

The Explicit Economics of Knowledge Codification and Tacitness

By

Robin Cowan,* Paul A. David,** and Dominique Foray***

*University of Maastricht, MERIT

**Stanford University & All Souls College, Oxford

***University of Paris-Dauphine & CNRS

First draft: November 1998

Revised draft: April 1999

This version: May 1999

Acknowledgements

This paper has been prepared under the EC TSER Programme's TIPIK Project, for presentation to the 3rd TIPIK Workshop held in Strasbourg, at BETA, University of Louis Pasteur, 24th April 1999. We are grateful for the comments and suggestions received from colleagues in the TIPIK Project, although it has not been possible for all of those to be absorbed in the present version.

Please do not reproduce or quote without permission of the authors.

Contact authors at:

P. A. David, All Souls College, Oxford OX1 4AL, United Kingdom

Fax: +44+(0)1865+279299; Tele: +44+(0)1865+279313

E-mail: paul.david@economics.ox.ac.uk; pad@leland.stanford.edu

R. Cowan, MERIT, Maastricht, P.O.Box 616, 6200 Maastricht, Netherlands

Fax: +31+43+3216518; E-mail: racowan@mud.cgl.uwaterloo.ca

D. Foray, IMRI, University of Paris Dauphine, 775 Paris Cedex 16, France

Fax: +33+(0)1+44.05.48.49; E-Mail: dominique.foray@dauphine.fr

The Explicit Economics of Knowledge Codification and Tacitness

by

Robin Cowan, Paul A. David, and Dominique Foray

1. Introduction

With increasing frequency these days references appear in the economics literature to "tacit knowledge". More often than not the meaning of this term itself is something that remains literally tacit— which is to say, those who employ it are *silent* as to its definition. Something is suggested nevertheless by the common practice of juxtaposing mention of tacit knowledge and references to "codified knowledge." What is all this about? Why has this distinction been made, and what significance does it have for economists?

Michael Polanyi (1958, 1967) introduced the term into modern circulation, by pointing to the existence of "the tacit dimension of knowledge," a form or component of human knowledge distinct from, but complementary to the knowledge explicit in conscious cognitive processes. Polanyi illustrated this conceptualization by reference to a fact of common perception: we all are often aware of certain objects without being focused on them. This does not make them the less important, as they form the context that renders focused perception possible, understandable, and fruitful.

Subsequently, the term tacit knowledge has come to be widely applied to forms of personal knowledge that are not readily transmitted as "information" which itself is thought of as an ideal-type good having peculiar economic features that differentiate it from other, conventional economic commodities.¹ As a consequence of the growing practise among economists of referring to tacit knowledge when alluding to what is in effect a residual category of knowledge, the original psychological and contextual usage of the term has been largely discarded. Tacit knowledge thus has come to signify an absolute type, namely: "not codified" knowledge. As such, however, the label now covers the implicit complement of a category of containing a various forms of information and social

¹ Most significant, from the economist's viewpoint, is the absence of super-additivity and the negligible marginal costs of transmitting information. These properties and their implications are discussed more fully in the following text, but the canonical references are Nelson (1959) and Arrow (1962).

knowledge, but which itself usually is left undefined and undifferentiated.

The result has been that a good bit of semantic and taxonomic confusion has been created. These confusions, unfortunately, are being exploited to advance economic policy arguments that claim to be grounded upon well-established propositions about the existence of different kinds of knowledge pertinent to scientific, technological and organizational innovation, whereas, in many instances, such claims are neither analytically nor empirically warranted. In this paper we review the way the concepts of tacitness and codification are currently employed, and suggest a re-conceptualization of this aspect of the economics of knowledge production and distribution.

To motivate our proposals we begin by examining the ways in which "tacitness" has come to figure in contemporary economic policy decisions. The paper's following section elaborates on the respects in which the terminology of tacitness now in circulation, having drifted away from its original epistemological and psychological moorings, has become increasingly amorphous; indeed, that it now obscures more than it clarifies. Among the things that have thus been hidden are analytical and empirical flaws in the emerging critique of the old economics of R&D, and in the new rationale for subsidies based on the intrinsic tacitness of technological knowledge. We contend that our proposed alternative, explicitly economic framework for the study of knowledge codification activities offers a more useful guide for further research directed to informing public policies for science, technological innovation and long-run economic growth.

2. The tacit dimension's wonderful new economics career

Why should economists become involved in discussions of the nature of knowledge, and begin sorting it into various categories? Isn't this something better left for epistemologists and others of similar philosophical inclination? Although one might be disposed to answer the latter in the affirmative, it is now too late to insist on ignoring the very meaning of something that economists and other social scientists seem bent upon discussing.

More than merely a topic of idle talk, tacitness has become an increasingly "loaded" buzzword, freighted with both methodological and policy significance for the economics of science and technology, and economic growth more generally. Indeed, recognition of a tacit dimension of human knowledge has been used both to attack and to defend government subsidization of science and engineering.

It is fair to say that the concept's recent popularization began with the

pioneering work of those who called themselves sociologists of scientific knowledge (SSK), by way of differentiating their approach from that of the Mertonian school in the sociology of science. The 'SSK' approach was more interested in the way sociological forces shaped the cognitive aspects of scientific work, rather than its institutional settings, reward structures, and the like. They observed that some kinds of knowledge deployed in scientific inquiry – such as the functioning of experimental apparatus and instruments, and the way to read and interpret the data obtained with the help thereof– did not seem to be transmitted among the researchers through any explicit statements. H. Collins' (1974) influential study of the construction of the TEA laser concluded: "The major point is that the transmission of skills is not done through the medium of written words." Subsequent contributors to the sociology of scientific and technological knowledge have read such observations as showing that "the diffusion of knowledge could not be reduced to the mere transmission of information" – to quote from Michel Callon's recent and authoritative survey in *The Handbook of Science and Technology Studies* (1995).

A contrast was thus posed between the "algorithmic model" of knowledge production, which involves generation and transmission of explicit, propositional statements, on the one hand, and the so-called "enculturation model" of scientific activity, on the other hand. This distinction was primarily invoked among philosophers and sociologists to challenge the notion that science, and the modern scientific method, was a source of "privileged" statements. The privilege in question implied that scientific statements could be stripped from the social contexts in which these had been formed and in which they had acquired meaning, and promulgated as belonging to the "codified", authoritative body of knowledge about the physical world. Challengers of that view leaned heavily on the importance of tacit knowledge in the actual conduct of scientific activities.

For a while, however, economists took little if any interest in the distinction being drawn between the algorithmic and enculturation models, and its possible sequelae. But things have moved on from that stage. Whether they have "advanced", however, remains another matter. Tacit knowledge is invoked today by those who challenge the "algorithmic" approach of modern economic analysis of all aspects of human behavior. The challenge is raised not only against neo-classical formulations, but to virtually every construction of rational decision processes as the basis for modeling and explaining the behaviors of individual agents. If crucial aspects of human knowledge cannot be explicitly formulated linguistically and articulated, how can it be made a basis for optimization, or indeed, consistent calculation of any sort? One may see the seeds of this flowering critique of neoclassical economics in the early attention paid to Polanyi's (1958, 1967) writings by Richard Nelson and Sidney Winter (1982); and, subsequently, in the rationale offered for the more thorough-going

rejection of microeconomic optimization in the evolutionary models of Schumpeterian competition.²

The challenge brought against algorithmic representations of knowledge generation and acquisition, however, goes much deeper than arguments for bounded rationality following Newell and Simon (1972); and it has been argued in far more sweeping, general terms by critics of the whole Artificial Intelligence program, such as Douglas Hofstadter (1979) and, more recently, Roger Penrose (1989, 1997). In such quarters, tacit knowledge has come to stand for the aspects of human intelligence that cannot be mimicked by any (computer) algorithm.

The lead in bringing the significance of "tacit knowledge" to the attention of economists was taken by Richard Nelson and Sidney Winter (1982), whose discussion of the parallels between individual human skills and organizational capacities (in *An Evolutionary Theory of Economic Change*, Ch. 4) gave prominence to Polanyi's concept. Those passages remain as perceptive and stimulating, and as fresh and balanced today as when they first appeared 17 years ago, and it is unfortunate that a larger proportion of those economists who now talk about tacit knowledge and its implications do not appear to have acquainted themselves with this "local source." Had things been otherwise, it seems only reasonable to suppose that we would have been spared at least the more serious confusions and unwarranted generalizations that have become commonplace in the literature.

What Nelson and Winter (1982) say about the nature and significance of tacitness in knowledge conveys not one sharply defined concept, but a nexus of meanings, each carrying somewhat distinctive implications. Their first reference to the term (1982, p. 73), for example, offers only a parenthetical clarification:

"The knowledge that underlies skillful performance is in large measure tacit knowledge, *in the sense that* the performer is not fully aware of the details of the performance and finds it difficult or impossible to articulate a full account of those details."
(Emphasis added.)

