1

This chapter describes a multi-level framework for the evaluation of research within a system of innovation, recognizing differences due to the scope and radicalness of research projects. Based on theories about the innovation process, a set of indicators is proposed that can guide investments in research and identify bottlenecks for policy makers to address.

A Theories-Based Systemic Framework for Evaluating Diverse Portfolios of Scientific Work, Part One: Micro and Meso Indicators

Gretchen Jordan, Jerald Hage, Jonathon Mote

Introduction

Recently, several articles (Arnold, 2004; Molas-Gallart & Davies, 2006) have appeared arguing for changes in the kinds of evaluations that are being made of research, technology, and development (RTD) programs. Among other suggestions, the following seem to be of special merit: (1) a more macro and systemic focus; (2) greater concentration on the processes of generating innovation; (3) using theory to guide the RTD evaluation; and (4) the identification of blockages and obstacles or what Arnold (2004) labels "failures." One of the great advantages for evaluators of concentrating on innovation processes is that it helps identify the causal chain connecting policy intervention and outcome including eventual societal impact, something that Molas-Gallart and Davies (2006) identify as lacking in typical medium and long-term evaluations. To this list of suggested changes, we would add that RTD evaluations need to investigate key indicators to provide data that policy makers need in order to know what policy *re*formulations should be made and how.

These recent concerns about the structure of evaluations stem from a number of challenges that governments face today. In particular, we would argue that three primary factors—the rising level of RTD expenditure in real terms, the importance of innovation for both economic and non-economic goals, and the increasing speed of development of innovative solutions—significantly increase the need for evaluations that better guide government policy formation and *re*formulation.

The objective of this chapter is to outline a theories-based Innovation Systems Framework (ISF) of indicators for RTD evaluations that can aid government policy makers in policy formulation and *re*formulation. The indicators that are proposed suggest protocols for performance monitoring and evaluation, and could form the basis of a new kind of data structure for science reporting agencies such as the National Science Foundation (NSF). Although the ISF we have developed is multi-level in nature, this chapter focuses on the micro- and meso-analytical levels. We address the issue of macro-indicators and their relationship to the micro and meso elsewhere (Jordan, Hage, & Mote, 2007). The systemic framework we suggest has significant potential for developing socio-econometric models that incorporate the innovation processes necessary for predicting innovation outcomes (or throughput), a request that Dr. Marburger (2006), the Director of the United States Office of Science and Technology, made recently.

In the first section of the chapter, we provide a brief overview of the theories-based systemic framework of indicators for innovation. Central to this is the idea of innovation network theory (Hage & Hollingsworth, 2000) that describes the innovation processes and identifies a number of potential blockages in the connectedness of innovation networks in technology sectors or regimes. The theory argues that there are six research arenas in the process of innovation (basic, applied, development, manufacturing, quality, and commercialization research) and robust linkages among these arenas are critical for continued innovation. Next, we articulate the critical role of the meso-level as the primary focus of analysis. In our framework, the meso-level connects to the micro-level which encompasses Jordan's (2006) theory of profiles and previous work on industrial innovation (Hage, 1999), and provides a larger context for discussing potential organizational obstacles to innovation. In addition, the meso-level provides focus when connecting to the macro-level, encompassing the various institutional theories in the national system of innovation literature (Hall & Soskice, 2001; Nelson, 1993) and the new work on institutional change (Campbell, 2004), as well as allowing for a discussion of obstacles to innovation created by various institutional rules.

