COMMUNICATION NORMS AND THE COLLECTIVE COGNITIVE PERFORMANCE OF 'INVISIBLE COLLEGES'

Modeling the Epistemological Functioning of Scientific Research Networks

by

Paul A. David *
All Souls College, Oxford & Stanford University

First draft: August 1996
This version : May 1997

**Note of Acknowledgment*

The present stochastic formulation of the role of communications networks in science has benefited immeasureably from my previous collaborative projects with Dominque Foray and Jean-Michel Dalle, on Markov random field models of network phenomena in other applications contexts. The particular line of work leading to this paper was encouraged by the comments of Patrick Cohendet and other participants at his CNET Workshop on the Economics of Networks, in Paris on 20th April 1995, where 1 described some very preliminary results from a model of research network size and consensus formation. The present version draws on material subsequently developed in a paper with Tom Flemming (1995) that was presented to the International Conference on Creation and Transfer of Knowledge: Institutions and Incentives, held in Castelgandolfo (Rome), 21-23 September 1995. Manuel Trajtenberg's insightful discussion, and the remarks of Adam Jaffe, Giorio Navaretti, Jacques Thiesse, and others on that occasion, have contributed substantially to improving the exposition. Elizabeth Johansson also helped, by diverting her time from more exciting activities to supply the answer to a problem in applied probability theory that arose in connection with the argument of section 4.2. I gratefully record, too, the probing questions and suggestions made by Albert Bolink, Sajeev Goyal, and Uskali Maki and others at the Economics and Philosophy of Science Seminar at Erasmus University, Rotterdam (13 November 1995); and the comments of Patrick Bolton and other participants of the Economic Theory Seminar of CENTER at the University of Tilburg (12 March 1996). Olivier Favereau's close critical discussion of the next draft of the paper, presented at the 1MRI Workshop of the University of Paris-Dauphine (11 April 1996), contributed materially to its further improvement; so, too, did my conversations with Kenneth Arrow, Arie Rip, and other participants in meetings of the Seminar on the Economics of Information Diffusion convened at All Souls College, Oxford, during Trinity Term 1996. Those whose kind willingness to help is recorded in the foregoing recounting of my intellectual debts must be fully exonerated from blame for such defects as have withstood their best corrective efforts.

Contact Author: P. A. David, All Souls College, Oxford 0X1 4AL, U.K.
Fax: 44 + (0)1865 + 279299; Email: <paul.david@economics.ox.ac.uk>

ABSTRACT

COMMUNICATION NORMS AND THE COLLECTIVE COGNITIVE PERFORMANCE OF 'INVISIBLE COLLEGES'

Modeling the Epistemological Functioning of Scientific Research Networks

Scientific research communities, or 'invisible colleges' are conceptualized in this paper as social communications structures formed by the overlapping of more compact networks of personal associations among researchers in particular disciplinary fields of inquiry. Within these 'colleges' circulate ideas and opinions regarding the "validity" of specific scientific propositions and methods. Transmission of tacit knowledge takes place through local social network connections, and is regarded to be critical in enabling individual researchers to participate in, and contribute to the collective epistemological performance. At the micro-behavioral level, the analysis posits a population of rational research agents who engage individually in continuous processes of experimental observation and Bayesian inference, but the interpretation placed upon their current empirical observations is influenced by the distribution of beliefs in their local social network. The agents' communication policies vis-a-vis others in their local network affect the transmission of both new propositions or scientific claims, and prevailing opinion as to the reliability of such statements.

Formulation of this structure in graph theoretic terms permits the use of results from Markov random field theory to establish the ways in which knowledge communications behaviors of members of local social networks affect the capacity of the invisible college (as the ensemble of these networks) to arrive at "closure" on cognitive issues. This yields a stylized, highly simplified model of the processes through which new knowledge-statements circulate, and may (or may not) acquire wider acceptance as valid, thereby being added to the stock of codified, "reliable knowledge." Its key features are shown to resemble closely the so-called "voter model" of Clifford and Sudbury (1973) and Holley and Liggett (1975), permitting application of results obtained for that stochastic system. A knowledge consensus (represented as a configuration of correlated belief orientations with regard to a new scientific statement) is shown to be an emergent and path-dependent property of the network ensemble. In addition to integrating a number of distinct disciplinary perspectives on science, the analysis underscores the central policy significance of the Mertonian norms of cooperation, disclosure and universalism. Through the influence exerted upon micro-level communications patterns of rational agents operating under a priority-based reward system, these cultural norms associated with participation in "open science" can be shown critically to affect the collective cognitive performance of research networks. Results from percolation theory, due to Hammersley and Welsh (1980), are used to formally underscore this aspect of the argument.

Keywords: economics of science, research networks, invisible colleges, Mertonian norms, social communications, collective epistemology, local interaction games, Markov random fields, percolation, path-dependence

COMMUNICATION NORMS AND THE COLLECTIVE COGNITIVE PERFORMANCE OF 'INVISIBLE COLLEGES'

Modeling the Epistemological Functioning of Scientific Research Networks

0. Introduction and overview

Scientific research communities may be studied as social networks within which ideas or statements circulate, acquire validity as reliable knowledge, and are recombined to generate further new ideas. Social networks also form the locus for the transmission of tacit knowledge and skills requisite to the interpretation and operationalization of scientific statements. These extensive, yet informal structures of inter-personal knowledge-transactions have been referred to as constituting "invisible colleges." This paper develops an abstract and highly stylized account of the communications structure of an invisible college, and examines its collective epistemological performance by employing concepts and results from Markov random field theory.

A central task in this undertaking is to characterize formally the relationship between the epistemological notion of scientific "closure", on the one hand, and the sociological notion of the organizational ethos and "norms" of invisible colleges and kindred scientific communities, on the other hand. The latter are taken to reflect themselves in micro-level stochastic behaviors affecting local social communications within a finite size system — which is constructed as the (global) network of (local) social communications networks. Using a particular stochastic structure (the "voter model" of Clifford and Sudbury (1973) and Holley and Liggett (1975)), a knowledge consensus, represented as a configuration of correlated belief orientations regarding a new scientific statement, is shown to be an emergent and path-dependent property of the global network. In that sense alone, however, "closure" is not held to be sufficient to establish a proposition's objective validity, nor is it tantamount to the stabilization of beliefs *over the long run.* I propose an approach to resolving the latter problem, by adopting a "weak relativist" position in regard to the objective "truth" or "falisity" of scientific statements, and embedding the stochastic process of consensus formation within a long-run evolutionary epistemological framework.

This analysis provides a new and more concrete foundations for asserting that the degree to which researchers' behaviors typically adhere to some among the so-called "Mertonian norms" associated with the ethos of "open science" is of critical significance for the collective work performed in the cognitive domain by "invisible colleges." In so doing, it reveals the different implications for the invisible college's short-run and long-run collective epistemological performances, respectively, of the norms affecting the openness of internal communications, and those norms affecting the network's openness to entry of a socially heterogeneous membership. As transparent as are the purposes and the logic of this paper, some background in the philosophy and sociology of science may be useful in order to situate it within the broader context of studies that have concerned themselves with the cognitive content of science and the intellectual and social organization of scientific communities. A brief orientation of that sort is provided by the discussion in Section 1. There I have tried to indicate how the present contribution to "the new economics of science" fits within that larger picture, and also to notice its connections with other recent lines of work by economic theorists -- specifically, those analyzing situations of strategic interdependence among agents that are embedded in local social networks, or whose information states can change only through word-of-mouth communications.

The organization of the paper's remaining sections follows in a straightforward manner the empirical motivations for, and analytical implications of the features of research networks that are captured and caricatured by the model. Section 2 motivates the emphasis placed upon *local* communications through social networks among scientists, pointing out the distinct but complementary roles of tacit knowledge and codified knowledge in the conduct of research, and noticing the importance of personal contacts for the (social) transmission of tacit knowledge. The premised micromotives affecting the behavior of the individual researchers who choose to belong to these social networks is considered in Section 3. Section 4 discusses the nature of the cognitive transactions in which they are engaged. Starting with a highly stylized account of experimental research activity as an iterative process of Bayesian belief revision (section 4.1), I then providing (in section 4.2) an illustrative model in which conformity to the consensus of local peer belief emerges as a rational reputational strategy for individual researchers engaged in interpreting their experimental results. A brief digression is then required, in order to clarify (in section 4.3) where this formulation leaves us in reference to recent controversies between "scientific realists" and "relativists" concerning the social construction of scientific knowledge. Finding the extreme subjectivism of the radical social constructionists unsatisfying, I propose the importation of arguments from evolutionary epistemology as the basic for a plausible resolution, a conciliatory synthesis between the antithetical claims of the two camps in the on-going "culture wars." This clears the way for the core of the network model to be set out in Section 5: the apparatus of graph-theoretic representation of connected local networks of research units, forming an "invisible college" is introduced by section 5.1; the correspondence between the dynamic stochastic process of network interactions specified by the preceding sections, and the Markov random field model known as the "voter model," is described formally in section 5.2, and basic results are presented for the cases of network that can be represented as one- or two-dimensional connected graphs. Section 5.3 discusses those properties of the dynamics of consensus formation in relation to other theoretical results on critical properties of stochastic communication structures — specifically those deriving from the branch of probability known as *percolation theory.*

In section 6 it is shown that the relationship between micro-level behaviors of the research units, on the one hand, and the macro-level cognitive performance of the invisible college represented by the network of local networks, on the other hand, have a quite direct correspondence with the functionalist sociology of science in the tradition of Merton (1973), and the significance of several key Mertonian norms is reconsidered in this light. Section 7 concludes the paper with a sketch of some directions in which it is possible readily to extend the analysis of the basic network model without extensive stochastic simulation exercises, namely, by examining the properties of the equivalent deterministic system. Consideration of the dynamics of endogenous network size adjustments, and the evolution of jointly determined network performance attributes is then quite tractable, as is the examination of the stability properties of growth paths for the "invisible college."

1. The larger context: Philosophy, sociology and the 'new economics of science[1]

The study of science as a cultural activity was for a long time fragmented into two domains. In one camp lived philosophers of many differing stripes, all of whom were preoccupied with the cognitive content of science — which is to say, with epistemological issues concerning the character of "scientific" statements and the way that these were arrived at.[1] With a few notable exceptions, philosophers and epistemologists of science have shared a primary interest in *normative* questions dealing with the nature of the beliefs that people *ought* to hold about their surroundings. The classical approach in such endeavors sought to settle questions on an a prioristic basis, through solely conceptual arguments.[2] A complicating departure from the classical tradition in epistemology, part of a movement to "naturalize" epistemology and the philosophy of science that is closely associated with the work of the philosopher W.V.O. Quine (1953), undertook to inject into such accounts facts about the actual epistemic situations in which people find themselves. Quine (1969) coined the term "naturalized epistemology", arguing that the positivist distinction between contexts of "discovery" and "justification" could be "naturalistically" translated into questions of perceptual psychology and sociology of knowledge, respectively. The former involves studying people's behavioral dispositions to associate words with situations, whereas the latter is studied by examining the communication patterns by which such dispositions become stabilized for a community. In a robust sense, the present essay could be categorized as a contribution to the latter programme, launched from the theoretical perspective of the economic analysis of individuals situated in social communications networks.

To understand how the latter point of departure was arrived at, however, it is necessary to go back to the second of the two domains into which the study of scientific activities was formerly divided. That comprised the territory inhabited by sociologists of science. Members of this clan, following the leadership of Robert K. Merton (1973), were absorbed in trying to understand the social context of modern science, the characteristic institutionalized modes of organization for this particular human activity, and the associated behavioral norms and cultural ethos which guided the participants' behaviors and permitted them to succeed collectively, if not always individually, in their pursuit of knowledge about the external world.[3] Although there were influential contributors to the

[1] See, e.g., Popper (1959, 1963), Quine (1962), Kuhn (1962), Latakos (1970).

[2] This picture of philosophy of science as a domain of concern completely disjoint from the sociology and history of science is deliberately overdrawn, for effect. See the more nuanced, but nonetheless compact account by Fuller (1994).

[3] See, e.g., Crane (1965), Suckerman and Merton (1971), Cole and Cole (1967), Cole and Cole (1973), Cole (1978).

history of science and technology who pointed to the role of external, material conditions affecting the physical phenomena that scientists undertook to study, and the ways in which they sometimes went about it, the canonical investigations in the sociology of science focused upon generic behavioral and organizational problems that were represented to be independent of the specific substantive concerns of the scientific discipline in question, and orthogonal to its cognitive content and claims at particular historical junctures.

Each of these philosophical and sociological tribes, appears collectively to have been disdainfully reluctant (and quite possibly also fearfully hesitant) to assert any claims to the territory staked out by the other; evidently, the two camps found it best to leave each other alone. In consequence of which, only occasionally, and rather tenuously, did their studies attempt to draw links between the ways in which communities of scientists were organized and rewarded, and the ways in which scientific statements were created and accepted within those communities.[4]

In the evolution of these academic transactions, as the case in other affairs, experience sometimes seems to reaffirm the aphorism of the French: *plus ca change, plus c'est la meme chose.* The cognitive content of science and the macro-sociological, organizational structures of scientific communities, once again, could not be held simultaneously within the focus of scholarly attentions. In the 1970's a new generation of sociologists took up the sociology of scientific knowledge — "SSK", as it came to be styled ~ they insisted with ample reason that the cognitive and the social dimensions of science should no longer remain compartmentalized as distinct, specialized fields of inquiry.[5] Instead, the subject matter had to be seen in reality to be inseparable. This was held to be so, because in "discourse" ~ within whose terms it was held possible to analyze everything — social and cognitive contexts are thoroughly inter-penetrating and mutually interactive. Left behind in the border-crossing rush of the SSK movement, however, were the old sociology of science's focii of attention, namely, the institutionalized reward systems, the social norms and the relationship of these to the organization of resource resource allocation within scientific communities.

But, both Nature and the spirit of academic enterprise seem to abhor vacuums, even partial vacuums. On the terrain recently deserted by those who had gone off to build the new sociology of scientific knowledge, leaving behind only a small remnant of organizational sociologists, there soon appeared a vanguard of the "new economics of science." Encouraged by Dasgupta and David (1987, 1988, 1994), these newcomers, although still comparatively few in number, have gone about enthusiastically setting up camp to study the microeconomic resource allocation properties of science reward systems and scientific research organizations. To date, their programme has encompassed both theoretical and empirical analyses of these particular "knowledge producing and distributing

[4] This is a gloss on a very extensive literature, for fuller discussion of which see Ben-David (1990), Callon (1995) and Leydesdorff (1995).

[5] See Barnes (1974, 1977), Bloor (1976), Mulkay (1979), Latour and Woolgar (1979), Knorr-Cetina (1981).

activities", in which application has been found for some of the familiar tools and insights that they had brought from industrial organization economics, game theory, and applied econometrics.[6]

Economists, it must be said, are for the most part quite happy to study the efficiency of resource allocation in producing and distributing goods and services without stopping to inquire even superficially into the specific natures and concrete shapes of those commodities, much less entering into serious discussion of how they come to be differentiated in perception and design.[7] After all, it has been a signal achievement for their profession simply to have recognized and drawn out the implications that follow from the otherwise obtrusive fact that ideas, knowledge and information belong to a category of commodities that really are qualitatively different from traditional economic goods.[8] Considered from that angle, it is perhaps not surprising that the "new economics of science" found it most natural to start by reworking the area of organizational analysis originally ploughed by Mertonian sociology of science, looking at the implications of certain institutional arrangements for allocative efficiency in the production of generic information that acquires a certain measure of reliability; but not troubling itself over the nature of reliability in this context, the details of the way that attribute of information might be acquired, or any of the other issues of socio-cognitive interaction that have occupied the sociology of scientific knowledge.