Yet, as was made clear shortly following this statement, Nelson and Winter accepted Polanyi's (1967) account of such situations as being contextual, rather than absolute: "the aim of a skillful performance" may "be

² See, for example, for example, Dosi (1988), Dosi et al. (1988), Dosi and Egidi (1991). Evolutionary modeling in economics now spans a wide range of practice regarded how "bounded" the bounded rationality of agents is assumed to be. Anderson (1994) discusses this and other issues in the algorithmic representation of the general class of "Nelson-Winter type" models.

achieved by the observance of a set of rules which are not known as such to the person following them." Reference was made to Polanyi's earlier philosophical work, *Personal Knowledge* (1958, p. 49), where an example is presented of a swimmer keeping himself buoyant by regulating respiration, yet remaining unconscious of doing so. In this case the operant rule ("never empty your lungs fully") plainly is one that is articulable, could be known to another person, and so might be transmitted verbally by a swimming instructor—were the latter aware of the principle of buoyancy. In other words, Nelson and Winter (1982, p. 77) do not insist, any more than did Polanyi, that tacitness implied "inarticulability"; even though the inarticulability of some (personal) knowledge logically implied that the latter would remain "tacit".

On the question of "awareness", Nelson and Winter (1982, p. 78) recognize that the skillful performer may have "subsidiary awareness" of the rules that are being followed, while being "focally aware" of some other—most probably novel—facet of the task in which she is engaged. This reinforces an appreciation of the contextual boundaries within which knowledge will be "tacit" rather than explicitly recognized and articulated. Yet, if one can describe behavior in terms of "rule conformity", then it is clear that the underlying knowledge is *codifiable*—and indeed may have previously been "codified", or at least articulated.

Most significant still for what we shall say about the more recent strain of the literature on tacitness is Nelson and Winter's (1982, p. 78) acknowledgement that this quality *is not inherent in the knowledge*. They write: "The same knowledge, apparently, is more tacit for some people than for others. Incentives, too, clearly matter: when circumstances place a great premium on effective articulation, remarkable things can sometimes be accomplished." In amplification of this point, they offer the example of an expert pilot giving successful verbal instruction via radio to a complete novice as to how to land an airplane— even though the "expert" had never had occasion previously to make explicit what was entailed in his successful performance of a landing. Indeed, the section on "Skills and Tacit Knowledge" concludes by emphasizing that "costs matter. Whether a particular bit of knowledge is in *principle* articulable or necessarily tacit is not the relevant question in most behavioral situations. Rather, the question is whether the costs...are sufficiently high so that the knowledge *in fact* remains tacit" (p. 80).

It is unfortunate that these more complicated aspects of the concept have been all but forgotten— were they ever widely grasped— when "tacitness" made its debut on the economic policy stage. Among the most notable of the uses to which the idea of tacit knowledge is being put on the more mundane levels at which most economists operate, and certainly the uses that have the greatest impact in economic policy circles, has been the qualification— and in some instances the outright rejection— of the

practical policy conclusions drawn from the classic information-theoretic analysis of the economics of R&D activities.

Following the seminal work of Kenneth Arrow (1955, 1962) and Richard Nelson (1959), an entire generation of economists treated scientific and technological knowledge as "information." To that degree, they reasoned, the knowledge generated by research activities possessed certain generic properties of public goods. Much of the case for government subsidization of science and engineering research, and for innovative activity more generally came to be grounded on the proposition that *qua* information, such knowledge could not be optimally produced or distributed through the workings of competitive markets.

But, nowadays, we are more and more frequently instructed otherwise. In the newer understanding of science and technology as being pursuits inextricably involved with tacit knowledge, it is claimed that the old public policy rationales are exploded; the essential understandings are said to be the portion of knowledge that remains uncodified, and so deprived of the public goods properties that would result in informational spillovers and market failure. Thus, as this argument concludes, the traditional economic case for subsidizing science and research in general collapses, as there is little or no basis for a presumption of market failure.

But, at the same time, tacit knowledge is invoked by defenders of government subsidization of science as part of a strategic innovation policy. A standard argument against public subsidy to science is that foreigners engaging in applied, commercially oriented R&D would free-ride (since information is a public good and travels freely) by exploiting the basic knowledge discoveries that our researchers vie to codify for disclosure in the scientific journals and similar archival publications. To this, the proponents of tacit knowledge reply, nations, and regions, like individual enterprises undertaking R&D investments, can count on the benefits of "sticky data" – to use Eric von Hippel's (1993) arresting term. Knowledge does not travel freely, a condition that rests largely on the importance of tacit knowledge residing only in the heads of the scientists and engineers engaged in its production. Codified knowledge may have low marginal costs of transmission and is thus slippery and hard to contain, but that is largely irrelevant if what one needs is its "sticky" , tacit counterpart. ³

³ In subsequent work von Hippel (1994) generalizes the idea so that it covers all situations in which there is an appreciable cost of transferring information, especially that relevant for innovative activities. In principle, at least, von Hippel's use of the notion of "stickiness" makes no distinction between transfer costs consisting of pure rents imposed by

The inherent "stickiness" of certain kinds of knowledge, consequently, enables business (or other) entities to protect their ability fully to appropriate the benefits derivable from their research investments by controlling access to the repositories of uncodified knowledge. For this, minimal recourse is required to the protection of intellectual property in the form of patents and copyrights; a mixture of trade secrecy law and labor law (master-servant relations) governing the behavior of current and former employees, may be enough. Thus, curious though it may seem, the tacit dimension of scientific and technological knowledge has found a new career for itself in science and technology policy debates: it is beginning to supplant its now dubious companion, "codified knowledge", as the core of a new rationale for government research funding intended to build national and regional "competitiveness" through innovation.

According to this application of the argument, even though the essential tacit knowledge concerning how to exploit what has been invented might be less than perfectly "sticky", what this implies is that its economic benefits are only available to be captured locally. In other, more formal, terms, it is asserted that the marginal costs of knowledge transmission rise very rapidly with "distance" from the context in which such knowledge was generated. Research by-products in the form of technological knowledge—being concerned with how best to get instrumentation involving chemical, mechanical, electrical and optical processes to work—are seen as inherently more strongly tacit in nature. That is held to be particularly beneficial for would-be commercial developers who are able to situate closer to the locus of such discoveries (see e.g. Pavitt (1987), Patel and Pavitt (1995), Nelson (1992)).

A broad policy implication, following from this, is that for an economy to have a strong, innovative manufacturing sector, it is necessary also to have correspondingly strong applied and basic research activities situated in close proximity to the production operations themselves. The following passage extracted from an article by John Kay (1999, p. 13) in *Science and Public Affairs*, is illustrative of the more elaborate statements of the new innovation strategy perspective that has now formed around the concept of tacitness in the business management literature:

" Since 'knowledge that'— the characteristic discoveries of natural science— is easily transmitted, one solution [to the problem of creating 'knowledge-based competitive advantages'] is to continually innovate and stay one step ahead. And that kind of innovative capacity depends on knowledge that isn't 'knowledge

the owners of intellectual property rights, on the one hand, and real social resource costs such as those entailed in physically transporting an expert for the purpose of demonstrating the proper use of a novel product or process in a distant location.

that', but 'knowledge how'— i.e. tacit knowledge.

" Tacit knowledge can take many forms, but it cannot be written down. It is unique to an organization— and therefore cannot be copied...The benefits of such tacit knowledge arise only through a culture of trust and knowledge-sharing within an organization...

" Tacit knowledge isn't a phenomenon exclusively existing within a single organization...It also exists between companies, consumers, suppliers and scientific laboratories..."

Thus, a notion that took its origins in the psychology of individual human motor skills has been wonderfully transmuted, first, from an epistemological category into a phenomenon of inarticulable inter-organizational relationships, and thence to one of the keys to corporate, and perhaps also national competitive advantage!

One corollary of this class of arguments is that the case for granting public subsidies and tax concessions to private companies that invest in R&D would correspondingly be weakened, were it not for the difficulties caused these firms by the circulation of their research personnel.⁴ Scientific and engineering staff are able to carry critical tacit knowledge off to potential rival firms that offer them better terms of employment, including equity ownership in " start ups" of their own. In the logic of this approach, recognition of the criticality of tacit knowledge argues for further strengthening of trade secrecy protections, to block those " leakages" and altogether eliminate the market failure rationale for governmental support of the performance of R&D by the private sector.⁵ That leaves the way open for those who wish to mount an essentially " techno-mercantilist" argument for R&D subsidies, grounded on the idea that the country can benefit from job-creation, etc., if its firms win the race to be first to launch new products in international markets. It is, in effect, a new strategic trade policy argument, grounded on the claim that tacit knowledge permits national appropriation of the direct and indirect benefits of monopolizing international product riches by being " first to

⁴ We can observe that the more things change the more they stay the same. We have moved from the view that the problem to be solved arises from the fact that a firm's *knowledge* is easily appropriated by other firms. Acknowledging the importance of tacit knowledge, and thus that the initial problem may not be so severe, we face a " new problem" stemming from the fact that a firm's *knowledge workers* are easily appropriated by other firms. In both cases the general issue remains however — fluidity of knowledge or information (whether transmitted through codified knowledge or labor mobility) is good for the economy but bad for the individual firm.

⁵ See, e.g., Kealey (1996) on industrial secrecy as the suitable " remedy" for the problem of informational spillovers from research, and the critique of that position in David (1997).

invest" .

The foregoing brief account of the wonderful career that has thrust " the tacit dimension" into the science and technology policy limelight serves to identify a number of issues that seem of sufficient importance to warrant further consideration. We see a need to put the economics of tacit and codified knowledge on analytically firmer and terminologically more precise foundations than those upon which most of the recent literature presently rests. The notion that the economic case for public support of science and engineering should now be based upon the inherently tacit and " craft" nature of research activities certainly is rather paradoxical. Taken at face value, it suggests that intellectual property protection is unjustified, since, in the " natural" state of things, there are no " externalities" of new knowledge. By implication, the patent system's exchange of monopoly of use for " disclosure" allows the patentee to retain the tacit knowledge without which the information contained in the patent really is useless.