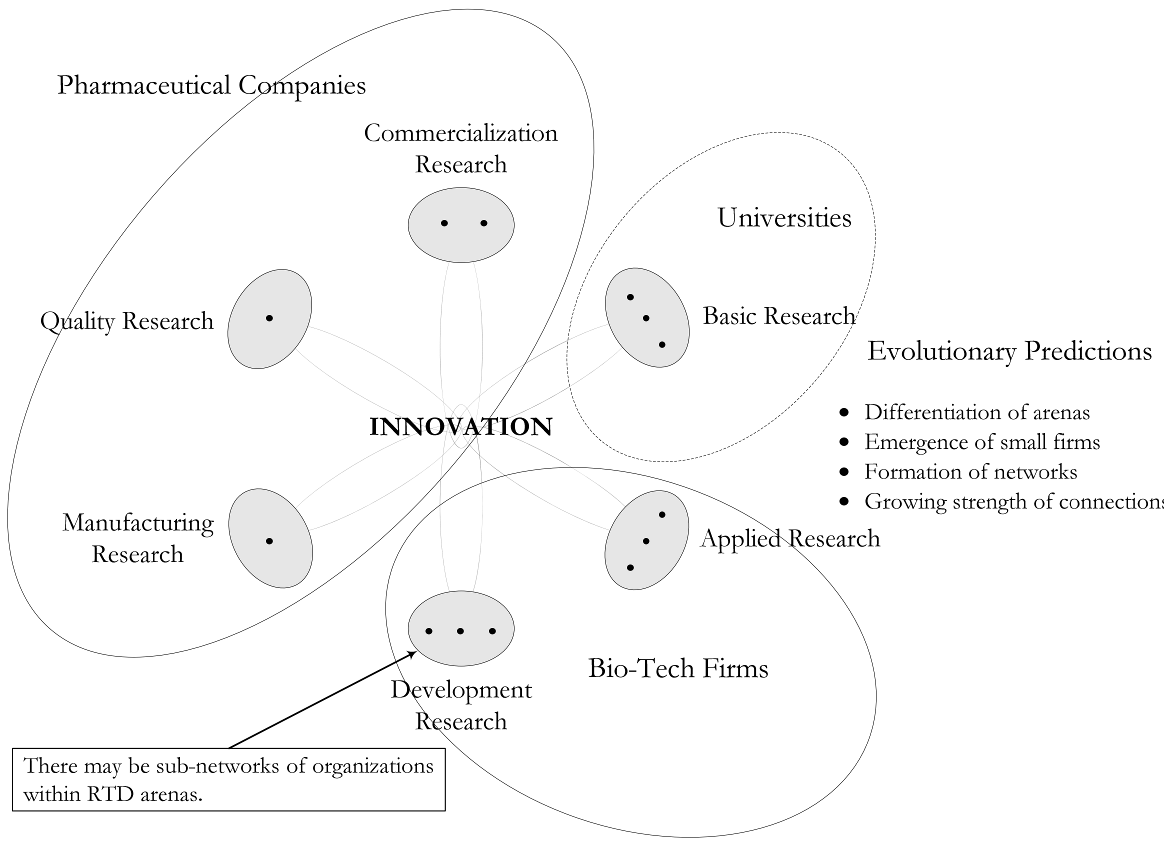
Together, the three levels answer a plea for a theory of knowledge production that contains these three analytical levels (Hage & Meeus, 2006) and provide the opportunity for contributing to other theories and frameworks such as organizational learning, knowledge communities, and for putting the throughput or black box of RTD into standard econometric input-output evaluation models.

Within each analytical level, we identify three sets of indicators that provide guidance for policy makers, as well as indicate specific possible blockages and obstacles. In the second and third sections of this chapter, we examine the selection of the indicators at the micro-, and meso-analytical levels respectively along with a discussion of the theories involved and some of the implications for the "science" of RTD evaluation. In general, micro-indicators focus on how to allocate funds using the criteria of balanced investments (public/private) across the six RTD arenas in a technological sector, across the portfolio of investments within each arena, and across selected research organizations with the appropriate organizational profiles for the portfolio choices. Similarly, in general meso-level indicators measure the outputs of each arena in real time, the strength of the connectedness between differentiated arenas, and the overall assessment of innovation performance including societal impact. We conclude with a summary of how this innovation systems framework can be used in RTD evaluation.

A Theories-Led Innovation Systems Framework for RTD Evaluation

As Weiss (1997) discusses, a theory-based evaluation allows one to address the most relevant mechanisms that mediate between processes and outcomes and better understand how programs work. A particularly useful way of capturing the complexity of the innovation process from scientific advance to socio-economic outcomes is contained in the idea innovation network theory of Hage and Hollingsworth (2000). The idea innovation network starts with a simple idea, namely that in many technological sectors, especially the more high tech ones (e.g., automobiles, pharmaceuticals, semiconductors, telecommunications) commercially viable product innovations necessitate research in the six arenas diagrammed in Figure 1. The theory builds upon the conceptual nonlinear model of Kline and Rosenberg (1986) but alters the focus to arenas within a technological sector, and adds the concept of quality research to the original five areas.

Figure 1. An Example of the Evolutionary Predictions in the Idea Innovation Network Theory: Bio-tech/Pharmaceutical Sector



