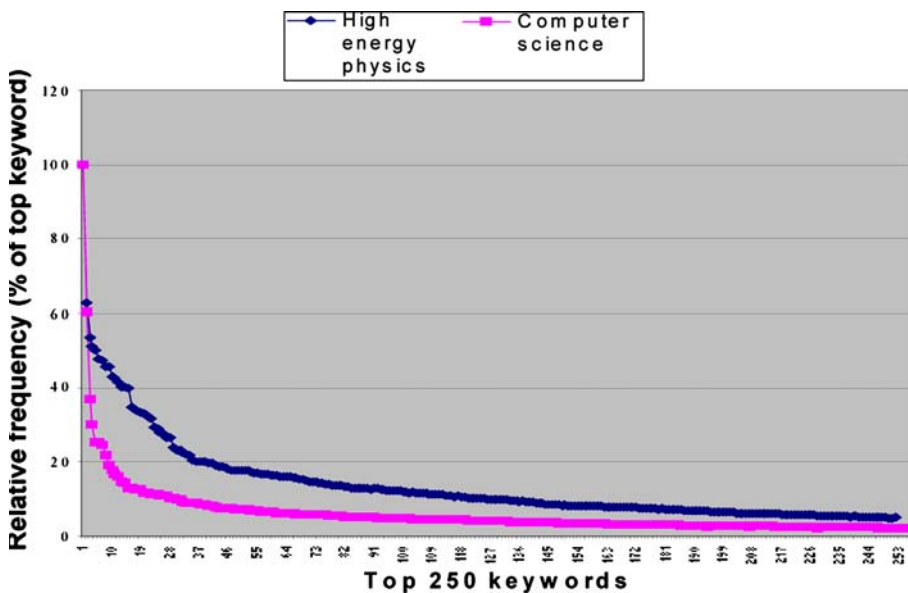


Fig. 3 Distribution of top 250 keywords in High energy physics and Computer science

used words reflect issues that are top in the agenda of scientists, and they would remain stable in a case of convergence. On the contrary, high turnover rate suggests turbulence, which we interpret as the effect of proliferation of new sub-hypotheses, so that the originating words decline in importance and are substituted more frequently by new ones.

In order to take the whole distribution into account, we compute a rank correlation index and plot the relation between the relative positions in the two sub-periods. Again, the differences are striking (Figs. 4, 5). In high energy physics rank correlation is 0.79, while in computer science it is only 0.49.

In the physics discipline, the top 50–100 words in the initial period are still the most widely used at the end of the decade. This is not the case in computer science, where a few words are located along the main diagonal. We believe this indicator, although crude, sheds light on the underlying dynamics.

Two main results are found:

- i) in the case of new science, new words are much less concentrated, suggesting a larger dispersion of new fields and then more diversity;
- ii) in this case, there is also lower correlation over time of the rankings of new words, suggesting a larger turnover, and then divergent patterns of search.

Therefore we find preliminary corroboration to the idea that new sciences not only grow more rapidly, but also grow more diverse. Further work is needed to develop new indicators for dynamic patterns. In [Bonaccorsi and Vargas \(2007\)](#), we

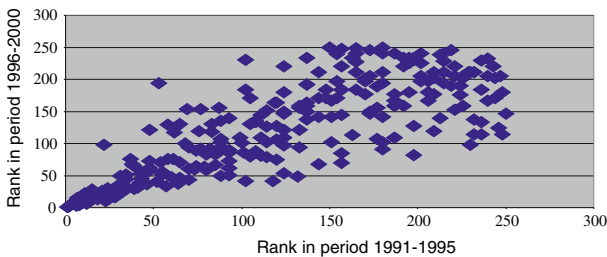


Fig. 4 Plot of rank correlation of top 250 keywords in high energy physics ($r = .79$)

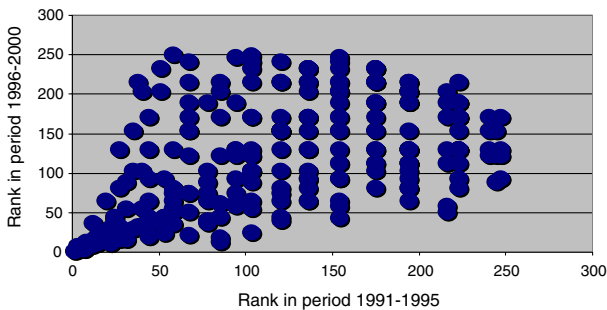


Fig. 5 Plot of rank correlation of top keywords in computer science

present a new indicator of divergence in science, based on the theory of bipartite graphs. Clearly there is a research agenda to be developed here.