But, before rushing to discard everything we know about the economics of R&D, however, scientific scepticism instructs us to ask whether the epistemological foundations of the argument are really all that solid. We should therefore question whether the economic functioning and attributes of tacit and codified knowledge are well understood by those who would invoke it in the context of current policy debates.

This essay responds to these felt needs by presenting what might be described as " the sceptical economist's guide to 'tacit knowledge'" . We agree that the questions economists and others have been discussing under the general rubric of tacit knowledge are truly important. Furthermore, we acknowledge that some of the now-classic contributions to the economics of R&D, when regarded from an epistemological perspective, appear unwarrantedly simplistic in their handling of some subtle questions concerning " knowledge" and " information" , and the relationship between the two. An explicit re-examination of some fundamental conceptual underpinnings in this area requires that we re-open some questions that readers coming to the topic from economics may feel are settled well enough for their purposes. We believe that the next section (3) makes a persuasive case to the contrary, and so establishes that there is a need for a new taxonomic framework in this area. That framework is set out (in sections 4 and 5), and we proceed to apply it in analyzing the economic determinants of codification (in sections 6 and 7).

3. Codification and tacitness

It will be easiest for us to start not with concept of tacit knowledge, but at the opposite. and seemingly less problematic end of the field, so to

speak, by asking what is to be understood by the term "codified knowledge." Its obvious reference is to *codes*, or to standards— whether of notation or of rules, either of which may be promulgated by authority or may acquire "authority" through frequency of usage and common consent, by *de facto* acceptance.

Knowledge that is recorded in some codebook serves *inter alia* as a storage depository, as reference point, and possibly as an authority. But information written in a code can only perform those functions when people are able to interpret the code; and, in the case of the latter two functions, to give it more or less mutually consistent interpretations. Successfully reading the code in this last sense may involve prior acquisition of considerable specialized knowledge (quite possibly including knowledge not written down anywhere). There is as a rule no reason to presuppose that all people in the world possess the knowledge needed to interpret the codes properly. This means that what is codified for one person or group may be tacit for another, and an utterly impenetrable mystery for a third. Thus *context*— temporal, spatial, cultural and social— becomes an important consideration in any discussion of codified knowledge.

In what follows, we make use extensively of the notion of a codebook. We use "codebook" to refer both to what might be considered a dictionary that agents use to understand written documents, but apply it also to cover the documents themselves. This implies several things regarding codification and codebooks. First, codifying a piece of knowledge adds content to the codebook. Second, codifying a piece of knowledge draws upon the pre-existing contents of the codebook. This creates a self-referential situation, which can be particularly severe when the knowledge activity takes place in a new sphere or discipline. Initially, there is no codebook, either in the sense of a book of documents, or in the sense of a dictionary. Thus initial codification activity involves creating the specialized dictionary. Models must be developed, as must the vocabulary with which to express those models. When models and a language have been developed, documents can be written. Clearly, early in the life of a discipline or technology, standardization of the language (and of the models) will be an important part of the collective activity of codification. When this "dictionary" aspect of the codebook becomes large enough to stabilize the "language", the "document" aspect can grow rapidly.⁶ But new documents will inevitably introduce new concepts, notation and terminology, so that "stabilization" must not be interpreted to imply a complete cessation of dictionary-building.

The meaning of "codification" intersects with the recent literature on economic growth. Much of modern endogenous growth theory rests on the

⁶ See Cowan and Foray (1997) for a further discussion of this issue.

notion that there exists a "world stock of knowledge," and, perhaps, also a "national knowledge-base" that has stock-like characteristics. This is true particularly of those models in which R&D is seen as both drawing upon, and adding to, "a knowledge stock" which enters as an input into production processes for other goods. How ought we to characterize this, or indeed any related conceptualization of a world stock of knowledge? Implicit in this literature is that this stock is codified, since part or all of it is assumed to be freely accessible by all economic agents in the system under analysis. Unpacking this idea only partially suffices to reveal some serious logical difficulties with any attempt to objectify "a social stock of knowledge," let alone with the way that the new growth theory has sought to employ the concept of an aggregate knowledge stock.⁷

The "new growth theory" literature falls squarely within the tradition emphasizing the public-goods nature of knowledge. So, one may surmise that the world stock of knowledge surely has to be the union of private stocks of codified knowledge: anything codified for someone is thereby part of the world knowledge stock. Such reasoning, however, may involve a fallacy of composition or of aggregation. One might reasonably have thought that the phrase "world knowledge stock" refers to the stock available to the entire world. But if the contextual aspect of knowledge and codification (on which see *supra*) is to be taken seriously, the world stock of codified knowledge might better be defined as the *intersection* of individuals' sets of codified knowledge— that being the portion that is "shared" in the sense of being both known and commonly accessible.⁸ It then follows that the world stock of knowledge, being the intersection of private stocks, whether codified or tacit, is going to be very small.

The foregoing suggests that there is a problem in principle with those models in the "new growth theory" which have been constructed around (the formalized representation of) a universal stock of technological knowledge to which all agents might contribute and from which all agents can draw costlessly. That, however, is hardly the end of the difficulties arising

⁷ See Machlup (1980: pp.167-169) for discussion of "the phenomenological theory of knowledge" developed by Alfred Schutz and Thomas Luckmann (1973: Chs. 3-4). The latter arrive at a concept described by Machlup as "the fully objectivated knowledge of society, a social stock of knowledge which in some sense is the result of a socialization of knowledge [through individual interactions involving private stocks of subjective but intersubjectively valid knowledge] and contains at the same time more and less than the sum of the private stocks of subjective knowledge...This most ingenious phenomenological theory of the stock of knowledge in society is not equipped to deal with ...the problem of assessing the size of the stock and its growth."

⁸ It is clear that the availability of two operators— union and intersection— when combined with two types of knowledge— tacit and codified— leads to a situation in which "the worldstock of knowledge" is going to take some defining.

from the primacy accorded to the accumulation of "a knowledge stock" in the recent literature on endogenous economic growth. The peculiarities of knowledge as an economic commodity, namely, the heterogeneous nature of ideas and their infinite expansibility, have been cast in the paradigm "new economic growth" models as the fundamental non-convexity, responsible for increasing returns to investment in this intangible form of capital. Heterogeneity implies the need for a metric in which the constituent parts can be rendered commensurable, but given the especially problematic nature of competitive market valuations of knowledge, the economic aggregation problem is particularly vexatious in this case.

Futhermore, the extent to which the infinite expansibility of knowledge actually is exploited, therefore, becomes a critical matter in defining the relevant stock -- even though in these formulations of the new growth theory this matter generally has been glossed over. Critics of these models' relevance have quite properly pointed out that much technologically relevant knowledge is not codified, and therefore has substantial marginal costs of reproduction and re-application; they maintain that inasmuch as this so-called "tacit knowledge" possesses the properties of normal commodities, its role in the process of growth approaches that of conventional tangible capital.⁹ If it is strictly complementary with the codified part of the knowledge stock, then the structure of the models implies that either R&D activity or some concomitant process must cause the two parts of the aggregate stock to grow *pari-passus*. Alternatively, the growth of the effective size of the codified knowledge stock would be constrained by whatever governs the expansion of its tacit component. Pursuing these points further is not within the scope of this paper, however; we wish merely to stress once again, and from a different perspective, that the nature of knowledge, its codification or tacitness, lurks only just beneath the surface of important ideas about economic growth.

Leaving to one side, then, the problematic issue of defining and quantifying the world stocks of either codified knowledge or tacit

⁹ This view could be challenged on the grounds that knowledge held secretly by individuals is not distinguishable from labor (tangible human capital) as a productivity input, but, unlike tangible physical capital, the existence of undisclosed knowledge assets cannot be ascertained. Machlup (1980, p. 175), in the sole passage devoted to the significance of tacit knowledge, adopts the latter position and argues that :
" Generation of socially new knowledge is another non-operational concept as long as generation is not complemented by dissemination...Only if {an individual} shares his knowledge with others can one recognize that new knowledge has been created. Generation of knowledge without dissemination is socially worthless as well as unascertainable. Although 'tacit knowledge' cannot be counted in any sort of inventory, its creation may still be a part of the production of knowledge if the activities that generate it have a measureable cost."

knowledge, we can now turn to a fundamental empirical question regarding tacit knowledge. Below, we address explicitly whether some situations, described as rife with tacit knowledge, really are so, but for the moment we can make an important point without entering into that issue.

Some activities seem to involve knowledge that is unvoiced— activities which clearly involve knowledge but which refer only seldom to texts, or put another way, which clearly involve considerable knowledge beyond the texts that are referred to in the normal course of the activity.¹⁰ Thus we can ask why is some knowledge silent, or unvoiced? There are two possible explanations: the knowledge is unarticulable, or, being capable of articulation, it remains unvoiced for some reason.

Why would some knowledge remain unarticulable? The standard economist's answer is simply that this is equivalent to asking why there are "shortages", to which one must reply "there are no shortages" when there are markets. So, the economist says, knowledge is not articulated because, relative to the state of demand, the cost, and supply price, is too high. Articulation, being social communication, pre-supposes some degree of codification, but if it costs too much actually to codify, this piece of knowledge may remain partly or wholly uncoded. Without making any disparaging remarks about this view, we can simply point out that there is some knowledge for which we do not even know how to begin the process of codification, which means that the price calculation could hardly be undertaken in the first place. Recognition of this state of affairs generates consensus on the uncodifiable nature of the knowledge in question. We raise this point to emphasize that the category of unarticulable (which may be co-extensive with uncodifiable) may be safely put to one side in what follows. That, of course, supposes there is still a lot left to discuss.