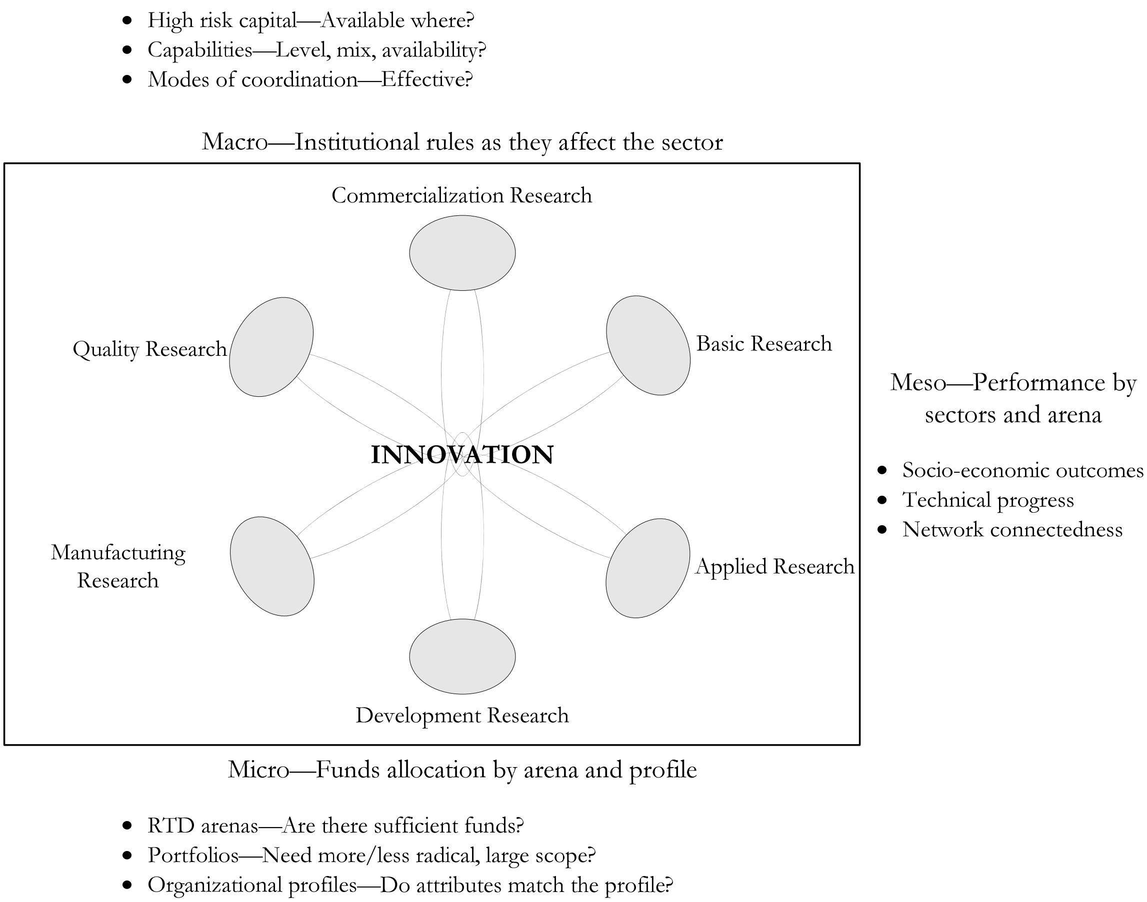
Why be concerned about all six arenas in the innovation process? In general, the process of generating scientific advance and innovation in products or processes that have socio-economic impact can be thought of as "throughput." The idea innovation network theory provides a necessarily complex view of the throughput of the innovation process and does so at the meso-technological sector level where indicators and socio-econometric models of innovative performance can be constructed. Another advantage of this theory is that it can expose research "gaps," an important issue in science policy today (Marburger, 2006). Some of these arenas have been ignored, with detrimental consequences for the competitive position of certain countries. For example, the American automobile industry continues to lose market share because of perceived poor quality of its products. This includes quality control in the sense of reduced defects and lower operating costs, but also research that reduces the various externalities of the products such as energy consumption, global warming, and health risks. Manufacturing research is also an important if frequently neglected arena. The importance stems from three considerations: (1) reducing costs and, potentially, the export of jobs to developing countries; (2) increasing customization so that multiple products can be produced on the same assembly-line; and (3) reducing externalities, as mentioned above.

Consistent with Kline and Rosenberg (1986), a good idea for an innovative product or service can start in any one of these six arenas. The process of innovation is nonlinear, with ideas moving back and forth between arenas multiple times, hence the use of the word "network" rather than chain. The real meaning of this term "network" is the argument that each of these arenas must be connected, and, if one of the arenas has a radical advance in knowledge, it must be strongly connected to the other arenas for the transfer of the tacit knowledge involved in the radical knowledge advance. Connecting research organizations in the same arena together in sub-networks to produce a radical advance in knowledge, as found in the concept of knowledge communities or research consortia, is also an important consideration. Strong connections are defined as face-to-face interactions with frequent meetings. Without these strong connections, then the radical advance in knowledge is not likely to be exploited in a timely fashion. Support for this hypothesis is found in the research on knowledge communities in basic research (Mohrman, Gailbraith, & Monge, 2005), on technical pools in the European Union (EU) study of the idea innovation network theory (van Waarden & Oosterwijk, 2006) and especially the work on the research consortium SEMATECH (Browning, Beyer, & Janis, 1995), which combined the semiconductor companies in the United States.