As there is still plenty left to do in the new economics of science along the trajectories that have already begun to be explored, perhaps it will be thought best to leave well enough alone. On the other side of the argument it must be said that to proceed by completely ignoring the foundation-level connections between the social organization of research and the nature of its information products is likely, sooner or later, to cause difficulties. Indeed, it has already been noticed that maintaining the focus exclusively upon the behavior of the individual researcher, a legacy from the approach of the older sociology of science, detracts attention from the study of collaborations; it encourages "representative agent" modeling of scientists, rather than analysis of role differentiation within larger working groups; it ultimately hinders the hinders collection of empirical evidence about patterns of resource allocation within and between teams, consortia, and institutionalized research networks. Perhaps even more seriously, in terms of its policy a minus, the individualist tradition tends to perpetuate the view that the cognitive performance of scientific communities is largely to be understood by simple aggregation of the independent behaviors of its constitutent members; that is to say, by simple aggregation, rather than by the complex interactions among the agents who form a research community. That view, unfortunately, encourages a policy-disregard for the respects in which scientific researchers' collective epistemological performance in the long-

[6] See, e.g., Arora, David and Gambardella (1995); Arora and Gambardella (1992, 1994); David (1994a, 1995, 1996); David and Foray (1995); David and Flemming (1995); David, Geuna and Steinmueller (1995); David, Mowery and Steinmueller (1992); Gambardella (1994); Trajtenberg, Henderson and Jaffee (1992). Some of the foregoing are noticed in the wider survey of the economics of science by Stephan (1996).

[7] See Bacharach, Gambetta et al.(1994) for an exceptional departure from this tradition.

[8] Indeed, an achievement that only came rather belatedly in the discipline's development, with the pioneering works of Arrow (1962, 1971), Hirschleifer (1971), and Marschak (1971).

run may affected by specific internal organizational features of such communities, including the norms that govern the informational and other transactions among the members. But, these are only some among the considerations that have prompted making a start in this paper on the systematic economic analysis of the relationship between individual behaviors and the cognitive performance of research networks.

1.1 Social network approaches to the economics of knowledge production

Conceptualizing scientific research as a social phenomenon involving actions and transactions conducted within *networks* thus offers a quite natural and appealing direction in which to extend formal modeling approaches in the new economics of science. Among its potential virtues may be that of bringing the latter programme into closer contacts with other social science programmes. Both SSK, and the "sociology of translation" programme of Gallon et al. (1989), have made illuminating use of "networks" as metaphors and tools in many case studies, bibliometric, and related empirical studies of science and technology. This need not stand alone as warrant for the methodological weight that is placed upon the concept of a research network in this paper's treatment of the production of scientific knowledge. Exploring formal models of network dynamics that hinge upon the communication of knowledge, as is done here, may also have some "spillover" benefits for other areas of economics where networks as entities for analysis are emerging with increasing frequency. A brief elaboration on the latter point will serve to establish the plausibility of this second claim.

It is undeniable that popular thought and expression have been encouraged in its use of network metaphors by the startlingly rapid and convergent advances in computer and telecommunications technologies, which recently have been transforming information industries and the users of their services. Whether in conscious or unconscious signification of the now trendier conceptualization of systems of production and social organization, economics too has taken up the study *of networks* — and not only in connection with the fashionable issues of telecommunication and computer network externalities and questions concerning compatibility standardization.[9] Social networks also are now treated formally in modelling many contexts involving strategic interdependence, where these structures are represented as conveying information, and forging mutual trust through repeated transactions ~ even though the connections among the players are highly localized and can be presumed to be effected without sophisticated technological supports.

Several lines of inquiry in game theory also have converged upon local network structures as the terrain for analysing the equilibrium properties of games characterized by strategic complementaries and interactive learning on the part of players. Interest has focused upon the strategic problem that arises when it is assumed that each player interacts directly with only some subset of the entire ensemble — those in the player's immediate "vicinity," and that the players are unable *to adapt their beh*aviors to deal individually with each of their "neighbors"; in such

---

[9] See, e.g., the literature surveyed by David and Greenstein (1990), and David (1993b).

circumstances every player must select a strategy that is uniform with regard to all their neighbors.[10] The conditions under which the dynamics of local interaction games of this kind will give rise to equilibria characterized by correlated beliefs or behaviors have been studied, both in deterministic decision frameworks, and in dynamic stochastic models that make use of results from Markov random field theory to show how local network externalities can lead to *de facto* standardization of technologies and the spontaneous formation of conventions.[11] The present paper will be found to be situated in that broader stream of the recent economics literature.

Whatever the causes,[12] it should be plain from the foregoing that work on the theoretical frontiers of economics has begun serious to embrace some of the essential ideas about the roles of informal social networks that already are thoroughly familiar in the sociology of science (see, e.g., Luhmann (1984, 1990), Ziman (1984), chs. 4,5). Nevertheless, the functioning of such networks, their relationship to the systemic properties of larger scientific communities, and their part in the linkages between those entities and the surrounding technological and economic environment, are all topics that remain to be explored much more thoroughly. Progress in that direction, thus far, has been made through micro-level studies of "knowledge network" structures, and research of this kind recently has begun to engage the attention of economists as well as other social scientists.

Economists interested in the relationships between the worlds of academic research and technological innovation also have made a significant start of their own into socio-cognitive interactions, by using bibilometric techniques of co-citation analysis to measure the extent and the degree of localization in patterns of knowledge transfers ("spillovers") from university research laboratories to industry.[13] These initiatives in the new economics of science have built directly upon the quantitative foundations laid down by Price (1965), Narin (1976), van Raan (1988), and their followers in "scientometrics" who apply bibliometric methods to the study of cognitive structures in science. In that regard, this line of inquiry implicitly accepts the position maintained among by the proponents of the SSK movement, and also by the sociologists in the "translation" school of Gallon et al. (1989), who hold that social networks of research and knowledge dissemination have corresponding linkages in the cognitive domain; that they give rise there to counterpart "connected

---

[10] See, e.g., Anderlini and Ianni (1993), Blume (1993), Ellison (1993), Bala and Goyal (1995), Morris (1996).

[11] See David (1988, 1992), Kirman (1992), David and Foray (1993, 1994), Dalle (1995), David, Foray and Dalle (1996), Ellison and Fudenberg (1995). The approach taken here will be seen to follow in the latter stochastic process vein.

[12] Perhaps the breath-taking advances in electronic communications networks have had a decisive influence in undermining economists commitments to atomistic individualism in the representation of human behavior, where previous opportunities for direct observation had for so long proved unavailing; there certainly is a striking temporal coincidence between the fascinating developments in network technologies and the acknowledgement of the importance of modeling human agents as socially "embedded" ~ entangled in webs of recurring, and inter-personal and intra-organizational transactions with identifiable "others". Technological progress can, indeed, have many unexpected cultural reprecussions.

[13] For this, see the path-breaking work of Trajtenberg, Henderson and Jaffe (1992); Jaffe, Trajtenberg and Henderson (1993); Henderson, Jaffe and Trajtenberg (1995a, 1995b).

clusters of connected nodes" in co-citation networks, formed among papers published in the scientific literature, patent applications, and other "inscriptions".

Beyond effecting these trans-disciplinary contacts, the extension of new economics of science in the indicated direction has something to offer that has hitherto eluded studies in the SSK tradition. In those and other explorations of the cognitive dimension of science, similarly based on networks and inscriptions and the artifacts (to which those inscriptions refer and are referred in turn, there is general difficulty in knowing how to resolve a larger picture from the microcosim of detailed filiations. One wants at the end of the day to know how the workings of these micro-level networks are connected to the properties of the macro-structure of scientific institutions; to the evolution of fields of scientific endeavors that become organized through the extension of social networks and their supporting apparatus of journals, conferences, professional societies, and suchlike. What characterizes social networks in science? How do the behaviors of the individual participants in scientific networks affect the ability of the collectivity to do its work, and does a network's cognitive-domain performance, that is to say, in creating and validating knowledge, affect its opportunities for growth? Questions of this sort crop up repeatedly on the critical "middle ground" between the older and newer sociologies of science. The new economics of science should have something to offer by way of answers. So, it is precisely in that potentially hazardous no-man's-land that the present contribution is meant to stand.

2. Tacit knowledge, local social networks and 'invisible colleges'

I shall begin with a characterization -- a caricature, to be more precise -- of the social context of communications among scientific researchers. Analysis of the economic logic of academic science reward system concurs with the functionalist sociology tradition in studies of the cultural ethos of modern science, by laying stress upon the centrality of the norm of public disclosure of knowledge among those who belong to the Republic of Science (see Dasgupta and David (1987, 1994)). Accordingly, we can agree with Ziman (1984: p. 58) that "the fundamental social institution of science is thus its system of *communication"*. Much attention properly has been given to the conditions affecting formal communications conducted through books, journals, other archival "publications". But, for reasons that are briefly indicated below, these "broadcast" modes of communication occupy a considerably less central position in the following analysis than do the informal and interactive transactions among the members of science communities.

Scientific communal interactions, specifically those among researchers acting *qua* scientists, tend to take place within the circle of those in the same research speciality, who thus form *invisible colleges*. The latter term has been applied in a variety of related ways, sometimes referring to informal (in the sense of non-institutionalized) clusters of scientists collaborating in specific, newly developing research frontiers, sometimes to more extensive collectivities within the same discipline, profession or research area, who are linked by a web of personal (non-anonymous) informational exchanges. The common features of invisible colleges in science are that they remain rather fluid as to membership and size, that internally they are not highly structured, and, especially in today's world of telecommunication technology and cheap air travel, neither are they localized institutionally, geographically, or nationally. Their members may communicate in personal conversations, by circulating working papers and pre-prints, via telephone conversations, and, now also by fax and email messages; whereas, in former times, they would have posted letters to one another.

Within these quite extensive communities, whose members are numbered in the hundreds, there exist rather smaller and communicatively more compact relational entities. These are referred to here as *local social research networks,* or simply as "local networks". Although more tightly clustered, the latter social groupings are research "cliques," rather than organized teams (indeed, then may encompass some or all of the members of different project teams). They also tend to be rather more localized ~ possibly in the dimension of problem specialization, or in those of geographical and institutional proximity which afford frequent occasions for face-to-face communications, or in some combinations of all of the foregoing, and still other social and personal dimensions.

Within the more restricted ambit of a researcher's local network will be circulating many bits of crucial knowledge, about experimental procedures, equipment functioning, data analysis routines -- all of which very often escape being codified and described with complete clarity in published accounts of research procedures and findings. Although the development and circulation of codified knowledge was traditionally taken to be of central interest in philosophical and sociological studies of science, the significance of non-codified, *tacit* forms of knowledge and its roles even in scientific endeavors has recently been gaining more general appreciation. Tacit knowledge, as conceptualized by Michael Polanyi (1966), refers to a fact of common perception that we are aware of certain objects without being focused on them. Lying outside the zone of conscious attention does not make them the less important, however; they form the context that makes focused perception possible, understandable, and productive. Tacit and codified knowledge should thus be viewed generally as complements, rather than substitutes in human cognitive processes. Like other human pursuits, scientific inquiry draws upon sets of skills and techniques that are acquired experientially and transferred by demonstration, by personal instruction and the provision of expert services. Knowledge of this sort may be highly precise and intricate, but it is most typically conveyed as a *gestalt,* and referred to by language and signs that is idiosyncratic, rather than being reduced to constituent elements and operations denoted by standard codes — from which might be assembled programmes of implementation.[14]

The reasons for the persistence of scientific knowledge in uncodified states are various, but it is relevant here to note the argument that these reasons have less to do with epistemological

[14] In such a process of conscious analysis of knowledge developed first in tacit form, some things are likely to be "lost in translation". Consequently, following Dasgupta and David's (1994) usage, David and Foray (1995: 25) say that information is "knowledge that has been *reduced* and converted into messages that can be easily communicated among decision agents"; that "'Codification' of knowledge is a step in the process of *reduction* and conversion which renders the transmission, verification, storage and reproduction of information all the less costly". (Emphasis added.) It is equally appropriate to observe that knowledge is formed through the receipt, interpretation and synthetic integration of information, in the course of which new tacitly understood constructs may emerge, to which those possessing them can attach idiosyncratic meanings.

problems arising in the intrinsic nature of the knowledge involved, than they do with the costs and incentives offered for rendering it in an explicit, codified form.[15] Nevertheless, it remains a fact that many "craft" aspects of scientific practice must be learned in modes of instruction akin to an "apprenticeship", by being afforded opportunities for first-hand observation of how they are done, leading to trials under the guidance and supervision of experts. Otherwise, something like the original process of acquiring mastery of such knowledge has to be repeated *ab initio,* guided and encouraged only by the belief that others have found this to be possible. A striking instance of the "craft knowledge" is deployed in science documented by Harry Collins' (1974) detailed and influential study of the construction of the TEA laser, which concluded that only those scientists who had themselves spent time in the laboratories where a properly functioning laser had been built were able to succeed in reproducing the apparatus at other laboratory sites.

The importance of this kind of "hands-on" experience in many laboratory and facility based research disciplines makes the problem of social communication of tacit knowledge especially germane for the cognitive work of those fields. But even in other fields where codification has been pushed to very high levels, such as pure mathematics, there are areas of advanced work where only initiates with immediate participatory experience in the research itself appear to be capable of following and checking the correctness of proofs developed by others, let alone producing new theorems.

In the modeling exercise undertaken in the following section, it is assumed for simplicity that codified knowledge alone can be effectively transmitted through non-network channels, that is, *broadcast* through a variety of public media that identify the authors of messages but are non-specific with respect to the identities of the recipients. But, no one-to-one correspondence is drawn here between the modalities for message transmissions containing codified information and tacit knowledge. Communications of the sort whose cognitive contents are uncodified (or incompletely codified, craft), and the codified scientific statements to which they relate alike are taken to be emitted locally in the first instance, and to diffuse first within the immediate social network contexts in which they arose.

Both as a matter of formal logic, and in practical affairs, knowledge may be either disclosed to others or kept secret, regardless of whether it exists in codified form or remains tacit.[16] The view presented by the preceding discussion is that for the ideas contained in scientific statements to be understood and rendered operational, researchers must possess the complementary tacit cognitive associations. The local network within which tacit knowledge is shared, thus, need not be considered to be a strategic instrument whose primarily functional role is that of "capturing" and exploiting the

---

[15] Dasgupta and David (1994), and David and Foray (1995) advance arguments for this economic approach, stressing technical and institutional determinants of knowledge producing behaviors. Cowan and Foray's (1996) analysis develops this approach more fully.

[16] See David and Foray (1995), where three distinct dimensions are recognized as defining a space in which knowledge products can be located: the codified-tacit axis, the disclosure-secrecy axis, and the public -private property axis.

benefits of tacit knowledge for its members. Indeed, I take just the opposite position and emphasize the benefits to the individual researchers of cooperative, reciprocated transactions in tacit knowledge possessed by specialists whose expertise enables those cooperating to increase their chances of solving complex, multi-step problems sooner than would be the case were they to work in isolation.[17] In keeping with this, we may suppose that local networks are not autarkic; by having some members in common with other, similarly local social groups, they can be rendered inter-communicative. Thus, complementary packets of codified and tacit knowledge, along with explicit and implied conjectures about promising lines of scientific inquiry, eventually may percolate outward from particular local networks and so become diffused throughout the wider community of researchers that constitutes the "invisible college".