It is worth emphasizing the existence of the two distinctions, and the degree to which they define co-extensive sets of knowledge. Knowledge that is unarticulable is also uncodifiable, and vice versa— if it is (not) possible to articulate a thought so that it may be expressed in terms that another can understand, then it is (not) possible to codify it. This is the source of the statement above that articulation pre-supposes codifiability. It is not the case, though, that codifiability necessitates codification; a paper may be thought out fully, yet need not actually be written out. Operationally, the codifiability of knowledge (like the articulable nature of a thought) cannot be ascertained independently from the actions of codification and articulation. But, when we consider the question of the status of knowledge with reference to multiple contexts, the preceding

¹⁰ We note that activities involving "unvoiced knowledge" are often assumed to involve thereby tacit knowledge. We argue below that this is too hasty.

strictly logical relations (implied by a single, universal context) are not exhaustive categories. Thus we see the possible emergence of a category: codified (ever) but not articulated (now, here).¹¹ This observation implies that care needs to be taken in jumping from the *observed* absence of codified knowledge in a specified context to the conclusion that only some non-codifiable (i.e. tacit) knowledge is available or employed.

It is within the realm of the codifiable or articulable-yet-uncodified, that price considerations come into play in an interesting way. Within this category reside decisions about the activity of codification, its costs and benefits. We shall discuss the factors entering into the determination of that knowledge-status more fully below.

4. A proposed knowledge topography

With these preliminaries behind us, we now proceed to examine a knowledge topography from which it will soon be self-evident that the realm of the tacit has been considerably reduced.

The topography we are about to propose is meant to be consulted in thinking about where various knowledge transactions or activities take place, rather than where knowledge of different sorts may be said to reside. We should emphasize that as economists, and not epistemologists, we are substantively more interested in the former than in the latter.

By knowledge activities we refer to two kinds of activities: the generation and use of "intellectual (abstract) knowledge" ; and the generation and use of "practical knowledge" , that is mainly knowledge about technologies and artifacts (how to use this tool, this car, or how to improve their performances) and about organizations.

Given that definition, we need to clarify the distinction between knowledge embodied in an artifact and codified knowledge about an artifact. The distinction between embodied and disembodied knowledge is a nice way for economists to capture features of intersectoral flows (of technologies), particularly in an input-output framework. Therefore, the fact that knowledge is embodied in a machine tool is not to be conflated with the codification problem. However, knowledge about the production and the use of artifacts, of course, falls in our set of issues about codification: does the use of this new tool requires the permanent reference to a set of

¹¹ In understanding these distinctions it is important to remember that we are discussing knowledge activities, and the kinds of knowledge used in them. Thus we can observe activities in which the knowledge has been codified at some point in history, but is not articulated in current endeavors.

codified instructions or not? We can make this point in a slightly different way. From the point of view of a producer, any artifact, from a hammer to a computer, embodies considerable knowledge. The artifact is often an exemplar of that knowledge, and can sometimes be thought of as a store of it, and the means through which the knowledge is marketed. From the point of view of the user, however, this is not the case. While any user will admit that the producer needed a variety of kinds of knowledge to produce the artifact, this is of little practical interest. The knowledge of interest, whether codified or not, and indeed that is often the issue, is how to use the artifact, rather than the knowledge embedded in it.

Part of the reason for this interpretation of what is located in our topography is simply that discussion about "where knowledge resides" is difficult without falling into, or attempting to avoid, statements about the relative sizes of the stocks of tacit and codified knowledge, and their growth rates. By and large entering into pseudo-quantitative discussions of that sort turns out not to be a useful line; indeed, possibly worse than unhelpful, it can be misleading. Although there is no scarcity of casual assertions made regarding the tendency toward increasing (relative) codification, the issue of the relative sizes of the constituent elements of the world stocks of scientific and technological knowledge, resists formal quantitative treatment. That is to say, we really cannot hope to derive either theoretical propositions or empirical measures in answer to queries about whether or not the relative size of the codified portion must be secularly increasing, or, alternatively, whether there is a tendency to a steady state. The fundamental obstacle to approaching such questions meaningfully is the vagueness regarding the units in which "knowledge" is to be measured.

To begin, we shall consider a topological tree structure in which distinctions are drawn at four levels. (The following text refers to Figure 1.)

The tripartite branching on the uppermost level breaks the knowledge transaction terrain into the zones labeled articulated (and therefore codified), unarticulated and unarticulable. Setting the third category aside as not very interesting, we are left with the dichotomy:

(a) *Articulated (and thus codified)*. Here knowledge is recorded and referred to by 'the group', which is to say, 'in socio-temporal context'. Hence we can surmise that a codebook exists, and is referred to in the usual or standard course of knowledge-making and -using activities.

(b) *Unarticulated*. Here, we refer to knowledge that is not explicitly referred to in the typical course of knowledge activities. Again, the concept of a context or group is important.

In case (a) a codebook clearly exists, since this is implicit in knowledge being or having been codified. In case (b) two possible sub-cases can be considered. In one, knowledge is tacit in the normal sense— it has not been recorded either in word or artifact, so no codebook exists. In the other, knowledge may have been recorded, so a codebook exists, but this book may not be referred to by members of the group— or, if it is, references are so rare as to be indiscernible to an outside observer. Thus at the next level, “unarticulated” splits into two branches: (b.1) In the situation indicated to the left, a source or reference manual does exist but it is out of sight, so we say the situation is that of a *displaced codebook*; (b.2) to the right lie those circumstances in which there truly is no codebook, but in which it would be technically possible to produce one.

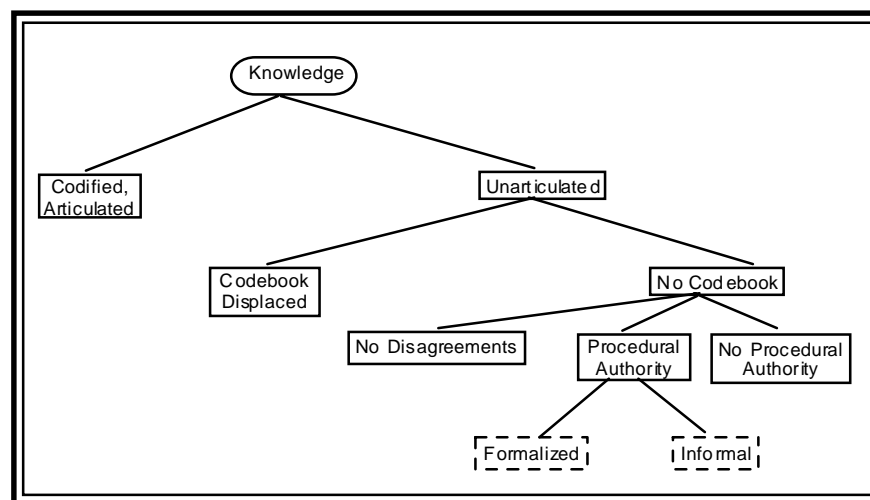


Figure 1: A tree structure for codified and uncoded knowledge

When a codebook exists, we still may refer to the situation in which knowledge is unarticulated because within the group context the codebook is *not manifest*; it is not explicitly consulted, nor in evidence, and an outside observer therefore would have no direct indication of its existence. (To the outside observer, this group *appears* to be using a large amount of tacit knowledge in its normal operations.) The contents of the codebook in such situations have been so thoroughly internalized, or absorbed by the members of the group, that it functions as an implicit source of authority. Here we can observe in passing that the ability to displace the codebook (e.g. pass closed-book exams) is often one criterion for entry, or part of the initiation, into the group.

A “displaced codebook,” implies that a codified body of common knowledge is present, but not manifestly so. Technical terms figure in descriptive discussion but go undefined because their meaning is evident to all concerned; fundamental relationships among variables also are not reiterated in conversations and messages exchanged among members of the

group, or epistemic community. (This often infuriates outsiders who complain vociferously about excessive jargon in the writings and speeches of physicists, economists, and psychologists...) In short, we have just described a typical state of affairs in what Thomas Kuhn (1962) referred to as "normal science"; it is one where the knowledge base from which the researchers are working is highly codified, but, paradoxically, its existence and contents are matters left tacit among the group unless some dispute or memory problem arises. We can also describe "normal technology" as the state in which knowledge about artifacts is highly codified but the codebook is not manifest.

Identifying such a zone in which knowledge is codified but the existence of codification is not manifest is an extremely important result. It raises, however, a very difficult empirical problem (or perhaps a problem of observation). This argument is crucial in understanding the economic problem raised by the management of knowledge in various situations: when the codebook is displaced and knowledge is highly codified, new needs for knowledge transfer or storage (or knowledge transactions generally) can be fulfilled at a rather low cost (it is the cost of making the existing codebook manifest), while when there is no codebook at all the costs will be very high (it is the cost of producing a codebook, which includes costs of developing the languages and the necessary models).