Analytical Focus at the Meso Level. The analytical focus of the ISF is the meso-level of the technological sector, which others have also identified as critical (Archibugi & Pianta, 1992; Guerrieri & Tylecote, 1997; Malerba & Orsenigo, 1993, 1997; Pavitt, 1984). This meso-level of arenas and networks is ideal for allowing one to connect to the macro-institutional level of the national system of innovation and to the micro-research organizations that are its constituent parts, thus also providing the often overlooked linkage between the micro- and macro-levels. This is important because the average scale of the research projects often differs, and the rates of technological change, and the pace of product innovation, typically vary from one sector to another. For example, some sectors, such as semiconductors, have radical breakthroughs in the performance characteristics of chips every 18 months. In contrast, some sectors, such as pharmaceuticals have a much slower pace, despite larger levels of RTD investment. Finally, policy makers are usually interested in intervening at the technological sector level to achieve their various goals. Hence, the meso-level provides an advantageous starting point for the development of a set of key indicators for an evaluation framework rather than the national system of innovation as the analytical focus. Further, since the meso-level in this framework is connected to the macro-level, it does not mean that one is ignoring this aspect of the innovation process.

An Overview of the Scheme of Indicators. A schematic of the general argument in the framework of indicators is provided in Figure 2. At each level—the micro, meso, and the macro—we have chosen three sets of indicators for evaluators to measure. The three sets of indicators at each level allow for a better appreciation of how policies impact on particular arenas of the idea innovation network as well as their connections, and thus can lead to more fine-tuning of these policies. With this information, the policy maker can assess what they have accomplished so far and then decide where they want to be in the next five to ten years. Best of all, evaluations with these sets of indicators at each analytical level are more likely to allow for the identification of the obstacles or bottlenecks that Arnold (2004) advocates as one of the main thrusts of systemic evaluation. Indeed, interventions may be essential if the mission goals of the policy makers are to be achieved and thus this aspect of the framework, clearly identifying organizational, network, and institutional failures, is critical. With this perspective, the evaluator can shift from simply measuring "what is done" or accomplished and move on to "what could be," a more important goal for evaluators.

Figure 2. The Innovation Systems Framework for RTD Evaluation



The Micro-Level of Indicators: Balanced Investments in Arenas, Portfolios, and Organizational Profiles

At the micro-level, evaluators need to be concerned about three key sets of indicators in order to assess and provide good policy guidelines for interventions. These three sets of indicators focus on aspects of the first and most common form of government intervention, namely the allocation of research funds. Usually when governments decide to achieve some mission such as national security or to become more competitive, they begin by increasing RTD spending. Before governments allocate more money, there should be an assessment of how the money (both public and private) is presently being spent and whether or not there are research "gaps."

The idea innovation network immediately focuses the evaluator's attention on the amount of investment in *each* of the six arenas and helps locate potential gaps blocking achievement of the desired throughput. Recognizing the complexity of the innovation processes also alters how one regards the disbursement of RTD funds to portfolio choices, classified by the degree of radicalness and the scope of focus. Each strategic choice of incremental versus radical, narrow scope versus broad scope must be evaluated in each of the six arenas. By the same logic, Jordan's (2006) theory of Research Profiles also results in a reconsideration of how funds are allocated to projects and research organizations to have the desired fit between strategy and organization profile. Furthermore, one can begin to assess whether organizational blockages are part of the reason for lack of technical progress in the strategic choices that the government has made.

Balanced Investments Across Six Arenas. We must begin with the objectives of the government policy makers, with the critical issues being the desired aims (innovation) and the identification of the most appropriate technological sector(s) to pursue these innovations. For example, the EU has recently decided to emphasize the technological sectors of health, pharmaceuticals, energy, environment, security, electronic equipment, and transport and logistics (Georghiou, 2006). Given the choice of these sectors for stimulating innovation, the policy maker or evaluator needs to determine the amount of RTD spending in each of the six arenas, including both public and private expenditures, in a sector. One issue for the assessment to determine is whether one or two of these arenas are being ignored or under-funded. As we have already observed, two likely candidates in the current United States context are manufacturing research and research on quality control or on the qualities of the product. Many American products have lacked the quality necessary to compete in a global market, which suggests a lack of funding at least by the private sector in this arena.

Some might question the necessity of funding RTD in all six arenas. The importance of research in all six arenas is highlighted in what are typically referred to as the high tech sectors in the United States. In these sectors, radical product innovations usually require radical advances in knowledge in more than one arena. Consider the case of semiconductors and Moore's law, which reflects not only radical advances in performance characteristics of the product but also necessitates radical advances in knowledge about how to manufacture the new generation of chips and how to achieve quality control. The same is true of radical new products in other sectors.