Level and Type of Complementarity

A third relevant dimension to characterize the search regime is the level of complementarity. By complementarity, we mean the extent to which different human or material resources are needed as inputs, in addition to the intellectual resources of the scientist himself (*superadditivity*: see Milgrom and Roberts 1990, 1992). We suggest that there are three types of complementarity in science: cognitive, technical and institutional. We are interested in evaluating whether new sciences make use of a different mix of complementarities than established sciences.

Cognitive Complementarity

The discussion on new sciences has shown two distinctive features: new sciences cut across different layers of complex hierarchical systems, and work at the boundary between natural and artificial. New sciences are cognitively more articulated than established sciences.

However, this complementarity is not driven, as the literature on Mode 2 has suggested, by the complexity of the societal or economic problems addressed by research (e.g. climate change, energy or environment). More deeply, complementarity is dictated by epistemic needs and is therefore organized around two main axes: the layers of reality observed, and the interface between natural and artificial. Scientific disciplines pursue research questions on particular objects of investigation that ultimately follow the layers of organization of matter. This articulation is not at all abandoned in new sciences: at each layer of reality scientists work with toolboxes tailored to that layer, and still claim that they produce valid explanations. At the same time, scientists and engineers may work on the same objects (e.g. a nanowire), although they will still use different cognitive approaches.

The emphasis on multi-disciplinarity severely underestimates the methodological weight of reductionism. We do not witness disciplinary fusion or integration, but rather a number of institutional and organizational solutions that support cognitive complementarity, such as new boundary roles (e.g. transfer sciences in life sciences), new PhD curricula, joint laboratories and shared instrumentation.

Technical Complementarity

Established disciplines were based on a dichotomy between large and dedicated facilities in big science (high energy physics, astrophysics, space, nuclear) and small scale laboratories in other fields (chemistry, biology). Separate institutions, usually at an intergovernmental level, were created for the former, while ordinary funding at the national level covered the latter.

In new sciences a different situation emerged, in which medium-size, general purpose facilities are needed, such as atomic force microscopy and STM in

nanotechnology, or bioinformatics databases in biotechnology. A variety of institutional arrangements are available for funding and management, and the scale exceeds the current funding of departments, but national and even regional policies are totally appropriate. This may partially explain the shift towards regional policies for science.

Institutional Complementarity

Finally, we introduce the notion of *institutional complementarity*. This refers to the degree to which research needs the contribution of scientists and teams working in different institutional environments, having different priorities and views as well as bringing different types of data and experience.

Again, this is a sharp consequence of the distinctive features of new sciences. A good case in point is biomedical research. Until the molecular biology revolution, clinical research was based on observations at the level of the entire human body, or main subsystems, organs, or specialised tissues. The main cognitive processes were clinical or expert judgment, induction and generalization. After the molecular biology revolution, the cognitive processes have been mostly made of causal reasoning at the cellular, intracellular or molecular level, creating a systematic link between clinical research and biology. Different hierarchical levels of the human body have been put in connection explicitly. But those that have access to information at different levels of organization of the human body are also working at different institutions: hospitals (data on patients), traditional laboratories (data on organs and tissues), biochemistry and molecular biology laboratories and academic departments (data on enzymes, proteins and genes). If these institutions are not coordinated, then data do not flow and people do not talk to each other.