Such a result should lead us to re-examine many empirical studies, which have been done in the near past and whose main conclusions are that tacit knowledge remains key in many activities. This is perhaps true but difficult to document convincingly; all these studies fail to prove that what is observed is "true tacitness" rather than highly codified knowledge without explicit reference to the codebook. Identifying the possibility that highly codified knowledge exists but is not manifest makes us aware of the serious empirical difficulties raised by this issue. Differentiating among the various cases certainly requires deep and careful case studies. By definition a codebook which is not manifest is not observable in that context, and there is a high risk that various situations can be confused.

When there is no codebook (case b), we again have a two-way division, turning on the existence, or not, of disputes. There may be no disagreements. Here there is stabilized uncodified knowledge, collective memory, convention and so on. This is a very common situation with regard to procedures and structures within organizations. The IMF has not written that there is only one prescription for all the monetary and financial ills of the world's developing and transition economies, but its advisers dispensing "identikit" loan conditions evidently behave as if such a "code" had been promulgated. Such uncodified-but-stable bodies of knowledge and practice, in which the particular epistemic community's members silently concur, will often find use as a test for admission to the

group or a signal of group membership to outside agents.

Where there are disagreements and no codebook is available to resolve them within the group, it is possible that there exist some rules or principles for dispute resolution. But, elsewhere, such "procedural authority" may be missing. This is the case of individual gurus, such as business management gurus like Tom Peters, and others who supply a form of "personal knowledge about organizational performance." Equivalently, in terms of the outward characteristics of the situation, this might describe "new age" religions— in contradistinction to structured ecclesiastical organizations that refer to sacred texts.

On the other hand, it may be the case that when disagreements arise there is some procedural authority to arbitrate among the parties. Here it is important to recall that by construction the situation is one in which the relevant knowledge is not codified, and different members of the organization/group have different tacit knowledge. When the tacit knowledge leads to conflict about how to advance the group's enterprise or endeavor, there is some way of deciding how to proceed— whether or not this has been explicitly described and recorded. Clearly, when the procedure is formalized (codified) we have a recurrence of a distinction paralleling the one drawn at the top of the tree, between codified and tacit. But this new bifurcation occurs at the meta-level of *procedures* for generating and distributing knowledge, rather than over the contents of knowledge itself. We can define a different type of group using this meta-level codified-tacit boundary, so the whole taxonomic apparatus may be unpacked again in discussing varieties of procedural knowledge.

5. Boundaries in the re-mapped knowledge space, and their significance

Across the space described by the foregoing taxonomic structure it is possible to define (at least) three interesting boundaries. The "Collins-Latour-Callon" boundary would separate articulated codified knowledge from all the rest— assigning observational situations in which there was a displaced codebook to the same realm as that in which learning and transmission of scientific knowledge, and *praxis*, was proceeding in the absence of codification. The line labeled "the Merton-Kuhn frontier" puts codified and codebook-displaced situations together on its left side, and would focus primary attention there— as it constituted the distinctive regions occupied by modern science; that would leave all the rest to general psychological and sociological inquiries about "enculturation processes" involved in human knowledge acquisition.

But, our branching structure recognizes that in the zone to the right of the Merton-Kuhn frontier there are things that allow *epistemic communities* to function. Such communities, which may be small working groups, comprise

agents who work on a mutually recognized sub-set of knowledge issues, and who at the very least accept some commonly understood procedural authority as essential to the success of their collective knowledge-building activities. They are separated from the region that lies to their right in the topography, by the line labeled the "functional epistemic community boundary." Beyond that border lies the zone populated by personal (and organizational) gurus of one shape or another, including the "new age" cult leaders in whom procedural and personal authority over the conduct of group affairs has been fused.

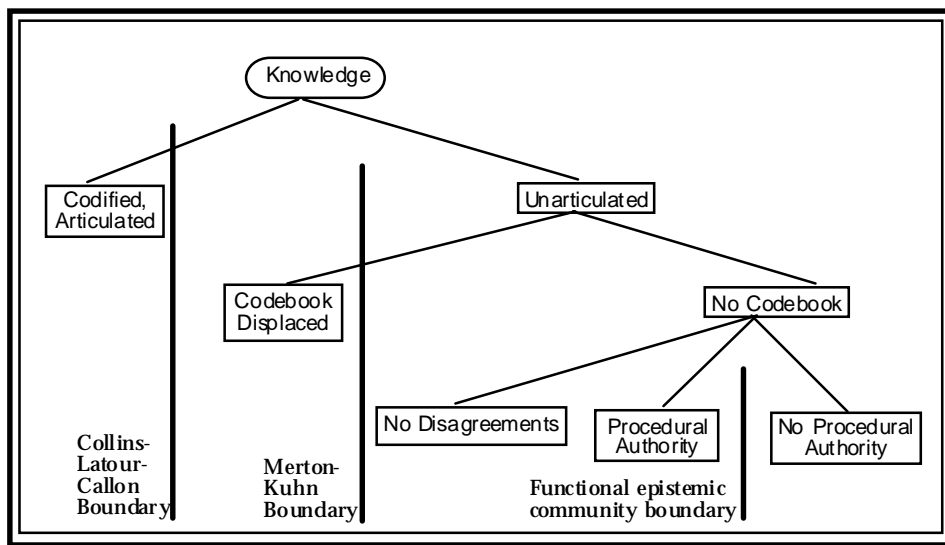


Figure 1a: Boundaries in the Knowledge Space

As is clear from the foregoing discussion, there are two quite distinct aspects of knowledge that are pertinent in the codified/tacit discussions, although they are often left unidentified. On the one hand, knowledge might or might not be presented or stored in a text. This is the notion associated with codification. On the other hand, there is the degree to which knowledge appears explicitly in standard activities. Here, we can think of knowledge as being *manifest* or not. Figure 2 elaborates these two properties in a tableau. For this purpose we have used a 3x3 matrix, in which one axis represents the extent of codification: Codified, Partially codified, and Uncodified; the other axis represents the extent to which the knowledge is manifest, or commonly referred to in knowledge endeavors: Manifest, Alluded to, and Latent.

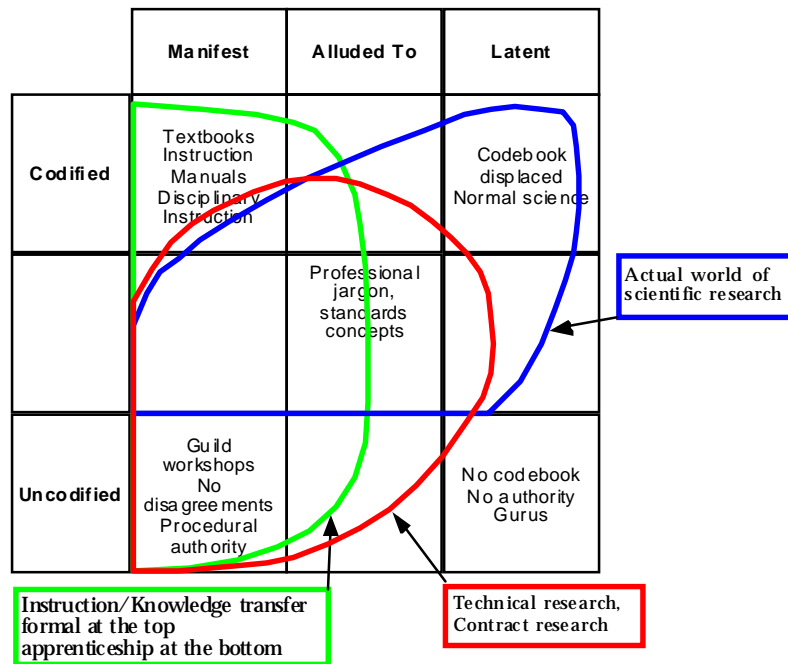
These divisions along the axes are patently arbitrary, for mixtures in the ordinary human knowledge activities form a continuum, rather than a set of boxes.

To make Figure 2 clearer, it may be useful to look specifically at the four extreme cases: the corners north-west, south-west, north-east and south-east. Both the codified-manifest case (the north-west corner) and the uncoded-latent case (the south-east corner) describe situations which

are easily comprehensible because the pair of criteria fit naturally together. The codified-latent case (the north-east) was described as a situation in which the codebook is displaced while knowledge is not tacit. Finally the uncoded-manifest case (south-west) describes situations in which agents start to make their discoveries, inventions, new ideas manifest (in order to diffuse them) but still cannot use a full and stabilized codebook to do so. And even "writing" a book with which to make this new knowledge manifest does not necessarily imply codification. The vocabulary or symbolic expression is still idiosyncratic, there are many ambiguities, and so on. This implies that while certain aspects of codification may be present (knowledge storage and recall for example), other important aspects, in particular communication, are not.

Nevertheless, the overly sharp co-ordinates give us a tableau that may be used heuristically to distinguish among major regions of the states-space within which knowledge-groups may be working at a given moment in their history. Instruction or deliberate knowledge transfer is thus roughly situated in the tableau's "manifest" column, spilling over somewhat into the "Alluded-to" column. Formal instruction comes near the top (codified), whereas apprenticeship lies near the bottom (uncodified) of the array. The world of scientific inquiry extends across the "ham-shaped" region that is oriented along the southwest-northeast axis (minor diagonal) of the array, leaving out the north-west corner. That corner is where codified knowledge is most plainly manifested, for purposes of didactic instruction, making use of textbooks, grammars and dictionaries, manuals, reference standards, and the like. Some apprenticeship training, as we have remarked, falls into the "codified and manifest space," although even formal computer-based tutorial programs occupy the liminal zone where codified knowledge is "alluded-to" rather than made explicit.