Even though codified information can be made accessible more quickly and cheaply through impersonal broadcast channels, in order to utilize much of that information directly, the view of the world presented here is one in which researchers must obtain the complementary tacit elements through their social network connections. That alone would make the conditions affecting the operation of local social networks the factors proximately constraining the speed and extent of effective diffusion of new discoveries and inventions. Indeed, it could be argued that recognition of the critical role of tacit skills and learning mechanisms ~ which has been such an a major theme in recent sociological research on "laboratory life"[18] -- should be read as justifying an even more stark caricature of a scientific communities or invisible colleges: nothing but an informally organized network of local social communications networks, or an interconnected set of "neighborhood clubs." But, it is neither useful nor necessary to go to the extreme of shutting one's eyes completely to the existence in modern scientific communities of a supporting apparatus of institutionalized procedures for publication via broadcast media, for the organization of global discussion fora, and the like. Indeed, it seems more plausible, especially, when one comes to try to formally characterize the way that new ideas are generated, to assume that at each moment in time the members of the invisible college are able to access the common archival knowledge-base containing scientific statements whose reliability is no longer a matter of doubt and active controversy.

3. Information-pooling, reputation and micro-motives for network membership

Contributors to the macro-sociology of scientific institutions have described the role of modern "invisible colleges" in channeling informal communications among scientists and forming the social substratum upon which the collegiate reputational reward systems of academic, open science are grounded.[19] Invisible colleges thus fulfill many functions. As seen by Derek de Solla Price,[20]

---

[17] This supposition is consistent with the arguments set out in the following section (3): small cooperative networks of information sharing can be supported ~ as an equilibrium of a game ~ among researchers, because some (enforceable) degree of cooperative behavior furthers their self-interests, even though they may be racing for *priority*.

[18] See, e.g., Collins (1974), Latour and Woolgar (1979).

[19] I have argued elsewhere that the central problem of informational asymmetry in the late sixteenth century patronage system was being addressed by the formation of networks of correspondence and the flowering of "invisible colleges" among mathematicians and the experimentalists in the new science. See *David* (1995). Such networks

they confer upon each scientist "status in the form of approbation from his peers, they confer prestige, and above all, they effectively solve a communication crisis by reducing a large group [such as the entire population of accredited scientists in a field such as economics, or chemistry] to a small select one of the maximum size that can be handled by interpersonal relationships"[21] But, insofar as these functions of an invisible college rest upon the members' sharing knowledge that might confer personal advantage if kept for their private use, subscribers to the theory of rational individual utility-maximizing behavior will want to know what induces researchers to seek admission to these "clubs," and conform by-and-large with club rules, by sharing knowledge that they have not yet divulged, and may never divulge by means of impersonal broadcast channels. What, then, do we suppose are the private incentives offered to the individual scientist, outside of a institutionalized setting in which such conduct is has been made a formal obligation (publish or perish!) of career advancement? Or is it necessary to presuppose the existence of some prior process of socialization into the role of "scientist" that implants in the individual an acceptance, and perhaps even an active subscription to professional norms of conduct in such matters?

Two answers come to mind that avoid such presuppositions. First, there is the obvious "exchange value" of information that is implicit in Price's (1963, 1986) discussion ~ except that some of the knowledge involved may be "expertise" conveyed by direct demonstration rather than being reduced to "information." The divison of intellectual labor confers advantages to those who need not work in isolation, but, instead, are able to draw upon the still-unpublished findings and tacit expertise of others in identifying or solving some particular problem. Access to such networks of assistance must in effect be "purchased up front," by proffers of worthwhile "material," or gained by accreditation which holds out credible promises of a future reciprocation of valuable intellectual help received at the outset.

Second, a certain status is conferred by being accepted as a "correspondent" by already established members of the scientific community. This carries some "signaling value" for the individual, as a mark of recognition and. approbation that can be displayed to advantage in transactions with third parties. The "price"extracted for elevation to membership in networks peopled by more prestigious scientists may be high ~ either in terms of attribution of credit for discoveries or techniques that could be claimed exclusively, or deflections of one's efforts to working on problems less centrally consequential for the advancement of one's own research

---

augmented by an expanding volume of printed pamphlets and treatises, provided an arena in which challenges could be issued, contests and competitions could be staged, and collegiate reputations could be both secured and widely broadcast. Challenges and contests among Europe's mathematicians were not an innovation of the seventeenth century, however. See e.g., Bell (1937), Boyer (1975), pp. 311-12, 341.

[20] Price (1976 :pp. 83ff) re-introduced and conceptually extended the seventeenth century term "invisible college," coined originally by Robert Boyle in referring to the small group of natural philosophers whose intellectual transactions with one another anticipated the formation of the Royal Society.

[21] See Crane (1972) for an early elaboration of Price's view of invisible colleges as vehicles for diffusing scientific information, an approach that sociologists of scientific knowledge have criticized as neglecting the social interaction processes through which tacit knowledge is formed and conveyed. Merton [1986] discusses the relationships between the concepts of reference group, invisible college, and deviant behavior in science.

program than to the work of others. Except where the reward system of the scientific community is characterized by strong personal patronage mechanisms rooted in teacher-student relationships, one should expect to find neophyte researchers, prompted by the first category of motives, entering into regular informal knowledge exchanges with disciplinary "peers" who are of approximately on a par with them in achievements and reputational status.

Unlike signaling value, the benefits of participating in a collegial research network (extending beyond the immediate group or team with which the individual is currently engaged) can be thought to have a substantive basis for both researchers and their patron-employers. This is so, whether or not the final disposition of the knowledge that is acquired will be public disclosure or private exploitation in some directly productive activity. Thought the medium of their "network" the individual members have at their disposal assistance in discerning unrecognized anomalies and research opportunities, and help in locating solutions to unanticipated problems that tax their own, or their research team's cognitive capabilities. The collectivity of the network thus operates somewhat like a "patent-pooling" arrangement without the patents—a loose coalition, in which some degree of non-compliance with the norm of mutual help will occur, and so yield an equilibrium in which the "common pool" is degraded by the private reservation of certain information. Nevertheless, access to help from "peers" remains a valuable asset for the individual.

To restate the thrust of the foregoing discussion, it is possible that cooperative behavior within a limited social sphere can emerge and be sustained without requiring the prior perfect socialization of researchers to conform, altruistically, to the norm of full disclosure and cooperation. This is a rather straightforward instance in which insights from the theory of repeated games are applicable to explaining cooperative behavior among potentially rivalrous researchers.[22] To sharpen this point, one should start simply by considering two researchers working towards the same scientific goal which involves the solution of two sub-problems, and suppose that each has solved one of the problems. Once each gets the other solution, it will be a matter of writing up the result and sending it off to a journal for publication, or reading the paper at a prestigious professional gathering —the first to do so being awarded priority. Now suppose, further, that the write-up time is determined by a random process, and that if both get the two halves of the problem at the same moment, each will have the same (one half) probability of being the first of the pair to submit for publication. Whether the winner will be awarded priority will depend, however, on whether or not some other researcher has obtained the full solution and sent it off already. The question is: should the first researcher to get any part of the solution follow the strategy (S) of sharing that information with the other one, or should she adopt the strategy (W) of withholding?

If, without prior communication, they play the strategy pair (S,S) they can proceed immediately to the write-up stage; if they play (S,W) the second member of the pair will be able to proceed to the write-up, and the opposite will be true if they play (W,S). Should they both withhold

---

22 The following parallels material developed in Dasgupta and David [1994].

(W,W), they must both spend further time working on the other problem. It is evident that if they are only going to be in this situation once, the "winner-takes-all" rule of priority alone will induce each of them to withhold, and they will end up (collectively, if not individually) at a relative disadvantage vis-a-vis other researchers who are hurrying to publish. If nobody else has the full solution yet, society also will have been forced to wait needlessly, because each member of the pair has a dominant (private) strategy of withholding what she knows. This game has the structure of a classic two-person "Prisoners' Dilemma."

It is well known, however, that an escape from the pessimal outcome and the attainment of the superior, cooperative equilibrium is possible under certain conditions: if the game is played repeatedly, if the future is not discounted too heavily, and if the players expect the other member of the pair to remember, and punish, their failure to play cooperatively by sharing. Indeed, there is a so-called "folk theorem" to that effect.[23]

Yet, that is not the end of the matter. As there are other researchers in the picture, we should really be considering an n-person game, involving the solution of an m-part problem, where n > m. Now the question of sharing information becomes one of sharing not only what you have learned yourself, and what you have been told by others. It obviously is advantageous to be in a group among whom information will be pooled, because that will give the group members a better chance of quickly acquiring all m parts of the puzzle and being the first to send it in for publication.[24] But, on the other hand, if there are individuals who behave opportunistically by exchanging what they

---

23

In addition, the value in the future of developing and maintaining a good reputation for sharing has to be large in order to discipline the self-interested researcher into adhering to the sharing mode of behavior in the current period. If repetitive play comes to an end, or if the future is valued only slightly, cooperation will unravel from the distant terminal point in the game, right back to its inception. For a non-technical introduction to the literature on the repeated Prisoners' Dilemma and its broader implications, see Axelrod (1984). The "folk-theorem" of game theory holds that (if future payoffs are discounted by each player at a low rate) in the "super game" obtained by repeating a finite, two-person game indefinitely, any outcome that is individually rational can be implemented by a suitable choice among of the multiplicity of Nash equilibria that exist. See Rubinstein (1979), (1980), and Fudenberg and Maskin (1984).

24 Campbell (1994 :p.l43) attributes to Michael Polanyi the analogy of a jigsaw puzzle-working group. Polanyi suggested, as a functional justification for the norm of open information sharing in science, that it would be "optimal" for the collective solution of the puzzle that every member of the group see the whole board and the pieces that each of the others was holding. Such heuristic examples may suffice to make the point about the advantages of cooperation, but they cannot substitute for empirical investigation of the structure of communications channels in small (local) networks that will be most effective. Some work in the field of social psychology as been devoted to this question: Harold Guetzkow and Herbert Simon [1955], for example, reported experimental results in group problem-solving, connecting the members of six-person teams in three different forms of communication networks, following the patterns previously studied by Alex Bavelas [1950]. These were (i) the "ring" structure, having as many channels as nodes (n=6), each node being linked to a node on each side; (ii) the "star" or hub-and-spoke structure, in which (n-1) = 5 channels suffice to link one node to every one of the others in binary couplings; (iii) the "fully" connected network, which requires n(n-l)/2 =15 channels. Guetzkow and Simon found that when they distributed fragments of a poker deck among the participants and asked them to select the strongest poker hand from among the cards they held, the hub-and-spoke structure performed best, regardless of ability of person positioned at the hub; the ring was worst. Interestingly enough, in repeated play it was observed that the fully connected structure tended to evolve, through disuse of channels, into a hub-and-spoke pattern of communications. It may be noticed that the latter pattern is precisely the one that is assumed here by the stochastic opinion-polling process in the "voter model" of local social network interactions.

have learned from one group for information from people outside that group, but do not share everything they know within their group, they can expect to do still better in their current race for priority of publication. However, because others would see that such "double-dealing" will be a tempting strategy, cooperation will be unlikely to emerge unless "double-dealers" (who disclose what you tell them to third parties, but don't share their knowledge full with you) can be detected and punished. What is the form that retribution can take? Most straightforward will be punishment by exclusion from the circle of cooperators in the future; and even more severely, not only from the circle that had been "betrayed" but from any other such circle. This may be accomplished readily enough by publicizing "deviance" from the sharing norms of the group, thereby spoiling the deviator's reputation and destroying his acceptability among other groups.[25]

What, then, is the likelihood that this form of effective deterrence will be perceived and therefore induce cooperative behavior among self-interested individuals? If a group, i.e., "a research network" numbering g players ($g \leq n$) is large, identifying the source(s) of "leaks" of information and detecting instances of failure to share knowledge within it, will be the more difficult. It is worth remarking that the power of a large group to punish the typical deviator from its norms by ostracizing him tends to be enhanced by the higher probability that all those individuals with whom potential deviators will find it valuable to associate are situated within the group. In other words, the expected loss entailed in being an "outcast" is greater when there is only a fringe of outsiders with whom one can still associate. But, this consideration is offset by the greater difficulties the larger groups will encounter in detecting deviators. Smaller groups have an advantage on the latter count, and that advantage also enables them to compensate for their disadvantage on the former count. The more compelling is the evidence that a particular individual had engaged in a "betrayal of trust," the more widely damaging will be the reputational consequences for the person thus charged. Hence, unambiguous detection and attribution of deviations (from recognized norms regarding the disclosure and non-disclosure of information) augment the deterrent power of the threat of ostracism that can be wielded by any group that remains small in relation to the total population of individuals with whom an excluded group-member could form new associations.

The foregoing suggests that small cooperative "networks" of information sharing can be supported among researchers, because cooperative behavior furthers their self-interest in the race for priority, and denial of access to pools of shared information would place them at a severe disadvantage vis-a-vis competitors.[26] Does this imply that the normative content of Merton's

---

25 See Greif (1989), and Milgrom, North and Weingast (1990) for analysis of repeated games of incomplete information that have this structure.

26 These "circles" or "networks" which informally facilitate the pooling of knowledge among distinct research entities on a restricted basis can exist as exceptions to both the dominant mode of "public knowledge" characterizing Science, or the dominant mode of "private knowledge" characterizing Technology. Thus, von Hippel (1988) and others have described how firms in fact tacitly sanction covert exchanges of information (otherwise treated as proprietary and protected under the law of trade secrets) among their engineer-employees. Participants in these "information networks" who accepted money or remuneration other than in kind, most probably, would be dismissed and prosecuted for theft of trade secrets.

communalistic norm of disclosure is really redundant, and plays no essential role in fostering conditions of cooperation among citizens of the Republic of Science? Not so! For, it can be seen that networks of cooperative information sharing will be more likely to form spontaneously if the potential participants start by expecting others to cooperate, than if they expect "trust" to be betrayed; and cooperative patterns of behavior will be sustained longer if participants have reason to expect refusals to cooperate will be encountered only in retaliation for transgressions on their part.

Furthermore, detection of deviant behavior warranting punishment, and implementation of the retribution of ostracism from a particular network, will have more broadly damaging reputational consequences when the norms of behavior involved (i.e., the "custom" within the network in question) is common knowledge, and part of the shared socialization among all the potential members of networks. It is evident from this that even if the process of socialization among scientists were weak and quite imperfect, the common "culture of Science" makes it much more likely that the rule of priority will not tempt individuals to engage in opportunistic withholding of knowledge, but, will engage the self-interest of researchers in reinforcing adherence to the norm of disclosure -- at least among those restricted circles of colleagues that form his or her local social network.