Another interesting case is the relation between academia and industry in new sciences. While most of the debate about the Ivory Tower and the commercialization of academic commons has been driven by concerns about institutional rules, few analysts have noted that the rise in collaboration has been taking place exactly in the same period in which new sciences emerged and grew exponentially in the scientific landscape. These new sciences are intrinsically based on a close relation between academia and industry, precisely because they work at the boundary between natural and artificial, and their discoveries, at the same time, need and contribute to technological achievements in an unprecedented way. New sciences are intrinsically based on institutional complementarities.

New Forms of Complementarity: Preliminary Evidence

While evidence on physical infrastructure may be found more easily, it is much more difficult to provide examples of institutional complementarities. We have been able to collect interesting evidence on some aspects of this issue in computer science, high energy physics and nanotechnology. They provide preliminary evidence, to be extended in future research.

Table 3 shows data on affiliations of top 1,000 scientists in computer science and high energy physics and their co-authors. We focus on patterns of co-authorship

Table 3 Structure of affiliations and institutional complementarities in publications of top 1,000 scientists in Computer science and High energy physics

	Computer science	High energy physics
Number of publications with full affiliations	6,386	34, 340
Number of different affiliations identified	1,857	6,439
% of papers with at least one industrial affiliation	21.3	13.1
% of papers with only academic affiliations	67.5	49.8
% of papers with at least one industrial affiliation and at least one academic affiliation	17.8	7.5
Total number of occurrences of affiliations in publications from the top 100 affiliations	8,304	68,863
Of which:		
- total number of occurrences of universities	7,136 (85.93%)	51,356 (74.58%)
- total number of occurrences of other research institutions	296 (3.56%)	17,507 (25.42%)
- total number of occurrences of companies	872 (10.50%)	0 (0.00%)

Source: see Appendix 1

between academicians and industrial researchers on the one hand, and researchers from government labs or other research centres on the other.

It is clear that the two disciplines exhibit completely different complementarity patterns. In computer science 21.3% of publications of top 1,000 scientists have been authored either by an industrial researcher who is in the top list himself, or they were co-authored by industrial researchers. Among the top 100 affiliations, there are 872 cases of company affiliation that represent 10.5% of the total. In the list of top 100 affiliations, following MIT, Stanford and the University of Maryland we find IBM, while the list also includes AT&T, Digital Equipment, NEC, Xerox, Microsoft and Lucent. We can say that large companies are an institutional component of top quality scientific research in computer science.

Interestingly, companies are also important in high energy physics, but with a very different role. 13.1% of papers have at least an industrial affiliation, but joint industry–academy papers make up only 7.5% of the total and none of the industrial affiliations shows up in the list of top 100 affiliations. This is not surprising, given the more fundamental nature of knowledge in this discipline. Here the institutional complementarities clearly go in a different direction: 40.7% of papers have at least one affiliation from government labs or other research centres. The institutional design must take into consideration the need for universities to get access to large scale facilities on an international basis.

With respect to nanotechnology, an interesting insight into the type of complementarities is given by the combination of data on publications, patents and start-ups. A large share of inventors of patents are also authors of scientific papers, and a large share of founders of new companies are also either inventors or authors. In particular, the set of patents can be divided into three approximately equal groups: patents invented by groups of inventors of which all have published at least a scientific article in nanotechnology (only authors); patents invented by

Table 4 Composition of communities of inventors in nanotechnology

	Only authors	Only inventors	Authors-inventors	Total
Number of patents	1,442	1,607	1,779	4,828
Number of individuals	2,202	2,400	4,662	9,224
Number of participations of individuals to patents	3,465	3,473	7,130	14,068

Source: our elaboration; see [Bonaccorsi and Thoma \(2007\)](#)

groups in which no one has ever published (only inventors); patents invented by groups with mixed origins (Table 4).

It is interesting to observe that the large majority of patents (3,221 out of 4,828) have at least one inventor who is also an active scientist. We also find a statistically significant difference in an overall index of patent quality based on a factor model: patents invented by communities with the highest level of institutional complementarity (author–inventor) have better quality than others. Therefore, we find evidence of a highly interconnected knowledge system, in which the transformation of scientific achievements into patentable results and of both into commercial ventures is very rapid, taking place through the multiple roles played by scientists themselves.