Figure 2: Classification of knowledge and knowledge generation activity on two axes



The boundaries of the world of engineering and applied R&D also extend upwards from the south-west corner (in which closed, proprietary research groups function on the basis of the uncoded skills (experience-based expertise) of the team members, and their shared and manifest references to procedures that previously were found to be successful). But, if we are to accept the descriptions provided for us by articulate academic engineers, those boundaries are more tightly drawn than the ones within which science-groups operate, and in particular do not reach as far upwards into the area where there is a large body of latent but nevertheless highly codified knowledge undergirding research and discovery.

6. On the value of this re-mapping

The topography of knowledge that we have presented serves to increase the precision and to allow greater nuance in the distinctions made among types of knowledge-getting and transferring pursuits. In addition, it permits amore fruitful examination of the influence of external, economic conditions upon the codification and manifestation of knowledge as information. A number of specific benefits that we see in this way of looking at the world bear further examination, to which we may now turn.

6.1 On the topography itself

(1a) Figures 1 and 2 clean up a confusion concerning the putative tacitness of the working knowledge of scientists in situations that we have here been

able to characterize by applying the 'displaced codebook' rubric. Previously, the literature seems to have fallen to reliance upon an overly casual observational test, treating situations where no codebook appeared in evidence as instances in which all that mattered was tacit knowledge, pure and simple. It is now seen that this fails to allow for the possibility that explicit references to codified sources of "authority" may be supplanted by the formation of "common knowledge" regarding the subscription of the epistemic community to that "authoritative" body of information.

(1b) One may see from Table 1 where the Collins-Latour-Callon boundary lies and the latter's positioning in respect to the regeneration of tacit knowledge; this makes it clear that such a process of regenerating tacit knowledge— in the course of frequent applications of formal algorithms and other codified material— clearly rests upon the pre-existing establishment of a well-articulated body of codified, disciplinary tools.

(1c) Economists' recent concerns with the economics of knowledge tend to lie in the 'no disagreements' (uncodified, manifest) box in Figure 2. We are talking here about the literature on 'sticky data' or 'local jargons' as methods of appropriation. This is the location that creates the 'quasi' aspect of knowledge as a quasi-public good, and thus exploration of the extent to which a type of knowledge is indeed only a quasi-public good will demand exploration of both the degree of codification and the degree of manifestation.

(1d) The concept of production of knowledge in Mode 2, advanced by Gibbons et al. (1996) as an emergent trend of the late twentieth century, has previously been discussed in terms that draw a contrast with the antecedent dominant organizational features of scientific inquiry associated with Mode 1, namely that the new form was more dependent upon tacit rather than codified knowledge, and typically was trans-disciplinary rather than disciplinary in the lines along which the recruitment of membership, as well as professional evaluations and rewards were structured. The conceptualization of Mode 2 as a self-sufficient successor that is now displacing an earlier, highly institutionalized system of research and innovation system has been disputed on historical and analytical grounds (see the discussion and further references in David, Foray and Steinmueller (1998), for example). But the additional point to notice in the present context is that such coherence and functionality as groups working in Mode 2 have been able to obtain would appear to rest upon their development of procedural authority to which the fluid membership subscribes.

6.2 On interactions with external phenomena

(2a) How do changes in information and communications technologies (ICT)

impinge upon the distribution of knowledge production and distribution activities within the re-mapped space described by our topography? The first and most obvious thing to notice is the endogeneity of the boundaries (which we discuss more fully in the next section). In the new taxonomy there are two interesting distinctions— knowledge activities may use and produce codified or uncoded knowledge; or they may use and produce knowledge that is either manifest or latent. We should re-state here that 'boundary' here is not being used in reference to the distribution of the world knowledge stock, but, instead to the prevailing locus of the activities of knowledge agents in a specific cognitive, temporal and social milieu. Nevertheless, it is most likely to be true that the location of a group's knowledge stock will be intimately related to, and possibly even coterminous with the location of its knowledge production activities.

(2b) Organizational goals affect the Manifest/latent boundary. Activities that couple teaching with research, for example, will be pushed towards the more fully "manifest" region of the statespace. This consideration will be important in studies of the economics of science, and one particular example will be in studies of the economics of university funding, and how different funding structures affect the types of activity within the university system. (See Geuna, 1998.)

(2c) The positioning of the endogenously determined boundary separating the codified from the uncoded states of knowledge will be governed by the following three setoff forces, which we shall examine at greater length, below. For the present it is sufficient simply to note that these include: (i) costs and benefits of the activity of codification; (ii) the costs and benefits of the use of the codified knowledge (data compression, transmission, storage, retrieval, management...); (iii) feedbacks that arise because of the way codified knowledge is used to generate further codified knowledge.

(2d) The age of the discipline (stage in the life cycle) affects both boundaries. The evolution of a discipline, a technological domain (or of a research group or a community of practitioners) must now be described as a movement in the two-dimensional plane of the tableau in Figure 2. Early in the life-cycle of a knowledge generation endeavor, we begin in the south-east corner of figure 2, as a disparate group of individuals with no common authority, highly tacit knowledge, and a very restricted scope for its transmission beyond the confines of the work-group. Subsequently the locus of the knowledge-activities moves westward, as agents make their discoveries and inventions manifest either in physical artifacts or in published papers, etc. Though scholarly texts may be produced, because the language in which these reports are couched has not yet been standardized, codification must still be considered very incomplete. Nonetheless, this is where the codification process begins. Thence the movement follows a northerly course, spreading towards the north-east as disputes arise from

inconsistencies in description and interpretation, and conflicts emerge over the way the language will be standardized. As they are resolved, the system takes an eastward course, landing up in a situation where most of the actual research activity is carried on within the "latent-codified" and "manifest-partially codified" domains.

7. The economic determinants of codification

Up until this point attention has been focused on the conceptual distinctions separating types of knowledge activities, and to describing the locations of knowledge activities in the space thus delineated. In any topographic discussion there is a temptation to treat boundaries between regions as having been imposed from outside the system being examined. In a sense this is proper, in that structures are often in principle distinct from the activities that make them. But inasmuch as we are dealing here with knowledge, and the latter is seen today to be so central to the process of economic growth, a treatment of the subject would not be useful were it to omit the question of the boundaries' genesis, and the forces that determine their positions. This becomes all the more relevant because our concern here is not primarily taxonomic; we are less interested in delineating the nature and varieties of human knowledge than in being able to explain and predict the changes taking place in the character of economically significant knowledge activities.

Another way of highlighting this issue is to return briefly to the previous discussion of the critique of the implicit assumptions of new growth theory regarding the composition of the knowledge stock. Both camps in this incipient debate implicitly have been regarding the "composition of knowledge by type," i.e. the codified-tacit mix, as exogenous— a matter for epistemologists, cognitive scientists and students of human psychologists. In other words, it seems to be something that is essentially pre-determined by non-economic conditions affecting the supply of knowledge. But, this approach mistakenly ignores the influence of the range of considerations impinging upon the *demand* for codified vs. uncoded knowledge. Some of these factors involve institutional arrangements affecting the structure of relative rewards for codification activities, whereas others have to do with the state of available technologies affecting the costs of rendering knowledge in codified form, and the storage, retrieval and transmission of information.

7.1 The endogeneity of the tacit/codification boundary

Any agent makes a decision about what kind of knowledge activity to pursue. Should the output be codified or uncoded? Are the inputs manifest or latent? For an economist, there is a simple one-line answer: It depends on the costs and benefits. This implies that where knowledge activities are

located (the extent to which agents codify their knowledge for example), will depend on economic considerations, and the boundaries may move in response to changes that are external to the knowledge system per se. To this we now turn.

The endogeneity of the tacit-codified boundary (or the Merton-Kuhn frontier of figure 1a) refers to the fact that the agents pursuing a knowledge activity have a choice regarding whether or not to codify the knowledge they use and produce. In analyzing the economics of this choice, we need— even more so than above— to consider only knowledge which is codifiable, the issue being to analyze the rationale for codifying knowledge. Different situations arise: knowledge can be in a state of true tacitness but codifiable; the codebook can exist or not; it can be displaced or not. Each situation will generate its own cost-benefit problems, which we address through the concept of the knowledge environment.

In practice, the extent to which knowledge is codified is determined by incentives: the costs and benefits of doing so. For example, many factors— such as, to take the simplest argument, the high cost of codifying a certain type of knowledge— can decrease the incentives to go further, by lowering the private rate of return on codification. This low rate of return can, in turn, induce the maintenance of a large community of people possessing the tacit knowledge. In this case, there will be a labor market that can be used to store and transfer the knowledge from firm to firm. Of course, the presence of a thick labor market as a way of transferring knowledge further reduces incentives to codify. (See the policy discussion above.)

A self-reinforcing process of this kind can generate multiple equilibria. If, for example, there are high returns to codification, more knowledge will be codified. This will decrease the value of a thick labor market as a means of maintaining and distributing (tacit) knowledge. As the labor market shrinks, the relative value of codification increases further. Thus there are two possible equilibria: one with significant resources devoted to codification and a resulting high incentive to codify; and one with few resources so devoted, a thick, active market for skilled labor as the mechanism for storing and disseminating knowledge, and thus low incentives to codify. This argument rests on there being substitutability in the production process between the types of knowledge transferred by these two mechanisms.