4. Cognitive communications: Beliefs, Bayesian learning, and scientific consensus

What is it that is being communicated in these transactions? For purposes of simplifying the following analysis, envisage two sorts of cognitive processes as taking place within the local social network structures. Information in the form of codified statements can be passed among the individuals (or among the teams, and other multi-person research organizations that institute communication policies binding upon their members), by the act of one of them sending a message, or "sharing" a piece of knowledge, and the other receiving, or "reading" it. The cognitive substance of the generic message-transaction comes in two parts. The first component contains or otherwise identifies a particular scientific *statement* about the natural world, or about the design of a measurement instrument or other artifact, or an experimental procedure, or a logically connected "bundle" of such statements.[27] The message's second part conveys the sender's present state of belief as to the "reliability" or "unreliability" of the accompanying statements.[28] I do not speak here either of "truth," or of degrees of reliability. At any moment the researchers (acting either in a team

[27] One may imagine that the cognitive content of such "statements" is produced by combining the texts of codified information already, or soon to be broadcast through public channels, and further comments of a qualifying , or clarifying sort resembling "marginalia" on those texts. The latter may be linguistically expressed, or conveyed by actions ('gestures"), and reflect the tacit knowledge of the sender ~ including particular, idiosyncratic rendering he or she may give to published "texts".

[28] It can be supposed that the channels of communication are multiplexed, and capable of bundling (or "packaging") cognitively interrelated propositions, so that it would be possible for them to handle more or less concurrently a flow of numerous messages of the foregoing sort about many distinct, and cognitively independent scientific statements. Each of those problematics could be assigned to its own "layer" over the network, and the resulting information-processing architecture thus would be enabled to simultaneously execute multiple consensus building routines in parallel.

organization or as a solo investigator) are of only one mind about the reliability of the particular statements they are emitting or transmitting, and so they mark them either as "acceptable", or "unacceptable". But their minds are open, in the sense that from moment to moment they can find cause to revise their labeling of the same statement(s).

This revision process is precisely where the cognitive content of the messages' second component (beliefs) comes into play. Two stories could be told at this about how scientific inquiry is conducted in the world of our caricature. In both accounts the researchers are seen to be engaged in revising their orientations with regard to the reliability or unreliability of statements that have reached them in ways that permit them to be influenced, if not completely "persuaded" by the information they receive concerning the beliefs held by colleagues belonging to their local social network. They also are mutually compatible, so that for our purposes it does not seem that one has to choose between them.

4.1 Beliefs and Bayesian experimental strategies

In the non-strategic story the researchers are assumed to be engaged in bounded Bayesian information processing, drawing observations from their own experiments, and the reports of similar observations made by others belonging to their network. Inasmuch as they are assumed to be following a common epistemological strategy (and referring to the same subjective probability thresholds when declaring a conjecture about an unobserved state of nature to be acceptable, or not) they view the *a priori* beliefs conveyed by neighbors as having been shaped by a Bayesian revision process, and therefore influenced by both the pooling of beliefs communicated by other neighbors, as well as being occasionally updated by the formation of the likelihoods of their own experimental observations.

The expressed beliefs of local colleagues about the acceptability or non-reliability of propositions con observations, are assumed in this story to weigh heavily in the process just described. Remember that these beliefs have been communicated by the folk with whom they regularly are interacting. So, another appeal a "folk theorem" provides some warrant suppose that even if the individuals involved had not been socialized to report accurately their current scientific views, the prospect of extended repetitive play would have led them toward some Nash equilibrium characterized by "cooperative" rather than opportunistically deceitful strategies in this regard.[29]

But, it is now necessary to take on board some of the doubts that sociological, philosophical, and psychological studies have raised regarding the degree to which scientific progress occurs through the experimental invalidation of conjectural propositions, especially when strong supporting

---

[29] See David (1995) for further discussion of the emergence of cooperation from "repeated play" in this context. Since individuals are free to revise their beliefs, discovery of "misrepresentation" would not be quite straightforward; when challenged on such a matter, a researcher would have to be have to acquit him- or herself by back-tracing the sequence of events that led from one belief orientation to its opposite.

beliefs surround the proposition in question.[30] The 1970s were marked by the rising chorus of doubts among philosophers and historians of science about the basis for the view that experimental falsification of hypotheses played a crucial role directing the progress of scientific knowledge. This scepticism emerged from the sociological and psychological critiques advanced by Kuhn (1962/1970), Lakatos (1970) and Feyerabend (1975), and reinvigorated the logical problems that the earlier writings of Duhem and Quine posed for Popper's (1959) account of the process of scientific discovery. As Quine (1953:45) summed up the problem: "Any statement can be held true, come what may, if we make drastic enough adjustments elsewhere in the system". Thus, it might be argued that whereas the whole edifice of scientific theory was at risk of empirical refutation, particular theories or hypotheses could escape experimental falsification.[31]

Thomas S. Kuhn, in *The Structure of Scientific Revolutions* (1962/1970), held that major changes in scientific understand occurred through shifts in paradigms, and that although competing paradigms each defined a context of "normal science" within which certain kinds of experiments played an important role in the evaluation of particular theories, alternative paradigms, essentially, could not be put to empirical tests on common experimental grounds. Lakatos (1970:119) also contended that experimental evidence would not suffice to reject a theory: "...no experiment, experimental report, observation statement or well-corroborated low level falsifying hypothesis alone can lead to falsification. There is no falsification before the emergence of a better theory." Feyerabend (1975) went farther, portraying changes in scientific opinion as propaganda victories by one group of scientists over others, and denying the validity of the modern scientific enterprise's claims to possess a superior method of gaining knowledge through experimentation.

In this climate of reappraisal of the traditional emphasis placed upon "crucial experiments" in the progress of science since the 17th century, historians of science cast a skeptical eye upon putatively epic events of empirical refutation and found them to be largely "mythical". Not only were doubts expressed as to the very occurrence of the experiment whereby the authority of Aristotelian mechanics, supposedly, was undermined by Galileo's finding of essentially the same rate of fall of two unequal weights dropped from the Tower of Pisa; modern replication showed that the observable difference in the speeds of the objects over their 200 foot descent would have been too large to justify Galileo's reporting it to have been negligibly small.[32] Similarly, revisionist history of science now instructs us ~ through Worrall (1976) ~ that Newton's corpuscular theory of light was

---

[30] Little will be said here about some of the early doubts about validation processes in science that were raised by research on the social psychology of perception and observational reporting, but Campbell (1994) offers a useful review from the social psychology perspective of sociological and philosophical literature noticed here.

[31] See Franklin (1986:4, 106) for the latter reading; Harding (1976) for discussion of the so-called "Duhem-Quine problem" regarding the possibility of scientific refutation, and the original essays of P. Duhem and W. V.O. Quine.

[32] Franklin (1986:p.2, n.7), citing the replication study by Alder and Coulter (1978), goes on to point out that an Aristotelian could readily have modified the theory to accommodate the experimental data.

not overturned, and the wave theory established by the experiments of Thomas Young, since corpuscular explanations were available for both interference and diffraction.[33]

To be sure, counter-arguments can be mobilized against the extreme position that scientific changes are wholly theory-dominated, and that rather than being driven by the force of experimental evidence, it is sociological forces, and even political and economic considerations external to the logic of particular scientific questions, that impinge in crucial ways upon the fortunes of contending theories.[34] There have been instances in which, clearly, experiments did prove "crucial" in over-turning a prevailing theoretical model. A striking modern case in point is that of the experimental results prompted by Lee and Yang's (1957) famous theoretical paper, suggesting that the theory of parity conservation (mirror symmetry) ~ which at the time was accepted by the physics community as characterizing the strong and electromagnetic interactions ~ might not hold for the weak interactions. Yet, the careful historical re-examination by Franklin (1986: Ch.l) concludes that this was perhaps an atypical case, and a really rather special set of circumstances afforded the experimental evidence the leverage they exerted.[35] The discovery of nonconservation of parity, therefore, may serve less to exemplify the "crucial" role of experiments, than to expose the special nature of the circumstances that would make it likely for experimental results to prove decisive in rapidly altering a scientific consensus.[36]

---

[33] Moreover, the early (pre-Fresnel) wave model could not account for the rectilinear propagation of light, which was as troublesome for that theory as interference was for the corpuscular model. See Worrall (1976), discussed by Franklin (1986:pp.2 n.8).

[34] Franklin (1986) reviews recent historical research on experimental physics (by Galison, Pickering and others), in addition to presenting his own case-studies of the role of experimental evidence. These include Millikan's 1913 "oil drop" experiments which established the quantitization of electrical charge and measured the electron's charge; and the experimental tests reported in 1957 papers by Wu et al.,, by Garwin et al., and by Friedman and Telegdi, which brought rapid acceptance of Lee and Yang's (1957) theory of parity nonconservation.

[35]The list itself is instructive: (1) the contending theories were mutually exclusive and exhaustive ~ they belonged to two qualitatively delimited classes, i.e., those that satisfied symmetry and those that did not, rather than simply making differing quantitative predictions; (2) a reinterpretation of the experimental results (in the logic spirit of the Duhem-Quine gambit) — contending that they did not establish the absence of symmetry specifically in the weak interactions, because parity might not be conserved in the "background" conditions which the experiment presupposed ~ would have led to the even more radical rejection of symmetry in all the forces of nature; (3) while there was extensive experimental evidence consistent with parity conservation in the other forces, as Lee and Yang had pointed out, there was no such evidence about the weak interactions; (4) the operation of the apparatus that was used depended only upon the strong and the electromagnetic interactions, and so, while the experimental design was "theory-laden", the results could not be compromised by circular dependence upon the theory that they appeared to refute; (5) there already existed "framing conditions" of a sort that could be seen as providing a precedent for the partial violation of the parity conservation law ~ that is, for symmetry to be violated in only one class of interactions. This was the theory of the conservation of strangeness, put forward by Gell-Mann and Mishijima during the immediately preceding period (1953-55). It accounted for certain well-established experimental observations by the conjecture that particles known as K meson and hyperons possessed a quantity called "strangeness" that was conserved in the strong, but not in the weak interactions. See Franklin (1986: 32-33, 39).

[36] Indeed, as Franklin (1986: Ch.2) describes, there were experimental results in the 1930s on the scattering of fast electrons by metals that, in retrospect, provided evidence for the nonconservation of parity in the weak interactions. The asymmetry findings were recognized at the time as anomalous because they were part of a larger body of experimental results that were at variance with the predictions of an other-wise well-corroborated theory of Dirac; but, because there was no theoretical context in which parity nonconservation evidence could be read, the latter became lost from view.

More generally, close examination of the ways in which runs of experimental physics results are generated, reveals how the latter can sometimes appear to imply parameter magnitudes that are conditioned by either theoretical expectations, or by the initial values that were obtained in contradiction of received theory, but, after extended replications, may undergo significant alterations causing them to converge to the theoretical expectations. Franklin (1986: 235-236) writes:

> In any experiment, the sources of error, particularly systematic error, may be hidden and subtle. This is particularly true of the technically difficult experiments.....The question of when to stop the search for sources of error is then very important. One psychologically plausible end point is when the result seems "right" (i.e., in agreement with previous results).

One explanation of the role of theoretical expectations in guiding experimental outcomes in the evaluation of physical constants, is that attributed to the physicist E.O. Lawrence:[37]

> In any highly precise experimental arrangement there are initially many instrumental difficulties that lead to numerical results far from the accepted value of the quantity being measured. It is, in fact, just such wide divergences that are the best indication of instrumental errors of one kind or another. Accordingly the experimenter searches for the source or sources of such errors, and continues to search until he gets a result close to the accepted value. *Then he stops!* [Emphasis in original.] But it is quite possible that he has still overlooked some source of error that was present also in previous work. In this way one can account for the close agreement of several different results and also the possibility that all of them are in error by an unexpectedly large amount.

Such attitudes, and the experimental experience that they illuminate, constitute what has been referred to as "intellectual phase-locking" (Franklin 1986: 236).

Should we then believe that experimentation cannot offer a check on theoretical consensus? Are agreements among scientists so powerful a force for conformity that they compel convergence of the experimental evidence to the prevailing "truths"? To assert this would be to ignore the mechanisms that are thought to be reliable in the long run to effect escape from serious experimental errors. The performance of quantitative measurements by "different" methods, and especially the repetition of measurements when new and more powerful experimental arrangements and instrumentation becomes available, are perhaps the most oft-cited mechanisms of auto-correction for "bandwagon effects" that would otherwise perpetuate experimental errors. But the reward system within which the experimenters work must repeatedly impel the search for such new techniques, and their swift engagement when they do become available.

In the case of an important magnitude, such as the fundamental physical constants, there are, indeed, many repetitions of experimental measurements. According to Franklin (1986:236):

[37] The reference in Franklin (1986: 239) for the following quotation is to R. Birge, "A Survey of the Systematic Evaluation of the Universal Physical Constants", *NC* [Suppl.], 6 (1957):51.

> "The histories of such measurements, particularly those for *e*, the charge on the electron, and c, the speed of light, often do show such major changes when a new, and presumably better, technique is used."

So, there is a selective element of built-in correction to the research process, because when experimental techniques change there is an opportunity to score a scientific "coup" by establishing a new and different value from the one previously accepted when the magnitude in question is "important" and widely relied upon. Alternatively, so long as there is some persisting acknowledged discrepancy between the theoretical expectation and previous experimental findings, there is hope of achieving recognition from their reconciliation in one way or another. These aspects of the science reward structure will thus set some bounds upon unintended tendencies towards conformity with prevailing theoretical expectations.

4.2 Conformity to local peer consensus as a reputational strategy

Introducing some further strategic considerations, however, may provide a direct rationale for the persuasive power that a prevailing consensus for, or against, a given proposition might exert in the formation of an individual researcher's reported beliefs about its validity. Unlike the foregoing arguments from the philosophy and history of experimental science, this story is meant to more generically "economic" in its appeal. In the spirit of arguments suggested by Dasgupta and David (1994), we may suppose that a scientist working in a collegiate reputational reward system will consider the nearer-term reputational consequences of current actions (including expressions of scientific opinion), as well as considering long-term payoffs possibilities in the form of lasting fame for "having gotten it right". Whether one will be found to have been "right" in the judgement of the invisible college depends upon the alignment of one's recorded (or remembered) beliefs in relation to the consensus that existed at the time among members of the local network to whom those beliefs were disclosed, and also in relationship to the global consensus that may emerge within the invisible college as to the "truth" of the conjecture in question. It is quite possible to envisage a structure of "expected reputational payoffs" that makes it a dominant strategy to be found to have had beliefs conforming with those presently held by most of one's local network, even when the ethos of the lone scientific hero would accord maximum kudos and immortal glory to a researcher who had not conformed with local peer opinion and whose beliefs eventually came to be shared by an overwhelming segment of the discipline at large.

Consider the following example, in which we denote conformity with the preponderance of scientific opinion in one's c, and disagreement with that consensus by d. For expositional convenience we will consider the case were the prevailing consensus hold a particular statement to be "reliable/true", denoted as R. (One can treat symmetrically the case in which the consensus among the agent's reference groups holds the statement to be "unacceptable/false", not-R or, equivalently W.) The consensus that can emerge eventually in the global network, i.e., the limiting configuration of beliefs among members of the invisible college, may either hold the statement to be "Reliable Knowledge", denoted by R, or not, denoted by W ("Wrong"). But we suppose that this eventual determination cannot be known with certainty when the researcher is deciding which opinion to give on the matter. It is in that sense an unobserved "state of (social) nature". If the individual researcher treats the local network as the reference group whose esteem matters, we then have the following notation for possible states to there will correspond "reputational payoffs" to the representative researcher:

{c, R|R}= b 1: being right, with the crowd;

{c, R|W}=b2: being wrong, with the crowd;

{d, R|R}= b3: being in a minority and wrong;

{d, R|W}=b_4: being in a minority and right.