Conclusions: The Search Regime Agenda

In the last half century, and accelerating after the 1970s, we have seen the emergence of a large number of new fields and of some new disciplines in which the search regime is characterized by extremely high rate of growth, high degree of diversity and new forms of complementarity. They might be defined as *new sciences*.

Although their disciplinary backgrounds are different, they share a number of fundamental dynamic properties:

- (a) *there is a permanent process of entry of new fields*, which often grow exponentially or, more generally, grow very rapidly after entry. Acceleration in post-entry growth rates is a prominent characteristic of these fields. Considering new disciplines as collections of fields, the turnover rate is high. This situation is in sharp contrast to scientific fields which originated in the nineteenth century or earlier, in which scientific change has taken the form of paradigmatic, long term, revolutionary change, while normal science was characterized by slower rates of growth;
- (b) *these fields are subject to divergent dynamics of search*. New hypotheses are generated from within established paradigms, leading to new research programmes that exhibit strong or weak divergence. Again, this is in sharp contrast to classical physics, chemistry, biology or medicine of the nineteenth and twentieth centuries, in which divergence was an exceptional event leading to major changes while normal science was mainly convergent;

- (c) *new forms of complementarities arise*, beyond the traditional big science scenario. These complementarities take the form of processes of cross-disciplinary competence building (but not necessarily integration), or new forms of design of infrastructural utilization, or institutional cooperation across different types of actors. In economic terms, new sciences make use of a different vector of inputs.

We have provided a theoretical discussion of the underlying rationale for the distinctive features of search dynamics in these fields, drawn from the scientific literature. Then we have translated these into economic dimensions, adopting an industrial dynamics framework: rate of growth, degree of diversity and type of complementarity in inputs. A preliminary operationalization, based on new types of indicators, was suggested along with some stylized evidence in the disciplines of high energy physics (established science), computer science and nanotechnology (new sciences) supporting the main propositions. We believe the evidence is original and systematic enough to corroborate the main arguments, although we recognize the need for further examination and large scale validation of our preliminary results.

These results suggest that the notion of search regimes might be generalized. Combining the three dimensions offers a rich framework for examining search regimes with different rates of growth (stationary regimes vs. fast growing), dynamics of diversity (convergent vs. divergent) and types of complementarity (cognitive, technical, institutional). Each search regime is an abstract characterization of dynamic properties, with peculiar implications in terms of institutional requirements and policy implications.

Several directions of research can be identified. First, science policy studies might address more systematically the way in which different institutional settings and policies are able to cope with the dynamic properties of search regimes. The framework should also be enlarged and include, together with the internal dynamics of knowledge, social and economic factors such as research funding or shifts in research priorities. We have initiated this exploration with respect to European science (Bonaccorsi 2007), but the agenda should clearly be developed.

Second, there is a need to conceptualize the hot debate on change in the social contract between scientists and society (Guston and Keniston 1994; Guston 2000; Mirowski and Sent 2002), as crystallized in the Humboldtian model of university and the Vannevar Bush endless frontier model (Martin 2003), and on the emergence of the third mission. While most of the literature focuses on external socio-political and economic factors, our approach places the attention on the internal pressure from knowledge dynamics in new sciences, particularly in creating new forms of complementarity.

Third, introducing industrial dynamics in the field of science and technology studies will require further conceptual work and empirical operationalization. New indicators will have to be identified, defined and measured. Subsequently, large scale validation exercises will be needed.

While the notion of search regimes is still in a conceptual stage, it seems to be able to suggest a number of promising research directions.

Appendix 1

Computer Science

In the community of computer scientists, the website managed by NEC Corporation (<http://citeseer.nj.nec.com/mostcited.html>) is considered as a reliable source. It is based on the automatic updating of information from a large list of journals in computer science that add to the total count of citations that each author receives.

We built up a database with personal information of top 1,000 scientists in the CiteSeer ranking. Curriculum vitae of all top scientists were manually downloaded and classified, given the lack of automatic software tools for the structuration of informal information such as CVs.