This focus on endogenous limitations indicates that costs and benefits and the resulting incentive structures are pivotal in shaping the dynamics of codification. Emphasizing the role of the incentive structures by no means implies that the codification of new forms of knowledge is an instantaneous process: moving the boundaries between codified and tacit parts of the stock of knowledge is a matter of long-term technological and institutional

evolution; involving changes in incentive structures, and in costs and benefits.

7.2 Costs, benefits and the knowledge environment

In order to understand the sources and magnitudes of costs and benefits, it is necessary to put them in the context of the knowledge environment. A first and straightforward point is that the incentives will depend to a very great extent on the possibility of proceeding to codification on the basis of pre-existing codebooks (languages, models and techniques, according to the framework of Cowan and Foray (1997)).

When the language and the model already exist, the fixed costs, those born to generate the now standard models and languages, have already been sunk in situations of this sort: languages and models have been developed by past work, and are known by codifiers and users. Such a situation describes both cases in which codebooks are manifest and those in which codebooks are displaced. The idea here is that some existing body of well-developed, stable, codified knowledge, often one that is displaced, contains the necessary concepts and relations with which to codify the knowledge in question. The only cost then is the variable one. On the other hand, if codebooks do not exist, or are incomplete or ambiguous, costs of codification entail more than simply the variable costs. Further, before a language has been standardized and is stable, linguistic ambiguity implies that codes which appear to represent codified knowledge can change their meanings as the language is developed and refined, and as vocabulary expands and changes. It is, thus, useful to differentiate between contexts of stability and contexts of change.

7.3 Costs and benefits in a stable context

In a stable context— when there is a community of agents who have made the necessary initial investments to develop a language and to maintain efficient procedures of language acquisition for new entrants— the transfer of messages can be assimilated to transfer of knowledge, and storing messages means recording knowledge.

On the benefit side, the efficiency of codification will be greater in very large systems having specific requirements regarding co-ordination among agents. We identify five classes of situations: (i) systems involving many agents and many locations; (ii) systems strongly based on re-combination and re-use and which take advantage of the cumulateness of existing knowledge (rather than on independent innovation); (iii) systems that require recourse to detailed memory; (iv) systems which need particular kinds of description of what (and how) the agents do; and lastly, (v)

systems characterized by an intensive usage of information technologies. We take these up *ad seriatim*.

First, codification will provide high benefits in stable systems characterized by specific requirements of knowledge transfer and communication. This kind of requirement can result from a tendency towards de-localization and externalization or from the development of co-operative research, entailing a spatial distribution with activity at many places. This first effect can be appreciated without any ambiguity, for example in science. It operates, however, within a given "clique" or network— that is a community which shares common codes and the tacit knowledge to interpret them.

Second, in (stable) systems of innovation where advances and novelties mainly proceed from re-combination, re-use and cumulativeness, benefits of codification are important. Gibbs (1994) claims that the very limited progress in the productivity of software engineering is due to an excessive dependence on craft-like skills (in contrast for example with chemical engineering). The schema that Gibbs has in mind is that once an algorithm is written as a piece of code, it can be used in many applications— at least in principle. The practical difficulty in doing so arises in part because of a lack of standardization both in the way code is written and the way algorithms are employed. This lack of technological rationalization impedes the full realization of the opportunities provided by the re-use and re-combination model.

Third, systems that require extensive memory and retrieval capacities (e.g. firms with long development cycles, high rates of turn-over, some particular demographic problems or institutions confronted by a big technological bifurcation) will greatly benefit from codification. In those systems, too little codification increases the risk of "accidental uninvention". MacKenzie and Spinardi (1995) showed, for example, that specific local and tacit knowledge was so important in the nuclear weapons design process that there was always a risk of losing critical elements of the knowledge base simply due to the turnover of scientists and engineers— a risk of technological retrogression, or at best costly reconstruction of previous tacit capabilities.

The same argument is readily extended to cover situations in which knowledge has been thoroughly codified in the form of algorithms, or operating instructions, but the text of the "source code" for these— or an understanding of the language in which it was recorded— has ceased to be readily decipherable, or has simply been misplaced, or destroyed. The result is a paradoxical one: the technology in which the knowledge has been embedded may continue to work, as is the case when the computer implements the machine-language version of its instructions. But, as has been found to be the case with some major pieces of "legacy software", the human agents

who can no longer read or write the source code are unable to emend or elaborate those machine-language encoded instructions, or repair defects in the original source code that have become evident. It is possible that even beyond the range of such algorithmic technologies, cultural inventions and culturally transmitted skills important for activities upon which social welfare depends— such as those involved in dispute resolution— may become lost because “ the market ” for agents possessing tacit knowledge of that kind is undermined by the competition of more fully codified (legal) procedures.

Fourth, systems, which need accurate descriptions of what agents are doing (either to meet quality standards constraints, or to patent innovations, or to enter into contractual relations with a partner), will benefit greatly from codification. Here we can also include systems confronted with inefficient market transactions, where the traditional mechanisms of legal warranty, insurance, reputation and test are not efficient means to mitigate the effects of information asymmetry (Gunby, 1996). Recording production practices, which is a form of codification based on recent language innovation (in the form of creating standards for record keeping) is aimed at reducing these asymmetries.

Fifth, and last but not least, a sort of cross-situation deals with the lack of productivity gains from the use of ITs, due to incomplete codification. Fully taking advantage of the potential productivity gains of IT typically demands not only the adoption of the technology but also organizational change (see Cowan, 1995.) But a firm undergoing organizational change does not want to lose functionality in the process. The firm must develop jointly the new technology and organizational structures that will reproduce old functions and create new ones (see David 1991, 1994). It is obvious that if too much of the old functionality resides in tacit knowledge, or depends heavily on it, this task will be extremely difficult. When the presence of tacit knowledge operates as a bottleneck, impeding the full realization of productivity potential, the firm can expect great benefits from codification (Baumol et al., 1989). This, indeed, may be the role of management consultants, referred to earlier.

In all these cases, where important operations of transfer, re-combination, description, memorization and adaptation of existing knowledge (to ITs) are required, it would be very costly and inefficient to keep this knowledge tacit. Thus, there can be under-investment in codification, co-existing with “ excess of tacitness.” Given the nature, degree and pace of recent technical change, it is likely that the current equilibrium involves an allocation of resources devoted to knowledge generation and transmission under conditions of tacitness which is neither socially nor privately efficient, because knowledge that should be made more widely available in codified form remains uncoded, whereas a lot of resources are poured

into the production of differentiated "information" that is idiosyncratically coded in the hope of being able to capture private rents.

Now, the way through which companies or other entities can manage a process of codification in a situation in which a large portion of the knowledge base is tacit raises many issues. Thus occasionally (in times of stress) the organization seeks the aid of some external management consultant to identify what things it really "knows how" to do. A great deal of management literature has by now been spun around that conceptualization, so it may be found reassuring to notice an implication of our topographic structure, namely, that collective procedural knowledge may be unarticulated, but is perfectly (or at least workably) "codifiable."

A more interesting issue for the skeptical economist to ponder in that connection is simply this: why is it that the organization—having somehow acquired and successfully deployed its "core capabilities" without needing to make them explicit—now requires, or finds it profitable to employ the costly services of outside management consultants to break the spell of tacitness. In most of the specific cases discussed in management literatures this question is not posed. But, there is a suggestion that the organization has either forgotten what it once knew and acted upon collectively, or that the situation in which it operates has been altered without prompting a timely revision of the collective awareness of the mismatch created between opportunities and capabilities to exploit them. The presumption, therefore, is that it will take too long, or be too risky, to go through a tacit, trial and error learning process. Bringing explicit analysis to bear—and so codifying the organization's understanding of itself and its situation—will prove either more expedient or less costly, or both, than continuing in tacit mode.¹²

7.4 Costs and benefits in the context of change

While many knowledge activities take place in a relatively stable context, some particular domains or sectors are characterized by acknowledge environment exhibiting ongoing rapid transformations.

Models and languages are fluid, and the community of agents conversant with the models and languages is itself changing. The fluidity of the language implies that there exists a certain amount of uncertainty about what the messages actually mean because there is uncertainty, and perhaps change, with regard to the vocabulary in which they are written. Even when scientific papers express new discoveries, or re-examine old results in some "natural" language, much jargon specific to the subject matter

¹²Reference may be made here to the work of Cobbenhagen (1999).

remains; 'terms of art' are employed whose meanings are lost on outsiders; and, in formal modeling, definitions of variables specific to the model may remain in flux as the model itself is modified and reconciled with observational data. In an important sense, the progress of research involves— and requires— the stabilization of meanings, which is part of the social process through which the stabilization of beliefs about the reliability of knowledge comes about.

To the extent that codification is taking place under those conditions, the benefits from it have substantial "spillover" elements, as they contribute largely to the modeling and language development parts of the exercise. There may be competition among different basic models, and so among the basic tenets and vocabulary of the language. Until this competition is resolved the community of potential knowledge generators and users will have difficulty communicating, and the value of knowledge codification that arises from dissemination will be reduced. Thus the codification process in this environment generates some immediate value (which derives both from worth of the content of the messages that agents can transmit and interpret with less effort and expense, and from the value to the agent of storage and retrieval of his own knowledge), but it has greater value as an investment good— a contribution to the resolution of the competition among variant languages and models.