A suitable general payoff structure for an epistemic community is one that assigns greater value to a individual research who is found to be "in the right" than to the individual who embraced the "wrong" view. But it is also plausible that being wrong in a crowd will be deemed not as bad (for her subsequent reputational standing) as being found to have been wrong more-or-less on one's own; and being "lonely yet right" was reputationally more glorious than being correct among a a crowd of one's peers. This scheme of valuation of the outcomes of the game against (social) nature corresponds to the condition : b_4>b1>b2>b3.

Now, let denote the individual's subjective probability assigned to the global consensus forming eventually on R. Where does this subjective probability come from? There will be some processes of consensus formation for which is the case that the best probability estimate of the configuration of beliefs in the college at large converging eventually on R will be the distribution of beliefs prevailing at the intial point.[38] But, even if we have warrant for assuming such is the case, the preceding discussions of the role of local social networks gives grounds for assuming that the individual researcher's field of information on this matter is restricted to the state of beliefs among members of her immediate, local network. So, the proportion of latter who presently hold say "R" when canvassed could thus provide the individual's probability estimate of the eventual "truth" towards which the consensus of the invisible college as a whole must converge, . Then, with a little algebra it follows the condition $[(b_1-b_3)/(b_4- b_2)]>[(l-)/]$ is sufficient for the strategy of conformity, c, to be the one that will maximize the individual researcher's expected reputational payoff. Conformity with the prevailing local network majority opinion implies that when the best estimate of is substituted into that condition, it simplifies to : $(bi-b_3) > (b_4-b_2)$.

This says that for conformity to be generally rationale it will be sufficient for the term $(b_4- b_2)$, to approach a small positive value. This is to say, deviance from the local peer consensus should not

[38] It will be seen below in section 5.2, that the stochastic process known as "the voter model" is one system for which this is true, as *Property 3* of the model states that — for one- and two-dimensional network graphs — the system achieves perfect correlation of beliefs on one or the other of the binary options, and the probability that can be assigned ex ante to either of those outcomes corresponds to the initial distribution of beliefs. The "voter model" is constructed by symmetricallly assigning to each node of the graph (each researcher) a belief revision strategy that reduces to probabilistically following the local consensus, and the illustrative micro-model provided here shows the possibility of here is consistent with that — since that strategy is seen to be individually rational under the specified payoff structure and the restriction that ties the individual's subjective probability estimate to the distribution of beliefs in her local social network.

be too strongly rewarded even when it turns out eventually to be "justified", whereas being mistaken should not deal servere damage to one's reputation when one has done so in plenty of company.

Keeping matters simple, we may postulate that such conditions do obtain and build upon the implied micro-level behaviors. In other words, can suppose that all the researchers are symmetrically situated in their respective local networks, and that each is (rationally) following a decision routine of the following sort: at a random intervals in time the opinion-messages emitted by others in the local network are polled, and if there is unanimity among them, either on the reliability or the unreliability of a particular proposition, the researcher at the polling node will accept the local consensus and accordingly adapt the messages she emits concerning the state of her belief about the validity of the proposition. When a disagreement is found by polling within the local network, the researcher takes a quasi-Bayesian approach, and probabilistically forms an opinion about the true state of the natural world by a procedure equivalent to tossing a coin whose loading mirrors the division of opinion among the set of neighbors. In this procedure, implicitly, equal influence is accorded to the opinions of the local reference network members.

How does this work? Consider these two trivially simple illustrative situations:
(i) If the social network was composed of three members then disagreement among them would take the form of an even split. So, the index-agent who was polling her "neighbors" would toss a fair coin (with a 1/2:1/2 loading) to arrive at an unequivocal position on the question under investigation. (ii) Alternatively, if the social network had four members disagreements among the index-agent's "neighbors" would take the form of 1:2 or 2:1 splits on the question of reliability. Accordingly, the index-agent would form an "opinion" on the question by tossing a coin loaded (1/3: 2/3) for reliability when the local consensus was against that view, and another coin with loading (2/3: 1/3) in favor of the statement's "reliability" when the reference group's balance of opinion was in favor.

The toss of the (local opinion-weighted) coin envisaged in this routine could be interpreted as a highly abstract representation of the conduct of an inherently ambiguous experiment or observational procedure, in which the "reading" of the outcomes would be strongly shaped by the experimenter's or observer's priors as to the validity of the hypothesis under examination. Only when opinion in the local peer group is quite evenly balanced would the testimony of the experiment exercise much leverage upon the experimenter's reported belief, and then, the conclusions that will be drawn from a "face value" reading of the results remains uncertain.

At this point two remarks may need to be made about this schema. The first is directed to readers who have noticed that these hypothetical scientists cannot be very bright if they are using this cumbersome procedure for gathering and processing the information about beliefs prevailing in their local network. In case situation (i) above, it evidently would be far simpler to equi-probably sample one of the two others in the network, and adopt his or her orientation on the question under consideration: were both in accord, it would make no difference to the outcome which of them was consulted; if they happened to disagree, the index-agent's orientation on the issue of "reliability" would be determined under either procedure by the toss of a coin with a 50:50 loading. A similar argument holds for situation (ii), in which the polling agent shares a network with three other researchers and, in the event of their disagreement, will have a one-third probability of polling a member whose views represent the 2:1 majority, and the same chance of drawing the minority belief. The outcome, in terms of the orientation of the polling agent who merely mimics that of a network member chosen (with equal probability) at random, is the same as that which would result from canvassing the entire local network and -- in the absence of unaminity ~ forming an opinion probabilistically from the information obtained about the distribution of beliefs among its members. The same argument can be generalized to larger groups than the ones just considered, and the two procedures will be exactly equivalent in their probabilistic outcomes, so long as the rules for coin-tossing in the absence of unanimity is designed in the symmetric form that sets the coin-loadings to mirror the pattern of disagreement among other members of the social network. As soon will be seen, in Section 5, this formal equivalence of mechanisms turns out to be very useful in analysing the properties of the stochastic diffusion of belief statements throughout the extended community (the invisible college) which can be formed by the intersections, or overlappings of local social networks which share two-or more members with some other social networks.

But, even though, for purposes of analysis it is strictly correct and convenient to treat them the two mechanisms as equivalent, from another vantage point it may be held that these two procedures for forming an opinion really are not at all the same. This brings us to the second of the promised pair of remarks, a digression addressed to those who protest that the problems of theoretical framing and the ambiguities of complex experimental observations notwithstanding, the preceding description is a sheer travesty of science that ought not to be allowed to stand. Surely it is excessive to have minimized the persuasive force of modern science's cognitive content to the point of portraying its practitioners as being simply engaged in efficient opinion sampling among their close colleagues. Worse yet is the caricature proposed in latter reduced-form polling routine, in which the actors are quite passive, and mechanically copy the view of the colleagues with whom they have just made a (random) contact. Can it be thought reasonable to so completely disregard the claims of scientists to possess other, more reliable methods, such as direct observation and experimentation, which they deploy to arrive at "closure" in discussions about the validity of scientific statements?

4.3 Realists v. relativists: "Culture Wars" and an evolutionary reconciliation

One style of direct answer already has been given to this sort of protest -- by adherents of the so-called Edinburgh "Strong Programme" in the sociology of knowledge, following the seminal arguments of Bloor (1976). These "social constructionsts" would maintain that inasmuch as nothing distinguishes science from other domains of human action, and insofar as scientists behave like other social actors, there is no cause to seek to explain agreement, disagreement, success and failure in science in *special* terms. A generic account of consensus formation, such as the mimetic process captured by the foregoing random polling model, might well suffice to make the general point.

Rather than having recourse to such an extreme rationale, it seems more satisfying to point out that it is only in the face of total unanimity among a close set of scientific colleagues that our researchers go along with "the accepted view", and here we can look to Kuhn (1962) and the philosophers and historians of experimental science who have documented the power of reigning scientific paradigms to withstand assaults by contradictory empirical observations. In the other and more frequent circumstances in which there is some disagreement, the researchers in our model adopt a position formed probabilistically, and so it may be imagined that there would be occasions (admittedly they will represent the rarer contingencies) when a contradiction of the weight of received opinion by the results of the researcher's own inquiries is sufficient for the former to be rejected.

Will this median stance be sufficient? Can it serve to stave off an outright refusal on the part of dogged critics of the "social construction" movement in the sociology and philosophy of science to suspend for even a moment further their disbelief in the foregoing "toy" model of a science research network? The current state of debate over the issues that have been raised is not entirely encouraging on this score. In a recent exchange in the *New York Review of Books,* for example, Gerald Geison (1996:68) characterized the unsympathetic review by Max Perutz of Geirson's book, *The Private Science of Louis Pasteur,* as another "salvo in the culture wars that now divide too many scientists from too many historians, philosophers, and sociologists of science." Perutz (1996:69) answered by reiterating his rejection of the "relativist's" view of knowledge, ultimately, is warranted by institutionalized communities whose acceptance of an empirical claim certifies it as true. The terms in which he did so suggest no imminent abatement of the ongoing hostilities:

> "These statements represent a fundamental misunderstanding of scientific knowledge. For example, the Second Law of Thermodynamics states that heat cannot be transferred from a cold to a warm body without performing work. This is neither an empirical claim, nor a social construction, nor a consensus by institutionalized science, but an inexorable law of nature based on the atomic constitution of matter. Scientific laws are different from social ones like 'extreme poverty breeds crime,' which may be called a consensus among liberal sociologists."

It is manifest that Perutz and other practicing scientists do not accept the point of the SSK inquiry into the way in which "law like statements" come to be accepted. Their notions of the progress of scientific knowledge are more than merely unreconstructedly Whiggish: "acceptance" does not enter into Perutz's vocabulary in connection with scientific statements; and the possibility that propositions may be accepted as being "true," only to be discarded at a later time for a "better truth," is not something that interests him. Such propositions, one suspects, would not be seen by Perutz to belong to Science, which he treats as being interesting as an epistemological category, and not as a cultural activity. Taking an essentially teleological position, Perutz regards "science" as knowledge upon which the processes of inquiry *eventually and inexorably* will converge. Thus, it is not the acceptance by contemporary peers, but by later generations of scientists that confirms the truth of a scientific statement. For Perutz, Pasteur's claim to greatness rests not only on having gotten a useful answer, but one that is "seen to be correct" in the light of subsequent experience:

> "[Pasteur's] greatness derives from a long series of fundamental discoveries which have stood the test of time and have much reduced human suffering....I accept the validity of Pasteur's work not merely because of his practical successes, but because his experimental results have never been falsified and their validity has been underpinned and given a strong theoretical basis by modern molecular biology. It is the hallmark of scientific genius to find the right answers even before the rationale for them becomes apparent."

Geison (1996: p.69) is quite justified, therefore, in characterizing this approach — which emphasizes the "end of the story" ~ as one that is "fundamentally *ahistorical* and inappropriate to any assessment of Pasteur's scientific achievements in the context of his own time". The last seven words of this sentence constitute its operative phrase, because, it would seem, Perutz cannot see the sense of undertaking such an assessment except as a prediction of what will be the ultimate "judgement of history."

The idea that accepted scientific truths are at least partially matters of convention, "social construction'"or consensus formation (that is, pieces of culture) seems virtually inescapable if one is seriously concerned to understand how scientific communities function in real time. At the same time, however, it does seem hard to deny that there is a point to Perutz's (1996) instistence upon taking a longer view; it is an important quality of "scientific inquiry" as a collective activity that it is able to escape *eventually,* even if not instantaneously, from the grip of consensus upon propositions that are falsified repeatedly by observational experience. Acknowledgement of this special quality, surely, is one of the things that gives studies of the social and economic organization of *scientific* communities a special, distinctive claim upon our interest. Otherwise, why would we have a sociology of scientific knowledge that was differentiated from that of bureaucratic knowledge, or legal knowledge, or theology? The Edinburgh "strong programme" for the sociology of scientific knowledge proposed that the mechanism of consensus formation should be one that was neutral with respect to the eventual conclusion as to the validity or falsity of the propositions. This was a useful departure, in asking sociologists of science to take notice of what they had learned of a generic nature about belief-formation in communities of specialized practitioners. It was a mistake, however, to suppose that such a programme by itself would be sufficient to generate a useful sociology of knowledge in modern *scientific* communities. This, surely is what Perutz (1996) means when referring to the Second Law of Thermodynamics as not merely a "social construction," but something objectively real about the atomic makeup of matter.

If that is so, then a resolution of the conflict between the "two cultures" in this regard would seem to be within reach: we may think of the development of scientific knowledge as *an evolutionary process* — which is to say, it may usefully be conceptualized as involving mechanisms of selection to which alternative propositions (ideas) recurring will be submitted, as well as processes that serve to regenerate the diversity of candidate propositions. This formulation corresponds in a general way to the approach of the so-called "model" school of evolutionary epistemologists.[39] To

[39] For further discussion and introductory references to this literature in addition to Campbell (1965, 1974), and Hull (1988); see David and Sanderson (1996b) on the limitations of cognitive learning in such "evolutionary" processes.

introduce it within the confines of our overly simplified model, the invalidation and ultimate destruction of a consensus supporting "objectively unreliable" statements -- i.e., statements that fail repeatedly to derive reinforcement from predictive successes in operational contexts. — must be treated as probabilistic occurrences, rather than being foreordained. Probabilistic failures of this kind will be triggered by "accidents" in the interpretation of ambiguous observational evidence that, in effect, lead researchers to state counter-consensual opinions, thereby injecting into the network of scientific communications some "seeds of doubt" about the currently reigning set of beliefs. Adoption of a Darwinian "selectionist" position, familiar enough from a variety of formulations to be found in the evolutionary epistemology literature following the work of Campbell (1965, 1974), would then allow us to imagine that such sparks of disagreement (to switch the metaphor) are more likely to be released frequently by the recurring collisions between research programmes that adhere to erroneous and inconsistent propositions, and the comparatively rigid boundaries set by the world of Nature. The key question then becomes one of whether or not these expressions of doubt and dissent will be quickly smothered by the weight of contrary opinion prevailing in the scientific community, or whether the conditions governing the propogation of ideas will permit them to escape rapid suppression. By surviving within some local networks, these minority opinions must raise the likelihood of their acceptance in a wider sphere, and so may ignite a spreading conflagration of revisionism that eventually reorients the configuration of beliefs throughout the community at large. So long as the latter outcome remains a possibility, then, even on a purely chance basis one may expect that through their more frequent collisions with obdurate Nature the "objectively unreliable" proposition would tend over the long run to get selected out.

Thus, we reasonably may hope to arrive at an evolutionary epistemological synthesis as a form of comprise between the camps in the "culture wars." Under the proposed terms of peace, both sides should agree that a scientific community can arrive via generic "social processes" at a consensus on the acceptability of certain statements about the material world; also, there are some rules governing the way those statements are presented, and treated by members of the community, which have the effect of preserving the possibility that such a socially constructed "truth" will tend to be revised; there is, indeed, a long-run expectation that such social constructions will be discarded should they be found repeatedly to be difficult to square with other "truths" ~ especially those which possess a higher measure of "fit" with the logical implications and inferences that can be drawn from the available body of empirical observations.

5. From local stochastic social communications, to a model of global network behavior

We are at last in a position to draw together the strands of the preceding arguments and examine the properties of the stochastic communications model that they imply. To simplify the analytical task , I shall need to put aside any further distracting questions, such as how local social communication networks among scientists come to be formed in the first place, and whether strategic or other considerations on the part of the researchers might entered into the process of their formation. Instead, these structures will be treated as exogenously given multidimensional "neighborhoods" that are defined for each research scientist in the epistemic community under consideration.[40] The social boundaries of that larger research community,the *global network* which is the "invisible college," must likewise be taken to be exogenously delimited. Nevertheless, as shall be seen, the degree of coherence and agreement of its members on cognitive issues is an attribute of its disciplinary identity and intellectual connectedness that is in some important measure *endogenously* determined through the interactions of its members. The invisible college as conceived of here, therefore, is not an organizationally distinct entity; it is constituted of the ensemble of local social networks, through the chains of informal communication channels that connect the neighborhood of one researcher to the overlapping but not identically composed neighborhood where his or her immediate "correspondents" are situated.