An extensive cross-validation of information was carried out. Together with CVs we built up from CVs the list of publications of computer scientists and downloaded them whenever possible. We manually inspected words and affiliations from full downloaded papers, cover pages, or titles. We only considered articles published in international journals, following the same broad definition of standard ISI data. Clearly the set of journals is larger than the one considered in the relevant section of ISI data, giving a more accurate picture of scientific activity in the computer science community.

High Energy Physics

The community of high energy physics manages its own website, where scientists upload all their publications.

The website at <http://www.santafe.edu/~mark/collaboration/> provides the list of top most productive scientists. Again, we downloaded the full list of publications of these authors. We manually inspected words and affiliations from these papers, following the same methodology as above.

Nanotechnology

We used the full keyword structure developed by Fraunhofer ISI Karlsruhe for nanotech (courtesy of Ulrich Schmoch). With this query, we built up the full list of publications in nanotechnology, from ISI, comprising more than 100,000 papers in the period 1990–2001, under a PRIME project on Nanodistricts, for which the use of ISI data was authorized. Data on publications have been manually matched with data on inventors, drawn from patent data at USPTO and EPO, and the combinations of authorships and inventorships have been derived. Details can be found in [Bonaccorsi and Thoma \(2007\)](#) and [Bonaccorsi and Vargas \(2007\)](#).

References

- Amato, I. (ed.). 2002. *Science. Pathways of discovery*. New York: John Wiley.
- Barrow, J.D. 1998. *Impossibility. The limits of science and the science of limits*. Oxford University Press: Oxford.

- Bonaccorsi, A. 2007. Better institutions vs better policies in European science. *Science and Public Policy*, June: 303–316.
- Bonaccorsi, A., and G. Thoma. 2007. Institutional complementarity and inventive performance in nanotechnology. *Research Policy*, April.
- Bonaccorsi, A., and J. Vargas. 2007. Ant models in science. Patterns of search in nanoscience and technology. Paper presented to the PRIME Nanodistrict Workshop, Paris, 7–9 September. (Under review).
- Bryant, R. 2000. *Discovery and decision. Exploring the metaphysics and epistemology of scientific classification*. Cranbury, NJ: Associated University Presses.
- Bucciarelli, L. 1994. *Designing engineers*. MIT Press: Cambridge, Mass.
- Callon, M. 1986. Four models for the dynamics of science. In *Handbook of science and technology studies*, ed. J.C. Peterson, G.E. Markle, S. Jasanoff, and T. Pinch, 29–63. London: Sage
- Callon, M., J.P. Courtial, W.A. Turner, and S. Bauin. 1983. From translations to problematic networks: An introduction to co-word analysis. *Social Science Information* 22: 191–235.
- Callon, M., J. Law, and A. Rip (eds.). 1986. *Mapping the dynamics of science and technology*. London, MacMillan: Sociology of science in the real world.