It could be the science may already be in place, but there is a need for some radical advances in one or two other arenas. An interesting case in point exists in cancer research. The National Cancer Institute (NCI) in the United States had to develop a radical solution to the problem of clinical trials so as to recruit enough patients willing to try experimental protocols under double blind conditions. The NCI created, interesting enough, a network system of hospitals, clinics, and physicians for recruiting patients and testing new protocols in double-blind experimental trials.

One of the advantages of this systemic framework for describing the complex processes of innovation is that it broadens our concept about the types of innovations beyond the usual distinctions. Product and process (manufacturing) innovation have been joined to research about commercialization and quality control including the development of new qualities. More importantly, the two forms of scientific research, basic and applied, are united together with industrial innovation, a connection that has been underemphasized particularly in the sociology of science (Jordan, 2006). The key point is that a radical new product might develop out any one or two of these arenas. Thus, the word "balance" means "assessing what is the appropriate amount for each arena." Although ideas do move back and forth in the idea innovation network, the research problems in the different arenas, which in turn necessitate increases in expenditure, occur across time. Many of the high-risk industrial innovations may take as many as ten years of research of various kinds. Thus, evaluators need to consider this aspect as well in assessing the amount of expenditure.

Usually within a technological sector, the expensive unsolved problems for research are quite well understood and can be isolated in a specific arena. In pharmaceutical research, for example, clinical trials are quite expensive and yet are fundamental for successful commercialization of a new drug. In automobiles it is more likely to be engineering research for product development whereas in aircraft research the need for applied research on fuel efficient engines or more light-weight materials for the body of the aircraft are perennial research questions. Radical new products may also require new forms of distribution and marketing as we have seen with cell phones to take a recent example. Since this framework has not been used extensively and because there is a tendency to focus primarily on basic, applied, and product development research, presently norms about the ideal amounts do not exist.

The Distribution of Projects by Portfolio Within Each of the Six Arenas. Governments choose missions and the question is not just how much to spend but whether these missions require radical solutions or solutions that necessitate a broad scope focus. To take a European example, France when they decided to develop high speed trains had to redesign everything including ticket sales. This means radical advances in basic research, applied research, product development, manufacturing, quality control, and commercialization of research. Reporting the distribution of

RTD expenditures must start with the basic question of the aims of that funding, that is, how radical an advance does the government desire in the specific technological sector being assessed. This leads naturally into another way of describing research gaps—by the distribution of spending in the portfolio.

What are some standard ways for describing portfolios? Jordan (2006) in her research on the research environment has suggested that the two basic strategic choices are the relative emphasis on incremental versus radical breakthrough or high-risk research and the relative emphasis on many narrow scoped projects versus a few broad-scoped programs. Practically, we are suggesting that the amount of money, both public and private, spent in each of these four categories should be assessed within each of the six arenas. For example, the United States government would like to have radically new technologies to reduce dependence on oil. Given this mission statement, then one examines the portfolio of projects in the various energy sectors associated with alternative energies and determines the relative emphasis on radical advances and/or broad scope projects designed to reduce dependence upon oil.

Given a decision to develop a radical product or service, the problem is then determining which arenas should have the radical advances in knowledge. This can vary from situation to situation. To continue with the example of developing alternative energies to reduce dependence upon oil, it varies depending upon the specific alternative. The science and technology for bio-fuels for automobiles has largely been developed. Here the problem in the United States is one of commercialization research to determine how best to secure customer interest and the creation of new distribution systems for this kind of fuel. In contrast, in research on photovoltaics, more of the need is for radical advances in the arenas of applied research and product development with more efficient thermo-couplings and in the materials employed to absorb sunlight.