5.1 Graph-theoretic preliminaries

Some concrete spatial examples may help in visualizing structures of this kind. First, think of a two-dimensional neighborhood in which there are four researcher-nodes (or research organizations) situated at the crux and the three extremities of a Y; the index node of this assembly is at the crux, and is connected to the rest of her social "neighborhood" by three communication channels. Two other such neighborhoods can be linked with this one, at the right and left upper branches, forming an arrangement like this: YYY. By now placing four more Y's at the four upper extremities of this structure, and positioning two more Y's below, so that their upper branches connect with the vertical's of the initial trio, a hexagonal honeycomb lattice can be extended in each direction, thus:

YYYY
YYY
YY

To take different spatial arrangment, we might start from a five-member neighborhood, in which the index member has neighbors to the north, south, east and west. A similar connected graph can be joined to that arrangement, in which two members of the first neighborhood will be shared by the second, and so on, adding still others until a square lattice develops. To keep everything perfectly symmetrical and leave no nodes in boundary positions, these two-dimensional graphs can be wrapped around in both the horizontal and vertical directions, connecting the right side edges to nodes on the left, and top side to the bottom — and so forming a *torus*.

In choosing to work with graphical forms of this kind, we are explicitly specifying that the neighborhoods, or local social networks of each researcher are symmetrical with those of all the others, and that they remain fixed for the purposes of our analysis. Both assumptions are convenient, which is the most than can be said for making them. Clearly, some scientific researchers have larger social networks, and networks of higher dimensionality at their immediate disposal than to others,

[40] Although the possibility is not pursued here, it would appears feasible to treat the local social communication networks explicitly as coalitions, and, following the lead of Kirman, Oddue and Weber (1986), to model their endogenous formation ~ and possibly also their ramifying interconnections — by applying concepts and analytical techniques from the branch of probability known as random graph theory (see Bollobás (1979)).

but we can simplify the analysis greatly by making the ensemble of social networks symmetric and homogenous.[41] In actuality, of course, invisible colleges are not temporally stable, as they consist of the researchers who happen at the time to be engaged with a particular set of cognitively interrelated scientific questions. Such populations in part or in whole may abandon the research field that formerly gave it intellectual cohesion and structured its members' communicative interactions; or extended invisible colleges may fragment into more compact areas of research specialization. The formal representation of an invisible college as *a network of localized social networks* abstracts from all these realistic complications, inasmuch as the fixed nature of the latter precludes researchers from migrating from one part of the larger network to another, or leaving the invisible college in question to participate in another.[42]

5.2 Markov random fields and the "voter model"

An undeniable attraction of the probabilistic routine for opinion formation set out in preceding section is that it corresponds directly with the heuristic "voter model" introduced in different contexts by Clifford and Sudbury (1973) and Holley and Liggett (1975), and best known in the form elaborated by Harris (1978). The latter heuristic model is based on Markov random field theory, a branch of probability theory that has lent itself to application in the study of social networks (see Kinderman and Snell (1980)), and, more recently, has been extended to the analysis of the dynamics of technological competitions in economic contexts that are characterized by the existence of local network externalities.[43]

Leaving aside technicalities, it is possible formally to set out this framework in a way that enables us to represent scientific communication and consensus formation in inter-linked local networks as a stochastic path-dependent process. Following the notation by Kinderman and Snell (1980), we may begin with the basic definitions relating to Markov random fields.

Let G = (O,T) be a connected graph with $O=(o_1, o_2,...,o_n)$ the vertices (i.e. the set of researchers, or research organizations) and $T=(t_1, t_2,...,t_m)$ the edges (i.e. the set of information transmission lines). For the moment we restrict the discussion to connected graphs of social networks that are defined in one or two dimensions. A *configuration* x is an assignment of an element of the finite set S to each point of 0. We denote this configuration by x = (Xo) where x_o is the element of S assigned to vertex o. If we let S = *[u, a*] represent assignments of the two possible

[41] Morris (1996), using mathematical tools other than those employed below, has shown that a number of the key properties of local interaction games, concerning the dynamic propagation of strategies chosen in particular locations, and the existence of equilibria of correlated equilibria, hold generally for a wide class of local (spatial) structures. On the other hand, the assumptions of symmetry and homogeneity are not wholly innocuous. For example, Bala and Goyal (1995) show that greater symmetry increases the speed of information diffusion in a local interactive learning game.

[42] It will be noted below (in section 7) that it is nonetheless possible within this framework to make allowance for newcomers to be inserted into the community — by attaching themselves, albeit independently, and randomly, to new social networks, or by occupying the places opened up in the network through the natural attrition of some of its members. But for the immediate purposes, such matters are left to one side.

[43] See, *inter alia,* David (1988,1993b); David and Foray (1993).

opinion orientations on the reliability of a given scientific statement (*a* standing for "acceptably reliable", u for "unreliable"), a configuration would be an assignment of either o_u or o_a to each of the points in O. A *random field* p is a probability measure p(x) assigned to the set X of all configurations, such that p(x)>0 for all x. By the "neighbors" N(o) of the point o we shall mean the set of all points o' in O such that (o'o) is an edge. A random field p is called a *Markov random field* if:

$$p\{Xo = s|x_o\text{-}o\} = p\{x_0 = s|x_{N(0)}\}.$$

That is, in trying to predict the value at o (either *u* or *a* in the example), given the values at all other points of O, we need know only the values assigned to the neighbors of o.

Assume now, following the "voter model", that associated with each point of a graph we have a researcher or research unit, and that with every such unit there is a reference set comprised of other units(the neighborhood available for random polling at a given moment in exponential time, to select one a neighbor whose orientation it copies). Then the process may be described informally as follows: each research unit selects an orientation in regard to the statement in question, *u* or *a.* At random points in time it will reassess its opinion. At these times it will commit to the choice *u* with a probability equal to the proportion of u-oriented research units in its reference set. The global process of migration between the alternative orientations of belief as to the reliability or unreliability of a given scientific statement now can be described by a finite state continuous time Markov chain, with states being configurations of the form

x = *(u, a, u, u, a,..., u, a, u),* where x(i) is the choice of research unit i.

A number of important properties of this well-studied process[44] may be immediately stated:

Property (1): It is evident on even the briefest consideration that the extremal states x^u = *(u, u, u, ..., u, u, u)* and x^a = *(a, a, a,..., a, a),* in which there is a perfect correlation of beliefs throughout the population, constitute absorbing states for this system. Once such a state is entered, there can be no further change. The existence of a multiplicity (two) of absorbing states tells us plainly that *this process is essentially historical, in the sense of being non-ergodic* - it cannot invariably shake loose from all initial configurations.

Property (2): A somewhat less obvious proposition, also true, is that for any starting state x the chain eventually will end up in either x^u or x^a. Thus, in the limit, *the process must become "locked-in" into one of its extremal solutions.* The system in this sense does eventually produce "closure" on the scientific issues submitted to it.

Property (3): There exists a limiting probability distribution over the macrostates (opinion configurations) of the system which is non-continuous, such that, starting in x, the probability that the chain will end in x^u is equal to the proportion of *u* in the initial configuration x (without regard to their position in the array); and the probability that it will end up in x^a is equal to the proportion of *a* in the initial configuration x. Therefore, although subject to random influences, *the nature of the*

[44] See Kinderman and Snell (1980b), and the fuller discussion in David (1993b).

asymptotic macrostate consensus in this system can be predicted (not with certainty, but probabilistically) from information on the initial configurations of opinions.

What is the import of these results -- a multiplicity of possible outcomes, eventual closure, predictability of the eventual outcome of scientific controversies on the basis of information concerning the distribution of the researcher's initial opinion orientations ~ in the present application context? Obviously, one thing they signify is the possibility of formally connecting the social organization of science (affecting communications) with an important performance attribute of science communities in the cognitive domain. Another rather pleasing formal reaffirmation of insights obtained via other routes is that "the details of history may matter" for the cognitive development of a scientific field. Further, in this light, the propensity of scientific communities to remark especially on instances in which new ideas have won eventual acceptance in the face of an initial consensus opposing them, is entirely understandable, for, at least in the near term such cases would constitute the rarer contingencies.

Several technical qualifications must be appended to the foregoing conclusions about the voter model's properties, especially as these also admit of interesting intepretations in the present context. First, the property of "complete closure", in the sense of lock-in to a consensus characterised by perfect unanimity, will not survive generalization of the framework to higher-dimensional graphs. It has been found that substantial correlations of opinions eventually would emerge in a community that was arranged with reference to three social dimensions. Plausibly enough, it may be surmised that as one goes to still higher-dimensions in the social space, and the community becomes in that sense "less compact", its ability to arrive even at a strong consensus would deteriorate further (see Kinderman and Snell (1980b). Perhaps the recurring formation of disciplinary sub-specialities in science serves as a "social compacting process", the latent function of which is to preserve network performance in producing closure. On the other side of the coin, higher dimensional social networks may be seen to improve the chances of the survival of heterodox opinions within small minority clusters of researchers who, in effect, shield one another from the conformity-inducing pressure of exposure to the preponderance of belief in the epistemic community at large.[45]

A second point of qualification is that the properties of lock-in to closure and predictability of the nature of the resolution are ones that strictly hold only *fox finite* populations. If the population of the network were to be constantly growing at a comparatively rapid rate - strictly, at a pace rapid enough to cause the introduction of newcomers (entering the field with randomly distributed beliefs about the scientific issues of the day) to overwhelm the pace of the process of random polling in the local social networks, then closure would no longer be assured. Moreover, under those conditions, the nature of the cognitive outcome would cease to be predictable on the basis of the system's initial configuration. It has been surmised by probability theorists that the dynamics of convergence to one or the other extremal (uniform consensus) configurations in a large finite system would approximate

[45] The broader significance of this will be further remarked upon below, in Section 6.

those of the infinite population case: such systems continue to migrate back-and-forth between extremal states, albeit with transit times between them that are quite prolonged. Here, then, we may remark upon another respect in which the cognitive performance of scientific communities could be related to their organizational dimensions and dynamic attributes of invisible colleges, such as the rate of entry of new members in relation to the speed of opinion-forming informational transactions within local social networks , and the "pre-entry orientation" of new recruits in reference to the prevailing distribution of scientific opinions held by those who currently constitute "the field". A substantial consensus can be expectedd to emerge even in the absence of strong pre-orientation as a criteron of eligibility for entry, once an invisible college "stabilizes" demographically ~ in the sense that its growth rate slows to the point that it is exceeded by the average internal opinion-polling rate. Here is where communications technology can help support the rapid growth of research communities that retain their epistemological functionality: a speed-up of the 'polling rate' will permit the mobilization of resources at a research frontier to proceed more rapidly without jeopardizing the network's ability to reach closure on the new questions that it has taken up for investigation.

Yet, another, and quite important class of qualifications regard arises from closer consideration of the assumptions of the basic "voter model" with regard to the symmetry of communication behaviors on the part of the research-agents. These can be brought out now by turning to consider the properties of a somewhat different stochastic communications structure, in which not all the actors follow the same policy in regard to the informational transactions with other research units.

5. 3 Imperfect communications, connectivity and critical percolation probabilities

The population of researchers portrayed in the "voter model" of the preceding sub-section is homogeneous in the special sense that the structure of communication links among them is completely symmetrical. Moreover, it is assumed to take a rather special form that renders the transmission of influence among them completely deterministic, even though the timing of interactions between pairs of agents within the local network is randomly generated. All of the researchers are always sharing their opinions with all who ask, as all others in their social network stand ready to do. Putting aside the possibility of entry, the source of randomness in the revisions of beliefs within the population has to do not whether or not particular researchers might be open to the influence of particular neighbors, but, rather with the direction of the re-orientation of beliefs that such influences would bring about. The concepts and terminology *of percolation theory* provides a precise way of describing these specifications, and showing their relationship to a more general specification of the model.

The *term percolation process* refers to the dual of a diffusion process. "Diffusion", to speak strictly, refers to the random movements of particles through an ordered, non-random medium — as in the case of the diffusion of molecules of salt in water; whereas, the term "percolation" conjures up the image of droplets of water moving under the deterministic pull of gravity through a dis-ordered, random medium — such as filtration tank filled with sand and pebbles of different sizes. When the water, entering at some source sites, eventually finds its way into enough open channels to pass throughout, wetting the entirety of the interior surfaces, *complete percolation* is said to have taken place. It is from this that the mathematical statistics describing the properties of analogous processes has acquired the label "percolation" theory (see Grimmet (1989)).

Adapting the notation of Hammersley and Welsh (1980) to the Markov random field framework, let **G** be a graph in which some, none, or all of the edges may be directed. Thus, as before, **G** consists of a set of research units (corresponding to the graph's vertices or nodes), $\mathbf{O}=(o_1, o_2,...,o_n)$; these are connected by a set of (possibly directed) edges representing channels of social communications, $\mathbf{T}=(t_1, t_2,...,t_m)$. An *operative path* in **G** from an research unit o_1 to another research unit, o_n, is a finite sequence of this form:

$$\{t_{12}\ o_2\ t_{23}\ o_3\ ...\ t_{[n-1]n}\ o_n\},$$

where t_{ij} denotes a relational line connecting o_i to o_j. The graph **G** is *connected* if for each pair of researchers o_i and o_j there is a path in **G** from o_i to o_j.

Now construct a *random maze* on **G**, as follows. Let each research node o of **G** be *open*, or ready to transmit messages that can influence any of its neighbors' opinions on the reliability of the statement at issue, with probability p_s; and, alternatively, it will be *closed* (unwilling to share its present knowledge on the question) with probability $q_s = 1- p_s$. Similarly, each line of inter-personal or inter-organizational communications t_{ij} may be thought of as potentially carrying messages that will be actually "read" with probability p_r, or it fails to do so with probability $q_r = 1-p_r$. Furthermore, we shall assume all these events are to occur independently of each other.

An operative path, $D = \{\ t_{12}\ o_2\ t_{23}\ o_3\ ...t_{[n-1]n}\ o_n\}$ from o_1 to o_n is said to be "open" if all its communication links are functioning and all its research organizations are ready to 'share' their knowledge-conclusions. Thus, the probability that the particular path D is "operational" in that sense is given by $(\ p_r\,p_s\ )^{n-1}$.

Let **Z** be some given set of "source" researcher units, from whom a particular "idea" or scientific statement emerges into G. The decisions to adopt that statement as reliable (or not) can flow along any open path from a source research unit and will then similarly reorient the other units on such a path ("wetting" them, to use the natural percolation metaphor). The *percolation probability* $P(p_r,p_s|\mathbf{Z},\mathbf{G})$ is the probability that **Z** can thus reorient some infinite set of nodes in **G**. We call the parameters p_r, p_s, respectively the probabilities of "reading" and "sending" or "sharing" information. In other words, in a large population, it can be expected that a proportion p_r are receptive to their neighbors' opinion on the reliability of a statement, while a proportion $1-p_r$ are unreceptive. The transactional lines of **G** connect neighboring pairs of researchers and the model supposes that a researcher already committed into a given scientific position has a chance p_s of "infecting" a neighbor, provided that the latter is receptive. Then $P(p_r,p_s|\mathbf{Z},\mathbf{G})$ is the probability that a belief initially established in the source research organizations **Z** will become adopted universally.