- Casti, J.L., and A. Karlqvist. 1996. *Boundaries and barriers. On the limits to the scientific knowledge*. Reading, Mass: Addison Wesley.
- Cohen, M. 1976. Reciprocity in materials design. *Materials Science and Engineering* 25 (3)
- Collins, H.M. 1985. *Changing order. Replication and induction in scientific practice*. Chicago, IL: The University of Chicago Press.
- Conner, C.D. 2005. *A people's history of science miners, midwives, and "low mechanics"*. New York: Nation Books.
- Corbellini, G. 1999. *Le grammatiche del vivente*. Bari: Laterza.
- Danchin, A. 1998. *La barque de Delphes Ce que révèle le texte des génomes*. Paris: Odile Jacob.
- Dasgupta, P., and P. David. 1987. Information disclosure and the economics of science and technology. In *Arrow and the ascent of modern economic theory*, ed. G. Feiwel. New York: New York University Press.
- Dasgupta, P., and P. David. 1994. Towards a new economics of science. *Research Policy* 23: 487–521.
- Dupré, J. 2005. Understanding contemporary genomics, Mimeo, *ESRC Centre for Genomics in Society*. University of Exeter.
- European Commission. 2003. *Third European report on science and technology indicators*. Luxembourg: Office of Publications.
- Feynman, R. 1963. *Six easy pieces*. CA: California Institute of Technology.
- Galison, P. 1987. *How experiments end*. Chicago: University of Chicago Press.
- Galison, P. 1997. *Image and logic: Material culture of microphysics*. Chicago: University of Chicago Press.
- Galison, P. 2003. *Einstein's clocks, Poincare's maps Empires of time*. Chicago: University of Chicago Press.
- Gibbons, M., C. Limoges, H. Nowotny, S. Schwartzman, P. Scott, and M. Trow. 1994. *The new production of knowledge. The dynamics of science and research in contemporary societies*. London: Sage Publications.
- Granovsky, Y.V. 2001. Is it possible to measure science? V.V.Nalimov's research in scientometrics. *Scientometrics* 52(2): 127–150.
- Gustason, W. 1994. *Reasoning from evidence*. New York: MacMillan.
- Guston, D.H. 2000. *Between politics and science. Assuring the productivity and integrity of research*. Cambridge: Cambridge University Press.
- Guston, D.H., and K. Keniston. 1994. *The fragile contract*. Cambridge, Mass: MIT Press.
- Hacking, I. 1992. The self-vindication of the laboratory sciences. In *Science as practice and culture*, ed. A. Pickering. Chicago, IL: University of Chicago Press.
- Hanson, N.R. 1958. *Patterns of discovery*. Cambridge: Cambridge University Press.
- Hartle, J.B. 1996. Scientific knowledge from the perspective of quantum cosmology. In *Boundaries and barriers. On the limits to scientific knowledge*, ed. J.L. Casti, and A. Karlqvist. Reading, Mass: Addison-Wesley.
- Holton, G. 1986. *The advancement of science and its burdens*. Cambridge: Cambridge University Press.
- Hut, P. 1996. Structuring reality: The role of limits. In *Boundaries and barriers. On the limits to scientific knowledge*, ed. J.L. Casti, and A. Karlqvist. Reading, Mass: Addison-Wesley.