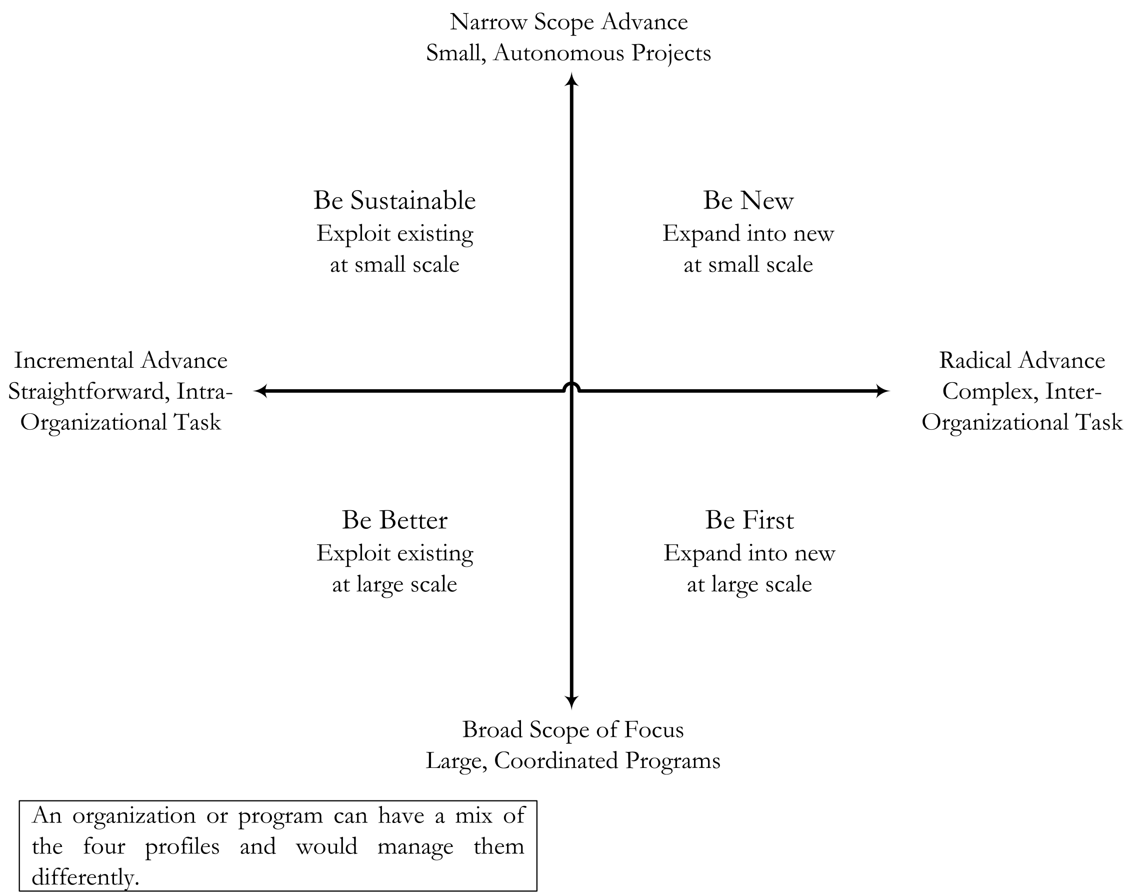
Likewise, certain missions necessitate a broad scope of focus for the research. The decision about the scope of the research project is more likely to be determined by the nature of the research work in one or more arenas. In the area of alternative energies the best example is fusion power research. Inherent in the problem is the need for very expensive equipment, a relatively large number of researchers, and expanded budgets. Here is an example of a broad scope of focus project that involves the goal of radical advances in a number of arenas, including manufacturing (reduce the cost of cooling), quality (reduce the risk of explosions), and commercialization (how to win acceptance of nuclear energy).

In this manner, an assessment would report the distribution of RTD funds, both public and private, in each of the six arenas in each of the following possible four kinds of research projects in a portfolio:

1. Narrow scope of focus and incremental advances
2. Narrow scope of focus and radical advances
3. Broad scope of focus and incremental advances
4. Broad scope of focus and radical advances

Figure 3 describes the four kinds of research projects. It is entirely possible that an arena would have a mixture of these depending upon the nature of the problem and the objectives of the government.

Figure 3. Dimensions of Strategy and Structure Define Four Research Profiles



The Distribution of Projects by Organizational Profile Attributes. The third set of indicators for evaluation at the micro level represents an inventory of the number of research organizations and projects within each arena of the technological sector that have appropriate characteristics for achieving the research objectives in a chosen portfolio. Rather than just report the number of research organizations, which is the more typical approach, our framework focuses attention on the number of research projects and organizations with particular attributes and characteristics. These characteristics are listed in Table 1 and have been developed from an eight-year research project developing a research environment survey (Jordan, 2005; Jordan, Streit, & Binkley, 2003). As one can observe there are the familiar themes of the organic organization, complexity or diversity in the division of labor, leaders with vision, and of course resources. The characteristics reflect organizations handling of tensions between flexibility and coordination and between inter-organizational collaborations and organizational control, to name two.

The assessment of the research environments would then report how many projects have the characteristics associated with the attributes related to radical advances in knowledge and how many have the characteristics associated with broad scope of focus and, of course, both. In each instance the issue is matching funding for each organizational profile to chosen strategies. Table 1 focuses on what might be called organizational failures, which Arnold (2004) called institutional failures but we believe should be kept quite separate from other kinds of institutional failures, meaning failures that are a consequence of institutional rules (Hage 2006a, 2006b; North, 1990).