Suppose that **Z** and **G** are fixed, that **G** is infinite and that we abbreviate

$$P(p_r,p_s|\mathbf{Z},\mathbf{G}) = P(p_r,p_s) = P,$$

then, clearly, the mixed percolation probability P is a non-decreasing function of p_r and p_s, and it follows that: P(0,0) = P(l,0) = P(0,l) = 0, while P(1,1) = 1. Thus, P,(p) = P(p_s, 1) and P_r(p) = P (1, p_r) are, respectively, called the *node percolation* and *connection percolation* probabilities of this system in the literature of applied percolation theory.

A fundamental mathematical property of the percolation process is that there exists some critical values of $p_r > p^*_r$ and $p_s > p^*_s$ beyond which there will be a positive probability that percolation occurs, and below which the percolation probability is zero (see Hammersley and Welsh (1980), Grimmett (1988)). In other words, the system undergoes a "phase transition" when these underlying critical probabilities are attained. There are corresponding critical values at which the node-percolation and edge-percolation probabilities, respectively, become positive. These define the endpoints of a region above which a "mixed-percolation process" (one for which it is not certain that either all nodes or all edges of the graph are open), will have positive probability of achieving complete percolation. Figure 1 illustrates the position and shape of this region. What these results from percolation theory tell us in the present context is that there is a minimum level *of persistently* communicative behavior that a finite size science network must maintain if ideas are to percolate within it, so that closure can be obtained.

6. A new perspective on the significance of the Mertonian "norms"

For a community of scientists to exist as a cognitively functioning entity, therefore, it has been seen that there is a formal necessity of attaining some critical measures of "expected connectedness," through the average communication behaviors of its constituent members. Another pertinent result from percolation theory is represented by placement of the endpoints of the boundary shown in Figure 1: there is an important asymmetry between the effects upon network performance of reducing the representative agent's communication probabilities. Thus, a given proportional reduction of the mean probability of sending messages (node openness) will do more to degrade the percolation performance of the system than will the equiproportional reduction of the mean probability of a communication link being open.[46] In view of this, the invisible colleges' first condition for functionality, in the sense of its most exacting requirement, is that the network must maintain at least the critical level of openness in regard to the behavior of a "representative node", i.e., in the expected proportional composition of member "types" with respect to disclosure of their scientific knowledge. The representative researcher, of course, is a purely statistical construct in this framework ~ an average of nodes that are *permanently* open and those that are *permanently* closed. The fraction of those who, being closed, will never "share" (or never "write") what they have learned, therefore, must not be allowed to exceed the critical level ($1-p_s^*$) if the invisible college is to retain a positive probability of reaching closure on scientific questions.

6.1 From the norm of disclosure to the probability of closure

[46] See Hammersley and Welsh (1980), for proof of the generalized asymmetry theorem; David and Foray (1993, 1994) for its application in the social communication network context.

What follows from the fact that the qualitative performance of the system undergoes this critical shift when the mixing fraction passes a specific level, but the precise value of that critical point is not likely to be known in advance, nor to become evident when the mean probability of node openness is progressively degraded ~ until it is too late? In such circumstances it would be sensible to protect the cognitive functionality of the global network by maintaining a "safety-first" policy of selectivity in regard to the recruitment of researchers to the college, requiring some prior test of their intrinsic propensity towards"openness" concerning what they will find and conclude in the course of their researches. By the same token, the existence of strong and universal norms requiring cooperative behavior on the part of researchers, especially in disclosing what they learn, and a reward system that elicits such disclosures as the basis for establishing a collegial reputation, can be seen to serve as important bulkwark protecting the invisible college's ability to "deliver" a clear consensus on the questions brought before it. Furthermore, inasmuch as $p_s^* > p_r^*$, in the design of the incentives for individual behavior in such a system, it is reasonable from the viewpoint of assuring "connectivity" that a concern to assure the consideration by others of one's own "findings" should take some measure of precedence over concerns about the arrangements and facilities that affected the average propensity of network members to attend to each others' messages. The foregoing observations might be taken as a formal communications-theoretic rationale for the stressed placed by Mertonian sociology upon the functional importance of the behavioral norm of openness among scientists, and the corresponding tendency in that literature to de-emphasize the significance of variations in the reliability of communications technologies upon the qualitative performance of scientific communities.

Still other implications of the Mertonian "norms" for the conduct of (non-proprietary) research can be brought into sharper focus by reconsidered their operation in the context of the stochastic models presented here. Viewed from the perspective of our earlier discussion (in section 4.3) of science as an evolutionary epistemological process, norms of "disinterestedness," "universality," and "disclosure" may be seen as crucial when taken together, precisely because they reinforce behaviors at the micro-level that permit an the tendency for more "objectively reliable" forms of belief to emerge by selection from the collision between consensus and contradictory observations. There has been an important failure of understanding among the relativists, the partisans of the "strong programme" in SSK who dismiss these "norms" as a self-serving ideology that is betrayed by human failing and competitive stratagems adopted among the practitioners of science. It is naive to suppose that the essential features of the qualitative performance of a mode of social organization will be lost if any one among its "norms" is violated by some members at some point in time. Any system of social organization that is so rigid and unrobust that it cannot tolerate a degree of deviance, is not likely to survive for very long. Furthermore, it has been seen — from the properties of both the voter model, and the percolation model -- that some degree of intermittent (random) interruption of communications, and, even some margin of persistently non-communicative actors, are formally compatible with the ability of the global network eventually to arrive at closure.[47]

6.2 Reducing the hazards of premature closure: The "universalistic" norm and network dimensionality

Just as deserving of criticism are those champions of the objectivity of "the scientific method," who celebrate its power to uncover "the truth," yet fail to grasp the connection between the success of science and the nature of the social organizations that scientific communities represent. It is a too human failing of scientists who miss this point that they quite typically attribute the progress of scientific knowledge to the succession of individual "genius". It may be quite proper to share the doubts expressed by Perutz (1996), about the philosophical representation of epistemological progress resulting from the dogged adherence to a formula-like "scientific method," and to deride as pure fantasy the picture of scientific researchers as just so many drones harnessed to the crank of a cleverly designed inferential engine, mechanically proceeding through the cycle after cycle of theoretical hypothesis and empirical refutation, until law-like propositions emerge, like so many sausages. Whether or not the process of inquiry at the individual, microcosmic level of the investigator is messy, irrational, driven by the insights of genius, or by human motives that are less than noble, really is not what matters. These are not the crucial aspects of the method of modern science upon which the discussion should be focused. Instead, it should be more widely understood that there is an aspect of the "scientific method" for which it is true that the efficacy of the enterprise depends upon the subscription of of most members of the research community. This is just the "method" or mode *of social* organization affecting the internal communications of invisible colleges.

It has been pointed out that in the analysis of the voter model that in the case of networks that are arrayed in three or more dimensions, one loses the property of "lock-in" to *complete* closure. Further, it has been noticed that perfect correlation of beliefs throughout the global network constitutes an absorbing, or "trapping" state for the system, so that compact, low dimensionality systems are exposed to the risk of being unable to free themselves from the grip of continually reinforced orthodoxy. The universalistic norm may therefore be seen to forestall restriction of entry into network membership by criteria that otherwise would tend to render the social dimensionality of the resulting communications structure more compact, and more likely to result in the rapid extinguishing of small pockets of dissenting opinion. The ability to shelter the views of dissenting minorities from the pressures for comformity with predominating peer beliefs, pressures such as are especially likely to exist in collegiate reputation-based reward systems, must thus be an important system attribute. It reduces the likelihood of premature consensus, and preserves the possibility of an evolution of beliefs away from those that persistently generate observational anomalies.

---

[47] Trust in that capability, of course, is what has been presented as underpinning the rational microlevel strategies of the agents engaged in polling their respective networks in the case of the voter model.

Thus, by enjoining members of the community to accept dissenting claims as worthy of examination, without regard for the economic, social, political or nationality status of the claimants; by insisting on disclosure as the condition for successful claims to the rewards that are attached to priority; by preventing a complete and enduring monopolization of knowledge and access to knowledge from protecting the consensus from challenge, these "norms" of the Republic of Science serve an important epistemological function. They allow signs of the collisions between the constructs of social communication and the constraints implied by the structures of the material worlds to be registered; they encourage the perception among individuals that such signs are to be interpreted as indicating opportunities for achieving recognition, and reward, rather than as areas in which they might stumble into heresies that will bring punishments.

We have arrived, via a somewhat novel route, at a fresh appreciation for the critical balancing of the influences exercised upon individual behaviors by the institutionalized norms associated with the practice of "open science." Jointly, these tend to provide a mechanism that leads to closure, while protecting existing diversity of opinion from the generic social pressure for perfect conformity and uniformatization that otherwise would arbitrarily extinguish the process of consensus- revision. Consequently, there is a basis for supposing that the stabilization of beliefs among communities whose members adhere to such rules will tend to be indicative of the attainment of that greater degree of fitness with the unchanging features of the natural world —which specially qualifies the (scientific) statements involved to be described not merely as knowledge, and information, but as "reliable knowledge". More precisely, we should recognize, the stabilization of community opinion can occur not because "the truth" has been uncovered, but because the material and intellectual resources at the disposal of working scientists ~ for example, the sheer limitations of the resolution power of their telescopes, the magnification power of their microscopes, the sensitivity of their particle detectors and the computational power available for numerical analysis of data — have intervened; that it is perceived that although matters are not fully resolved, the opportunities for individuals to win recognition for contributions that will effect a revision of the consensus are a potentiality that, for the time being, remains unripe for exploitation.

7. Towards extensions of the model: Closure, creativity and the size of invisible colleges

A further range of insights is available if the analysis is focused on the performance of open science communities, all of whose members conform probabilistically to the norms of openness. Here we return to the world of transient random mazes represented by the voter model, where the communication probabilities pertain to the homogeneous behavior of the researchers in regard to their sending and receiving of messages. Under those conditions (see Cox (1989)), the expected length of time that is required for "closure" will be increased by all movements away from certainty that a randomly chosen node, and a randomly selected communications channel will be actually "open" when probabilistic polling occurs. The macro-level performance implications of alterations in the mean probability of node (and channel) openness is remains. But, what amounted to a critical qualitative change in the case of percolation through a permanent maze (in the case of the pecolation model), is transformed into a continuous modification of the speed of attaining closue when one consider the case of random mazes (as in the voter model). Exploring the dynamic implications these and related results about system performance in presumptively "open science" research networks is a task that must be left for the future, but the direction of advance may be briefly indicated here.

The basic model of a research network that has been presented in this papercan be elaborated quite easily to allow dynamic feedbacks, whereby the macro-level performance characteristics of the global network would affect the micro-level communication behaviors of the research-agents in their respective local social networks. In the most elementary implementation of this extension (by David and Flemming, 1996), the latter are just the homogeneous mean probabilities of sending and receiving information, or "sharing" and "reading". The resulting closed stochastic system can be transformed into an "equivalent deterministic system" by analyzing the *expected* performance characteristics of a science research *network viewed as an entity.* Pursuing the latter approach in initial explorations is especially convenient, as it dispenses with the need for extensive stochastic simulation experiments in order to establish the main qualitative aspects of the endogenous dynamics of network growth and performance.

In David and Flemming's (1996) implementation of the equivalent deterministic system modeling approach, the expected percolation time for establishing "closure" with regard to an idea, and the expected rate of "creativity" through recombinant generation of ideas, are each modeled as dependent upon endogenously determined features of the global network's performance. Key among these are the global network's size, and the mean probability that within local social networks all ideas will be shared freely (transmitted). Making use of a result due to Cox (1989) on the expected time to percolation on a three-dimensional torus, the mean speed of closure can be taken to be proportional to the product of the mean probabilities of sending, and of receiving messages, and inversely proportional to the cube of network size. The mean probability of a "new scientific proposition" being generated at any given moment somewhere in the connected network, or the invisible college's "creativity" rate, can be represented as a (positive) power function of the network's size, and of the probability that new ideas are "shared". The latter specification may be rationalized by appeal to findings in cogntive psychology and the social psychology of science, and in the history of invention, all of which point to the importance of "recombinant creativity," and the sensitivity of latter phenomen upon the size and diversity of the set of ideas and techniques that are available for permutation and recombination by members of an epistemic community.[48] In this context, emphasis is placed not upon the determinants of individual creativity, and there is no assumption as to the automaticity of the cumulative process through which new ideas are generated. Rather, like "closure", "creativity" is an emergent property of social networks that have certain features, among which size and openness are critical.

---

[48] See, e.g., the contributions in Shadish and Fuller (1994), Usher (1982), espcially Ch.IV. Weitzman (1995) builds a model of endogenous economic growth around the core concept of "recombinant inovation," taking the process of hybridization of existing plant varieties to obtain new and improved strains as a metaphore for the process through which "knowledge can built upon itself in a cumulative interactive process" involving the combination and recombination of ideas.

When the global network size can be varied exogenously, a variety of networks displaying a wide range of performance attributes all turn out to be equally sustainable equilibrium solutions. If the "creativity" effect of greater network size is strong enough to overcome the latter's effect in slowing the expected speed of closure, the larger networks belonging to this family, generally will be those which are consistent with a greater (probability) measure of "openness" at the micro-behavioral level. Hence, by virtue of both their size and (the assumed diversity of the ideas held by their members),the higher probability that the network's members will communicating those ideas openly, the bigger networks are able also to sustain higher levels of systemic "creativity".

A fully dynamic equivalent deterministic system also may be obtained by taking the further step of modeling the influence of those latter (average) performance features upon the ability of the network to maintain its size, and even to grow by recruiting members from a surrounding eligible population. This entails the formation of expectations about the implications of the overall network's performance for the benefits that potential members would derive by joining it, in comparison with those obtainable in competing employments. It is possible to consider either situations in which the latter opportunity costs are exogenously set, or "closed" models, in which they are determined endogenously by forces related to the performance of the network in question.[49]

There are at least two useful aspects of the demonstration that such simple formal models can generate results that reinforce the intuitions, and generalize the insights gained through careful and detailed empirical research in the sociology of scientific communities and the institutional and cultural contexts in which research is conducted. First, itsuggests that future work in this vein may yield some more subtle insights into the connections between micro-behaviors and macro-phenomena that are of interest for science policy formulation. Second, there is a potentially significant heuristic benefit: formal exercises of this kind do open up an attractive avenue that should entice economists proficient in the requisite modeling skills into learning more about the social and intellectual organization of scientific research communities, so that they may begin to deploy their talents in what one hopes will be increasingly sophisticated analyses of the production and distribution of scientific knowledge by networks of researchers.

[49] See David and Flemming (1996). It is even possible to extend the model to analyze the dynamics of competition between two networks that can draw recruits from a common pool of researchers, but this is left for future work.

References

Adler, C. G. and B. Coulter (1978), "Galileo and the Tower of Pisa Experiment," *American Journal of Physics,* 46: 199-201.

Alba, R. (1973), "A graph-theoretic definition of a sociometric clique," *Journal of Mathematical Sociology,* Vol. 1.

Anderlini, L., and A. Ianni (1995), "Path dependence and learning from neighbours," forthcoming in *Games and Economic Behavior.*

Arora, A., P.A. David and A. Gambardella (1995), "Reputation and competence in publicly funded science: estimating the effects on research group productivity," paper presented at the National Bureau of Economic Research Productivity Workshop, summer 1994; revised for the ICER Conference, Turin, Italy, June.