- Joerges, B., and T. Shinn. 2001. Research-technology in historical perspective. An attempt at reconstruction. In *Instrumentation between science, state and industry*, ed. B. Joerges, and T. Shinn. Dordrecht: Kluwer Academic Publishers.
- Joly, P.B. 1997. Chercheurs et laboratoires dans la nouvelle économie de la science. *Revue d'Economie Industrielle* 79: 77–94.
- Kay, L. 1993. *The molecular vision of life. Caltech, the Rockefeller Foundation and the rise of the new biology*. Oxford: Oxford University Press.
- Kelly, K.T. 1996. *The logic of reliable inquiry*. New York: Oxford University Press.
- Kevles, D.J., and L. Hood (eds.). 1992. *The code of codes. Scientific and social issues in the human genome project*. Cambridge, MA: Harvard University Press.
- Klein, E., and M. Lachièze-Rey. 1999. *The quest for unity. The adventure of physics*. New York: Oxford University Press.
- Kline, S.J. 1985. What is technology. *Bulletin of Science, Technology and Society* 1: 215–218; also published in R.C.Scharff, V.Dusek, eds. *Philosophy of technology. The technological condition: An anthology*. Malden, Mass: Blackwell Publishing.
- Kline, S.J. 1995. *Conceptual foundations for multidisciplinary thinking*. Stanford: Stanford University Press.
- Kline, S.J., and N. Rosenberg. 1986. An overview of technological innovation. In *The positive sum strategy*, ed. R. Landau, and N. Rosenberg. Washington: National Academy Press.
- Knorr-Cetina, K., and M. Mulkay (eds.). 1983. *Science observed. Perspectives on the social study of science*. London and Beverly Hills: Sage.
- Kuhn, T.S. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Laredo, P. 2001. Benchmarking of R&D policies in Europe: Research collectives as an entry point for renewed comparative analyses. *Science and Public Policy* 28(4): 285–294.
- Laredo, P., and P. Mustar. 2000. Laboratory activity profiles: An exploratory approach. *Scientometrics* 47(3): 515–539.
- Latour, B. 1987. *Science in action*. Milton Keynes: Open University Press.
- Latour, B., and S. Woolgar. 1979. *Laboratory life. The construction of scientific facts*. London: Sage.
- Lewontin, R. 1991. *Biology as ideology*. Ontario: Anansi Press.
- Lewontin, R. 1994. *Inside and outside. Gene, environment, and organism*. Worcester, MA: Clark University Press.
- Leydesdorff, L. 1987. Various methods for the mapping of science. *Scientometrics* 11(5–6): 295–324.
- Llerena, P., and F. Mayer-Krahmer. 2003. Interdisciplinary research and the organization of the university: general challenges and a case study. In *Science and innovation. Rethinking the rationales for funding and governance*, ed. A. Geuna, A.J. Salter, and W.E. Steinmuller. Cheltenham: Edward Elgar.
- Lynch, M., and S. Woolgar (eds.). 1990. *Representation in scientific practice*. Cambridge, Mass: MIT Press.
- Maddox, J. 1998. *What remains to be discovered*. New York: Simon and Schuster.
- Martin, B. 1996. The use of multiple indicators in the assessment of basic research. *Scientometrics* 36: 343–362.
- Martin, B. 2003. The changing social contract for science and the evolution of the university. In *Science and innovation. Rethinking the rationales for funding and governance*, ed. A. Geuna, A.J. Salter, and W.E. Steinmuller. Cheltenham: Edward Elgar.
- Martin, E., and D.N. Osherson. 1998. *Elements of scientific inquiry*. Cambridge, MA: MIT Press.
- Metzger, N., and R. Zare. 1999. Interdisciplinary research: from belief to reality. *Science* 283(5402): 642–643.
- Meyer, M. 2001. Patent citation analysis in a novel field of technology: an exploration of nano-science and nano-technology. *Scientometrics* 51(1): 163–183.
- Meyer, M., and O. Persson. 1998. Nanotechnology-interdisciplinarity, patterns of collaboration and differences in application. *Scientometrics* 42(2): 195–205.
- Milgrom, P., and J. Roberts. 1990. The economics of modern manufacturing: Technology, strategy and organization. *American Economic Review* 80: 511–528.
- Milgrom, P., and J. Roberts. 1992. *Economics, organization and management*. Englewood Cliffs: Prentice-Hall.
- Mirowski, P., and E. Sent. 2002. *Science bought and sold. The new economics of science*. Chicago: University of Chicago Press.
- Morange, M. 1998. *A history of molecular biology*. Cambridge, MA: Oxford University Press.