It is in the context of change that we expect to find situations of "excess codification." That is to say, the accumulation of successive generation of codes can prevent the development of radically new knowledge, simply because explicating and understanding it would require entirely new codes. As argued by Arrow (1974, p.56) codification entails organizational rigidity and uniformity while increasing communication and transaction efficiency: "the need for codes mutually understandable within an organization imposes a uniformity requirement on the behavior of participants. They are specialized in the information capable of being transmitted by the codes, so that they learn more in the direction of their activity and become less efficient in acquiring and transmitting information not easily fitted into the code."

It is clear, therefore, that codification can have unfortunate consequences for creativity and radical changes. Like a larger category of co-ordination mechanisms, codified knowledge can be a potent "carrier of history" — encapsulating influences of essentially transient and possibly extraneous natures that were present in the circumstances prevailing when particular codes took shape; and, having that power, it can become a source of "lock in" to obsolete conceptual schemes, and to technological and organizational systems that are built around those.

The second problem we have thus identified deals with "excess inertia" . There are high fixed costs to be borne in the process of codification,

especially when the cognitive environment is changing. Roughly put, costs of learning and developing languages in which new codes are being written will be incurred during the period when the knowledge environment is in flux, whereas benefits will accrue (from some of those investment) during a subsequent period of stabilization and widespread dissemination of the information. During a period of change, infrastructure is developed, languages and models are built, learned and standardized, and a community of agents with shared tacit knowledge grows. All of these investments contribute to a reduction in the fluidity of the knowledge environment, and conduce to hastening the enjoyment of the increasing returns from more widespread application that are permitted by the stabilization of organizational and technological knowledge. As a network of users of the knowledge expands, learning costs continue to decline and co-ordination externalities are likely to grow more significant as a source of social benefits.

But if developing new languages and models allocates the fixed cost to one generation while many future generations benefit from the new infrastructure to codify knowledge, there is an inter-generational externality problem which can result in a lack of adequate private (or social) incentives for allocating resources to the development of more powerful codes, and systematizing those that already exist. Solutions that would help mitigate this kind of time inconsistency problem entail the development of relevant markets (which may significantly increase the benefits even for the first generation of developers), or the creation of infinitely-lived institutions that do not discount the future so strongly, or with the cultivation of altruistic preferences for the welfare of coming generations to which a greater stock of useable knowledge can be bequeathed (see Konrad and Thum, 1993).

8. Conclusions and the direction of further work

This paper has looked intensively and critically at one of the several dimensions David and Foray (1996) identified that in their schematic description of the space in which "knowledge-products" were distributed. Our focus has been maintained on the most problematic and, for many economists, the most esoteric of the three axes defining that space: the dimension along which codification appeared at one extremum and tacitness occupied the other. This has permitted some further unpacking of the economic determinants of codification decisions, and the resources committed thereto, and it has revealed that the term tacit is being used so loosely in the current economics of science and technology literature that important distinctions, such as the one separating that which is uncoded in a particular context and that which will not (likely) be codified at all, are blurred, or entirely lost. Also lost from view in many modern discussions of codification is the important difference between

codification of procedural knowledge, as distinct from factual propositions. By bringing these into greater prominence, a number of puzzles and inconsistencies which detract from the logic of the analysis of the social return to investment in codification can be resolved.

REFERENCES

- Anderson, Esben Sloth (1994). *Evolutionary Economics: Post-Schumpeterian Contributions*. London: Pinter Publishers.
- Arrow, Kenneth J. (1955). "Economic aspects of military research and development," RAND Corporation Memorandum D-3142, August 30. (Unpublished document).
- Arrow, Kenneth J. (1962). In *The Rate and Direction of Technical Change*, R. Nelson. ed., New York: NBER.
- Arrow, Kenneth J. (1974). *The Limits of Organization*, Norton.
- Baumol, W., Batey Blackman, S. and Wolff, E. (1989), *Productivity and American leadership: the long view*, Cambridge: MIT Press.
- Callon, M. (1995). "Four models of the dynamics of science," in S. Jasanoff (ed.) *Handbook of Science and Technology Studies* Thousand Oaks, CA.: Sage.
- Cobbenhagen (1999) "Managing Innovation at the Company Level: A study on non-sector-specific success factors" PhD dissertation (to be defended), Maastricht University.
- Collins, H. M. (1974) "The TEA Set: Tacit Knowledge in Scientific Networks," *Science Studies*, 4: pp. 165-186.
- Cowan and Foray (1997). "The Economics of Codification and the Diffusion of Knowledge" , *Industrial and Corporate Change* vol. 6(3), pp. 595-622.
- Cowan, R. (1995). "The informatization of government as an economic opportunity," *STI Review*, no.16.
- David, P. A. (1991) "General purpose engines, investment and productivity," in *Technology and Investment*, eds., E. Deiaco, E. Hrnell and G. Vickery. London: Pinter Publishing.
- David, P. A. (1994) "Why are institutions the 'carriers of history'? : Path dependence and the evolution of conventions, organizations and institutions," in *Economic Dynamics and Structural Change*, 5(2): pp. 205-220.
- David, P. A. (1997). "From market magic to Calypso science policy: A review of T. Kealey's *Economic Laws of Scientific Research*," in *Research Policy* 26, May: pp. 229-255.
- David, P. A. and D. Foray (1995): "Accessing and Expanding the Science and Technology Knowledge Base," *STI*, no.16: pp. 13-68.
- David, P.A., D. Foray and W.E. Steinmueller (1998). "The Research Network and the New Economics of Science: From Metaphors to Organizational Behavior" , in A. Gambardella and F. Malerba (eds.) *The Organization of Inventive Activity in Europe*. Cambridge: Cambridge University Press.
- Dosi, Giovanni (1988): "Sources, Procedures and Microeconomic Effects of Innovation," *Journal of Economic Literature* 26: pp.1120-1171.
- Dosi, G. and M. Egidi (1991). "Substantive and Procedural Uncertainty," *Journal of Evolutionary Economics*, 1(1): pp. 145-168.
- Dosi, G., C. Freeman, R. Nelson, G. Silverberg and L. Soete , eds. (1988). *Technical Change and Economic Theory*. London: Pinter Publishers.

- Ferguson, Eugene S. (1993). *Engineering and the Mind's Eye*. Cambridge, MA: MIT Press.
- Geuna, Aldo (1998). *Resource Allocation and Knowledge Production: Studies in the Economics of University Research*. MERIT, Maastricht University PhD. Dissertation [forthcoming from E. Elgar Publishers].
- Gibbons, Michael et al. (1996). *The New Production of Knowledge*. London: Sage Publications.
- Gibbs, Q. (1994). "The crisis in software", *Scientific American*.
- Gunby, P. (1996). Explaining adoption patterns of process standards, The University of Western Ontario, Dept. of Economics, PhD Dissertation.
- Hofstadter, Douglas (1979). *Gödel, Escher, Bach: An Eternal Golden Braid*. New York: Basic Books.
- Kay, John (1999), "Money from Knowledge," *Science and Public Affairs*, April: pp. 12-13.
- Kealey, Terence (1996). *Economic Laws of Scientific Research*, London: Macmillan.
- Kuhn, Thomas (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Machlup, Fritz (1980). *Knowledge and Knowledge Production (Vol. I of Knowledge: Its Creation, Distribution and Economic Significance*. Princeton, N.J.: Princeton University Press.
- MacKenzie, D. and G. Spinardi, (1995). "Tacit knowledge, weapons design and the uninvention of nuclear weapons" , *American Journal of Sociology* , vol.101, no.1.
- Nelson, Richard R. (1959), "The Simple Economics of Basic Scientific Research" , *Journal of Political Economy*, vol. 67, pp. 323-348.
- Nelson, Richard R. (1992). "What Is 'Commercial' and What is 'Public' about Technology and What Should Be?" , in N. Rosenberg, R. Landau and D. Mowery (eds.), *Technology and the Wealth of Nations*, Stanford University Press.
- Nelson, Richard R. and Sidney Winter (1982). *An Evolutionary Theory of Economic Change*, Cambridge, MA: Harvard University Press.
- Newell, A., and H.A. Simon (1972). *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Patel, Pari and Keith Pavitt (1995). "Patterns of Technological Activities: Their Measurement and Interpretations" , in P. Stoneman (ed.), *Handbook of the Economics of Innovation and Technical Change*. Oxford: Blackwell.
- Pavitt, Keith (1987), "The objectives of technology policy," *Science and Public Policy*, 14: pp. 182-188.
- Penrose, Roger, and Martin Gardner (1989). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford : Oxford University Press.
- Penrose, Roger. (1997). *The Large, the Small and the Human Mind*. with Abner Shimony, Nancy Cartwright, and Stephen Hawking; edited by Malcolm Longair. Cambridge; New York: Cambridge University Press, 1997.
- Polanyi, Michael (1967). *The Tacit Dimension*. New York: Doubleday.

- Polanyi, Michael (1958). *Personal Knowledge: Towards a Post-Critical Philosophy*. London: Routledge & Kegan Paul.
- Schutz, Alfred and Thomas Luckmann (1973). *The Structures of the Life-World* [Translated by Richard M. Zaner and A. Tristram Engelhardt, Jr.], Evanston, IL:Northwestern University Press.
- von Hippel, Eric (1991). "The Impact of 'Sticky Information' on Innovation and Problem-Solving," Sloan School of Management, MIT, Working Paper BPS 33147-90 (Revised), April.
- von Hippel, Eric (1993). "Trading in Trade Secrets," *Harvard Business Review*.
- von Hippel, Eric (1994). "'Sticky Information' and the Locus of Problem Solving: Implications for Innovation," *Management Science*, 40(4), April: pp. 42-439.