Table 1. Organizational Profile Attributes Associated With Strategic Choices

Process Attributes for Radicalness	*Process Attributes for Large Scope*
Encourage exploration, risk taking	Clearly define goals and strategies
• Time to think and explore	• Research vision and strategies
• Pursuit of new ideas	• Sufficient, stable funding
• Autonomy in decision making	• Investing in future capabilities
Integrate ideas, internally and externally	Plan and execute well
• Internal cross-fertilization of ideas	• Project planning and execution
• External collaborations and interactions	• Project-level measures of success
• Integrate ideas and R&D portfolios	• Lab-wide measures of success
Encourage change and critical thinking	Build strategic relationships
• Sense of challenge and enthusiasm	• Relationship with sponsors
• Commitment to critical thinking	• Champion foundational research
• Identify new projects and opportunities	• Reputation for excellence

The Meso-Level Indicators: Six Arena Outputs, Network Connectedness, Overall Sector Performance

Increasing expenditures that are then allocated across the idea innovation network in accordance with strategic decisions made about the portfolio and the awarding of funds to research organizations capable of performing these strategic choices are not the only kind of policy intervention. More important is to know if the government is achieving what it desires. For example, suppose that the government has decided it wants radical innovations in a newly created set of bio-tech companies as did Germany (Casper, 2006). The issue is whether these new created bio-tech companies are in fact achieving radical bio-tech innovations. To have leading indicators for sector performance, one needs to assess the technical outputs *in real time* of each of the arenas and the strength of the network connectedness within and between arenas, and then relate these to the overall performance of the sector.

Real-time Technical Achievements in Each of the Arenas. In the Hage and Hollingsworth network theory (2000), each arena is perceived to have an output. In turn these outputs can be evaluated on the basis of how radical the advance is in that specific arena, whether it is basic research, applied research, product development, or manufacturing research. The radicalness of the advance in the technical output is a question of the context and how a radical advance is defined. Radical advances reflect a large percentage increase in the specific indicator defining the achievement in an output. Automation that improved the throughput of dishwashing machines of 300 percent clearly represented a radical advance as a consequence of manufacturing research. New paradigms or theories while rare are usually considered radical advances in science and are later recognized as such (e.g., by Nobel Prizes). Where does peer review fit into this measurement framework? Peer reviewers usually have the knowledge necessary to define typical rates of advances in an area as well as what is considered incremental and what would be a radical advance.

Some suggestions as to the kinds of measures of technical progress in each of the six arena's output are presented in Table 2. Our experience in measuring these outputs in scientific and technological research in different research organizations in the United States is that one can obtain a certain sense of progress in the second year of a project (Mote, Jordan & Hage, 2006).

The outputs have to be measured in real time, usually at least annually, for a number of reasons. One of the major ones, which we have already alluded to, is to provide quick feedback to policy makers. Papers and patents appear two to five years after the completion of a project in many cases; citations unfold even more slowly (Coryn, 2006). A second reason is that for policy or management interventions to be effective, they must occur while the project is still in progress. Although the reasons for lack of technical progress can adhere in many places in the innovation system, a good place to begin is with the management of the research project and whether the attributes appropriate for the specific strategic choices listed in Table 1 are present and to what degree.

Table 2. Indicators of Technical Output for each Functional Arena in the Idea Innovation Network

Functional Arena	*Measures of Scientific/Technical Advanced in Output*
Basic Research	• Percent increase in the modeling of some scientific behavior • Solution to a central problem • Identification of new concepts or processes
Applied Research	• Percent increase in control over some desired attribute

Product Development or product innovation	• Percent increase in different performance characteristics weighted by their importance • Addition of new properties to the functionality of the product
Production Research or process innovation	• Percent increase in productivity • Percent increase in customization
Quality Control Research and research on qualities	• Percent decrease in defects • Percent decrease in operating costs • Percent decrease in various externalities weighted by their importance
Commercialization Research	• Percent increase in customer satisfaction • Percent decrease in delivery time

Still a third reason to measure technical progress in each arena is to establish the links between short-term evaluations and medium- and longer-term evaluations. This speaks to two problems in systemic evaluations identified: the tenuousness of system evaluations of the medium and long term and the selection of quantifiable indicators that are easy to collect, but which deny the complexity of the innovation process and run the risk of irrelevancy. We believe Table 2 helps solve this problem. Beyond this, by establishing the missing link between short and medium or long term evaluations at the systemic level, one is also constructing a theory of the national system of innovation and developing a number of insights about institutional theory.

Measuring the technical achievements in the six outputs may appear to be an expensive and formidable task. Since each of these outputs tends to be the concern of a particular agency or ministry, the cost can be spread among them because the information about each arena individually is of value and collectively more than the simple sum of the six parts. For example, the ministry of the environment would want to know if the products are being manufactured with qualities that protect the environment in various ways, the ministry of commerce is interested in the establishment of new methods for advertising and distributing products, the ministry of technology, if there is one separate from either science or industry, is more concerned with research on manufacturing and product development.