Arbra, A., and A. Gambardella (1992), "Public policy towards science: picking stars or spreading the wealth?", paper presented at the Rosselli Foundation Workshop on Path-Dependence in Economic Processes, Torino, Italy.

Arora, A., and A. Gambardella (1994), "The changing technology of technological change: general and abstract knowledge and the division of innovative labour," *Research Policy* 23, 523-32.

Arrow, K.J. (1962), "Economic welfare and the allocation of resources for inventions," in R.R. Nelson (ed.), *The rate and direction of inventive activity: economic and social factors* (Princeton, NJ: Princeton University Press).

Arrow, K.J. (1971), "Political and economic evaluation of social effects and externalities," in M. Intriligator (ed.), *Frontiers of quantitative economics,* Contributions to Economic Analysis No. 71 (Amsterdam: North-Holland).

Bacharach, M., D. Gambetta et al. (1994), "The economics of salience: A research proposal," Unpublished paper, Oxford Institute of Economics and Statistics.

Bala, V., and S. Goyal (1995), "Learning from neighbors," Econometric Institute Report 9549, Erasmus University, Rotterdam, September.

Barnes, B. (1974), *Scientific Knowledge and Sociological Theory,* London: Routledge and Kegan Paul.

Barnes, B. (1977), *Interests and the Growth of Knowledge,* London: Routledge and Kegan Paul.

Bloor, D. (1976), *Knowledge and Social Imagery* ,London: Routledge & Kegan Paul.

Blume, L.E. (1993), "The statistical mechanics of strategic interaction," *Games and Economic Behavior,* 4: 378-424.

Bollobás, M. (1979), *Graph Theory: An Introductory Course.* New York: Springer-Verlag.

Gallon, M. (1995) "Four models for the dynamics of science," in S. Jasanoff, G.E. Markle, J.C. Petersen and T. Pinch (eds), *Handbook of science and technology studies* (London: Sage Publications).

Gallon, M., and J.-P. Courtial (1989), *Co-word analysis: a tool for the evaluation of public research policy* (Paris: Ecole Nationale Supérieure des Mines).

Campbell, D.T. (1965), "Variation and selective retention in socio-cultural evolution," in H.R. Barringer, G.I. Blanksten and R.W. Mack (eds), *Social change in developing areas* (Cambridge, MA: Schenkman Press).

Campbell, D.T. (1974), "Evolutionary epistemology," in P. A. Schilpp (ed.), *The philosophy of Karl R. Popper* (LaSalle: Open Court).

Campbell, D. T. (1994), "The Social Psychology of Scientific Validity: An Epistemological Perspective and a Personalized History," Ch.2, in *The Social Psychology of Science*, New York: The Guilford Press.

Caudill, Maureen and Charles Butler (1990), *Naturally intelligent systems* (Cambridge, MA: MIT Press).

Clifford, P., and A. Sudbury (1973), "A model for spatial conflict," *Biomethka* 60, 581-8.

Cole, J. and S. Cole (1973), *Social Stratification in Science,* Chicago: Chicago University Press.

Cole, S. (1978), "Scientific Reward Systems: A Comparative Analysis," in R. A. Jones, ed., *Research in Sociology of Knowledge, Sciences and Art,* Greenwich, Conn.: JAI Press, pp. 167-190.

Cole, S. and Cole J. (1967), "Scientific Output and Recognition," *American Sociological Review,* 32:.377-390.

Collins, H.M. (1974), "The TEA set: tacit knowledge and scientific networks," *Science studies,* 4: 165-86.

Cowan, R. and D. Foray (1996), "The Economics of Codification and the Diffusion of Knowledge," Unpublished working paper: IMRI, University of Paris-Dauphine, (March).

Cox, J. S. (1989), "Expected percolation times for the voter model on a d-dimensional torus," *Annals of Probability.*

Crane, D. (1965), "Scientists at Major and Minor Universities: A Study in Productivity and Recognition," *American Sociological Review,* 30:.699-714.

Dalle, J.-M. (1995), "Dynamiques d'adoption, coordination et diversite: la diffusion des standarads technologiques," *Revue Economiques,* (Juilliet).

Dasgupta, P., and P.A. David (1987), "Information disclosure and the economics of science and technology," in G. Feiwel (ed.), *Arrow and the ascent of modern economic theory,* New York: New York University Press.

Dasgupta, P., and P.A. David (1988), "Priority, secrecy, patents and the economic organization of science and technology," CEPR Publication No. 127, Stanford University (March).

Dasgupta, P. and P.A. David (1994), "Toward a new economics of science," *Research Policy* 23, 487-521.

David, P.A. (1988), "Path-dependence: putting the past into the future of economics," Institute for Mathematical Studies in the Social Sciences Technical Report No. 533, Stanford University.

David, P. A. (1993a), "Historical economics in the long run: some implications of path dependence," in G.D. Snooks (ed.), *Historical analysis in economics* (London: Routledge).

David, P.A. (1993b), "Path-dependence and predictability in dynamic systems with local network externalities: a paradigm for historical economics," in D. Foray and C. Freeman (eds), *Technology and the wealth of nations* (London: Pinter Publishers).

David, P.A. (1994a), "Positive feedbacks and research productivity in science: reopening another black box," in O. Grandstrand and P. Jacobson (eds), *Technology and economic change* (Amsterdam: Elsevier).

David, P. A. (1994b), "Les standards des technologies de l'information, les normes de communication et l'État: un problème de biens publics," *Analyse économique des conventions,* ed. A. Orléan (Paris: Presses Universitaires de France).

David, P. A. (1995), "Reputation and Agency in the Historical Emergence of the Institutions of 'Open Science[1]," Paper presented to the National Academy of Sciences Colloquium on the Economics of Science and Technology, held at the Beckman Center, U.C. Irvine, 20-21 October 1995.

David, P.A. (1996a), "Science reorganized? Post-modern visions of research and the curse of success," in *Measuring R&D Impact — Proceedings of the Second International Symposium on Research Funding,* (Ottawa: NSERC of Canada).

David, P. A. and T. Flemming (1995 & 1996), "Communications, creativity and research network dynamics: An elementary economic model of scientific communities," Paper presented to the International Conference on Creation and Transfer of Knowledge, held at Castelgandolfo (Rome), September 1995; revised for presentation to the All Souls College Seminar on the Economics of Information Diffusion, convenned in Oxford, May 1996.

David, P. A., and D. Foray (1993), "Percolation structuers, Markov random fields and the economics of EDI standards diffusion," in G. Pogorel (ed.), *Global telecommunication strategies and technological change* (Amsterdam: Elsevier).

David, P.A., and D. Foray (1994), "Dynamics of competitive technology diffusion through local network structures: the case of EDI document standards," in E. Leysendorff (ed.), *Evolutionary economics and chaos theory: new developments in technology studies* (Amsterdam: Elsevier).

David, P.A., and D. Foray (1995), "Accessing and expanding the knowledge-base in science and technology," *STI Review —Science, Technology and Industry, 16,* (OECD: Paris).

David, P. A., D. Foray and J.-M. Dalle (1996), "Marshallian externalities and the emergence and spatial stability of technological enclaves," forthcoming in *Economics of Innovation and New Technologies.*

David, P.A., A. Geuna, and W.E. Steinmueller (1995), "Additionally as a principle of European R&D funding," Report for the STOA programme of the European Parliament, MERIT Research Memorandum 2/95/012, Maastricht.

David, P. A., and S. Greenstein (1990), "The economics of compatibility standards: an introduction to recent research," *Economics of Innovation and New Technology* 1 (1 & 2), Fall 1990, 3-42.

David, P.A., D. Mowery, and W.E. Steinmueller (1992), "Analyzing the payoffs from basic research," *Economics of Innovation and New Technology,* Vol. 2, No. 4.

David, P. A., and W. C. Sanderson (1997), "Making use of treacherous advice: Cognitive learning, Bayesian adaptation and the tenacity of unreliable knowledge," forthcoming in *Frontiers of the New Institutional Economics,* J. V. C. Nye and J. N. Drobak, eds. (San Diego, CA: Academic Press.)

Ellison, G. (1993), "Learning, local interaction, and coordination," *Econometrica,* 61: 1047-1071.

Ellison, G., and D. Fudenberg (1995), "Word-of-mouth communication and social learning," *Quarterly Journal of Economics* 110(i): pp. 93-126.

Feyerabend, P. (1975), *Against Method*X London: Humanities Press).

Franklin, A. (1986), *The Neglect of Experiment* (Cambridge: Cambridge University Press).

Fuller, S. (1994), "A Guide to the Philosophy and Sociology of Science for Social Psychology of Science," Ch. 17 in *The Social Psychology? of Science,* New York: The Guilford Press.

Gambardella, A. (1994), *Science and innovation,* Cambridge: Cambridge University Press.

Geison , G. L. (1996) "Pasture and the Culture Wars: An Exchange, *New York Review of Books,* XLII(6), April 4: pp.68-69.

Grimmett, G. (1989), *Percolation* (New York: Springer Verlag).

Hammersley, J.M., and DJ. Welsh (1980), "Percolation theory and its ramification," *Contemporary Physics* 21 (6).

Harding, S. (1976), *Can Theories Be Refuted?* (Dordrecht: D. Reidel).

Harris, T.E. (1978), "Additive set-valued Markov processes and percolation methods," *Annals of Probability* 6.

Henderson, R., A. Jaffe and M. Trajtenberg (1995a), "Universities as source of commercial technology: a detailed analysis of university patenting 1965-1988," National Bureau of Economic Research Working Paper No. 5068, March.

Henderson, R., A. Jaffe and M. Trajtenberg (1995b), "The Bayh-Dole Act and trends in university patenting 1965-1988," paper presented to the Conference on University Goals, Institutional Mechanisms and the 'Industrial Transferability' of Research, March (Stanford: Stanford Center for Economic Policy Research); revised August.

Hirshleifer, J. (1971), "The private and social value of information and the reward for inventive activity," *American Economic Review* 61, 561 -74.

Holley, R., and T. Liggett (1975), "Ergodic theorems for weakly interacting systems and the voter model," *Annals of Probability* 3, 643-63.

Hull,, D. (ed.) (1988), *Science as a process* (Chicago: University of Chicago Press).

Imai, K. and Y. Baba (1989), "Systemic innovation and cross-border networks," paper presented to the International Seminar on the Contributions of Science and Technology to Economic Growth, June (Paris: OECD Division of Science, Technology, and Industry).

Jaffe, A.B., M. Trajtenberg and R. Henderson (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations," *Quarterly Journal of Economics.*

Kindermann, R., and J.L. Snell (1980a), "On the relation between Markov random fields and social networks," *Journal of Mathematical Sociology* 7.

Kindermann, R., and J.L. Snell (1980b), *Markov random fields and their applications: Contemporary mathematics, Vol. 1* (American Mathematical Society).

Kirman, A. (1993), "Ants, Rationality and Recruitment," *Quarterly Journal of Economics, 108:137-156.*

Kirman, A., C. Oddou, and S. Weber (1986), "Stochastic Communication and Coalition Formation," *Econometrica,* 54(1): 129-138.

Knorr-Cetina, K. (1981), *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science,* Oxford: Pergamon Press.

Kuhn, T. S. (1962/1970), *The Structure of Scientific Revolutions,*.First/ Second Edition (Chicago: University of Chicago Press).

Lakatos, I. (1970), "Falsification and the methodology of scientific research programmes," in I. Lakatos and A. Musgrave (eds), *Criticism and the growth of knowledge* (Cambridge: Cambridge University Press).

Latour, B., and S. Woolgar (1979), *Laboratory life* (Beverly Hills: Sage Publications).

Leydesdorff, Loet (1995), *The challenge of scientometrics: the development, measurement, and self-organization of scientific communities* (Leiden: DSWO Press).

Luhmann, N. (1984), *Soziale systeme. Grundrisz einer allgemeinen Theorie* (Frankfurt a.M.: Suhrkamp).

Luhmann, N. (1990), *Die Wissenschaft der Gesellschaft* (Frankfurt a.M.: Suhrkamp).

Marschak, J. (1971), "The economics of information," in M. Intriligator (ed.), *Frontiers of quantitative economics,* Contributions to Economic Analysis No. 71 (Amsterdam: North-Holland).

Merton, R.K. (1973), in N.W. Storer (ed.), *The sociology of science: theoretical and empirical investigations* (Chicago: University of Chicago Press).

Morris, S. *(1996),* "Strategic Behavior with General Local Interaction," University of Pennsylvania Department of Economics Working Paper. March.

Mulkay, m. (1979), *Science and the Sociology of Knowledge,* London: George Allen and Unwin.

Narin, F. (1976), *Evaluative bibliometrics* (Cherry Hill, NJ: Computer Horizons, Inc.).

Perutz, and M.F. (1996), "Pasture and the Culture Wars: An Exchange, *New York Review of Books,* XLII(6), April 4: pp.68-69.
Polanyi, M. (1966), *The Tacit Dimension* (London: Routledge and Kegan Paul).

Popper, K. R. (1959), *Logic of scientific discovery* (London: Hutchinson).

Popper, K. R. (1963), *Conjectures and refutations* (London: Routledge & Kegan Paul).

Price, D. J . de Solla (1965), "Networks of scientific papers," *Science* 149, 510-15.

Price, D. J. de Solla (1986), *Little science, big science and beyond (New* York: Columbia University Press).

Quine, W. V. O. (1953), "Two Dogmas of Empiricism," in *From A Logical Point of View,* (Cambridge, Mass.: Harverad University Press).

Quine, W. V. O. (1962), "Carnap and logical truth," in *Logic and language: studies dedicated to Professor Rudolf Carnap on the occasion of his seventieth birthday* (Dordrecht: Reidel).

Quine, W. V. O. (1969), *Ontological Relativity,* New York: Columbia University Press.

Schelling, T. (1971), "Dynamic models of segregation," *Journal of Mathematical Sociology* 1.

Shadish, W. And S. Fuller, eds., *The Social Psychology of Science,* New York: The Guilford Press.

Trajtenberg, M., R. Henderson and A. B. Jaffe (1992), "Ivory tower versus corporate lab: an empirical study of basic research and appropriability," National Bureau of Economic Research Working Paper No. 4146, August.

Usher, Abbott Payson (1982), *A History of Mechanical Invention,* Revised Edition, New York: Dover Publications.

van Rann, A.F.J., ed. (1988), *Handbook of quantitative studies of science and technology* (Amsterdam: Elsevier).

Weitzman, M. L. (1995), "Recombinant Growth," Harvard University Economics Department Working Paper. March.

Worrall, J. (1976), "Thomas Young and the 'Refutation' of Newtonian Optics: A Case-study in the Interaction of Philosophy of Science and History of Science," in *Method and Appraisal in the Physical Sciences,* ed. C. Howson, (Cambridge: Cambridge University Press).

Ziman, J. (1984), *An introduction to science studies: the philosophical and social aspects of science and technology* (Cambridge: Cambridge University Press).

Zuckerman, H. (1977), *The Scientific Elite: Nobel Laureates in the United States*, New York:The Free Press.

Zuckerman, H. and R. K. Merton (1971), "Institutionalization and Patterns of Evaluation in Science," in R K. Merton, *The sociology of science: theoretical and empirical investigations,* N.W. Storer, ed., Chicago: University of Chicago Press, 1973: pp.