- Nelson, R. 2005. The uneven evolution of human know how. *Research Policy*, April.
- Noyons, E.C.M. 2004. Science maps within a science policy context. In *Handbook of quantitative science and technology research*, ed. H.F. Moed, W. Glanzel, and U. Schmoch. Dordrecht: Kluwer Academic Publishers.
- Olson, G.B. 1997. Computational design of hierarchically structured materials. *Science* 277(5330): 1237–1242.
- Penrose, R. 2004. *The road to reality*. New York: Knopf Publishers.
- Pestre, D. 1997. La production des savoirs entre académies et marché. *Revue d'Economie Industrielle*. Special issue: L'économie industrielle de la science, ed. M. Callon, D. Foray, vol. 79, pp. 163–174.
- Pestre, D. 2003. *Science, argent et politique. Un essai d'interprétation*. Paris: Institut National de la Recherche Agronomique.
- Pickering, A. (ed.). 1992. *Science as practice and culture*. Chicago, IL: University of Chicago Press.
- Pickering, A. 1995. *The mangle of practice. Time, agency and science*. Chicago, IL: Chicago University Press.
- Polanyi, M. 1962. *Personal knowledge*. Chicago: University of Chicago Press.
- Price, D.J. 1951. Quantitative measures of the development of science. *Archives Internationales d'Histoire des Sciences* 14: 85–93.
- Price, D.J. 1961. *Science since Babylon*. New Haven: Yale University Press.
- Rafols, I., and M. Meyer. 2006. Knowledge-sourcing strategies for cross-disciplinarity in bionanotechnology. In *SPRU Electronic Working Paper Series* no. 152, July.
- Randall, L. 2005. *Warped passages*. New York: Harper Collins Publishers.
- Rescher, N. 1978. *Scientific progress. A philosophical essay on the economics of research in natural science*. New York: Basil Blackwell.
- Rescher, N. 1996. *Priceless knowledge? Natural science in economic perspective*. Lanham: Rowman & Littlefield Publishers.
- Richerson, P.J., and R. Boyd. 2005. *Not by genes alone: How culture transformed human evolution*. Chicago: University of Chicago Press.
- Rip, A. 1988. Mapping of science: Possibilities and limitations. In *Handbook of quantitative studies of science and technology*, ed. A.F.J. van Raan, 253–273. Amsterdam: Elsevier Publishers.
- Rosenberg, N. 1994. *Exploring the black box*. Cambridge: Cambridge University Press.
- Rossi, P. 1962. *I filosofi e le macchine*. Milano: Feltrinelli.
- Ruelle, D. 1991. *Hasard et chaos*. Paris: Odile Jacob.
- Schrödinger, E. 1944. *What is life? The physical aspect of the living cell*. Cambridge: Cambridge University Press.
- Schummer, J. 2004. Multidisciplinarity, interdisciplinarity and patterns of research collaboration in nanoscience and nanotechnology. *Scientometrics* 59 (3): 425–465.
- Shi, Y. 2001. *The economics of scientific knowledge*. Cheltenham: Edward Elgar.
- Simon, H.A. 1981. *The sciences of the artificial*. Cambridge, Mass: MIT Press. (original edition 1969).
- Skinner, R.E. 1995. *The usage of engineered artefacts for scientific abstractions*. Mimeo: Stanford University. November.
- Smith, C.S. 1981. *A search for structure*. Cambridge, Mass: MIT Press.
- Stephan, P. 1996. The economics of science. *Journal of Economic Literature* 39: 1199–1235.
- 't Hooft, G. 1997. *In search of the ultimate building blocks*. Cambridge: Cambridge University Press.
- van Leeuwen, T., and R. Tijssen. 2000. Interdisciplinary dynamics of modern science. Analysis of cross-disciplinary citation flows. *Research Evaluation* 9 (3): 183–187.
- Van Raan, A.F.J. 1997. Scientometrics: State-of-the-art. *Scientometrics* 43: 129–139.
- Van Raan, A.F.J. 2004. Measuring science. Capita selecta of current main issues. In *Handbook of quantitative science and technology research*, ed. H.F. Moed, W.G. Glanzel, and U. Schmoch. Dordrecht: Kluwer Academic Publishers.
- Vincenti, W. 1990. *What engineers know and how they know it*. Baltimore: Johns Hopkins University Press.
- Wagner, P. 1998. *La machine en logique*. Paris: Presses Universitaires de France.
- Weinberg, S. 1992. *The dreams of a final theory*. New York: Pantheon Books.
- Whitley, R. 1984. *The intellectual and social organization of the sciences*. Oxford: Oxford University Press. (2nd edition 2000).
- Wible, J.R. 1998. *The economics of science. Methodology and epistemology as if economics really mattered*. London: Routledge.
- Wightman, W. 1951. *The growth of scientific ideas*. New Haven: Yale University Press.

- Ziman, J. 1978. *Reliable knowledge*. Cambridge: Cambridge University Press.
- Ziman, J. 2000. *Real science. What it is, and what it means*. Cambridge: Cambridge University Press.
- Zitt, M., and E. Bassecoulard. 1994. Development of a method for detection and trend analysis of research fronts built by lexical or co-citation analysis. *Scientometrics* 30 (1): 333–351.
- Zucker, L., and R. Darby. 2003. Grilichesian Breakthroughs: Inventions of Methods of Inventing and Firm Entry in Nanotechnology In *NBER working paper* no. 9825.