Further, Table 2 provides suggestions for each of the six arenas, but it may not be necessary to measure all six. This is a question of how much functional differentiation has occurred. For example, if basic, applied and product development research are combined in a bio-tech company, and the manufacturing, quality and marketing research are combined in a pharmaceutical firm, one can concentrate on the technical outputs of product development of the bio-tech companies and the product outputs of the pharmaceutical firms. This represents only a first approximation and again, it may be necessary to examine the outputs of the other arenas within the bio-tech companies and the pharmaceutical firms because the blockages can be organizational with bottlenecks between basic and applied research in the bio-tech companies or between manufacturing research and quality research in the pharmaceutical companies. For this, internal network analysis can be quite valuable (Mote, 2005).

As more and more of the arenas become functionally differentiated then one is forced to measure the technical outputs of each arena. It speaks to the issue of understanding the innovation processes at the level of the idea innovation network in a technical sector, avoiding the errors in the business systems literature and other institutional studies that tend to generalize from one technological sector to all others (Hollingsworth, 1997; Whitley, 1992 a, 1992b).

The Strength of Connectedness Between Arenas. It should be clear that the six arenas of the idea innovation network need to have robust connections to have the desired result. In the past, when all arenas were within the same organization such as Siemens, DuPont, and Proctor and Gamble, the issue of connectedness did not present a significant problem. However, over time, connectedness has become problematic, even within the same organization. For example, disconnectedness occurred between Bell Laboratories and AT&T and between the research department (PARC) of Xerox and the main company. The real problems of connectedness start to grow as an entire functional arena becomes disconnected, as for example all of basic research is located outside the other organizations involved in some technological sector. The van Waarden and Oosterwijk (2001) EU study indicates how these evolutionary processes have unfolded in telecommunications and pharmaceuticals in Austria, Finland, Germany, and the Netherlands. A number of new sub-networks in specific arenas emerged to handle the problems of technical advances and in turn these were connected to the larger sector network of knowledge transfers.

As each arena becomes differentiated into separate research organizations, it becomes necessary to be concerned about the extent of the connection between these differentiated arenas. Even more important is the strength of the connection. Hage and Hollingsworth (2000) argue that as the radicalness of the technical achievement in a specific arena increases, the more frequent and intense interaction with other arenas must become to transfer the tacit knowledge involved in the radical advance.

What are some indicators of the strength of the connection? (1) The transfer of people from one research group to another, both within and among organizations; (2) joint research projects involving face to face collaboration among researchers, as distinct from long distance collaboration; (3) joint publications; (4) the strength of managerial, financial, and research ties among organizations in joint ventures; and (5) the strength of ties among actors in research consortia (Nieminen & Kaukonen, 1999). In van Waarden and Oosterwijk (2001), they observe a large number of different ways in which connectedness was established: joint ventures, user groups, product teams, patent pools, collective trademarks, technology clusters, partnerships, alliances, and even virtual firms. But it should be observed that number of these mechanisms present does not necessarily reflect the transfer of tacit knowledge. Just as one wants to measure the knowledge advances in each arena, one also needs to measure exchanges of tacit knowledge within and between arenas, especially the differentiated ones.

One of the more exciting literatures that measures the exchange of tacit knowledge, which has emerged recently in United States evaluation studies, is the work on the emergence of knowledge communities as an attempt to measure the value added by investments in R&D (Mohrman & Gailbraith, 2006; Mohrman, Gailbraith, & Monge, 2005). In the case of the study of the emergence of modeling and simulation technology at a national laboratory, the authors use a general framework that suggests networks form within an ecology of dynamic and overlapping communities and the co-evolution of social, human, and intellectual capital.

The idea innovation network highlights two kinds of networks that should be of concern for policy makers and thus evaluators to measure:

1. Sub-networks of small research organizations within an arena
2. Networks of organizations, whether small or large, across differentiated arenas

The former are important when governments are interested in creating technical pools, which was clearly the objective of the American government when it changed its anti-trust laws to encourage research consortia. In this instance some of the research consortia involved quite large organizations and they were primarily concentrated in the basic and applied arenas of research leaving product development to individual companies that desired to pursue particular market niches. In contrast, the networks of organizations become especially important in linking differentiated arenas. But that said, the more critical point is that it is not necessary and even counter-productive to have networks and sub-networks everywhere. Here is a strategic location for more research to guide evaluators and policy makers.

Recently a large number of network studies have emerged including the new research on visualization tools that mines large data sets involving papers, patents, citations of either papers or patents or both (Börner, 2006; Wagner, 2005). Many of these efforts are attempts to measure the pay-off from investments in RTD. However useful these network analyses are for examining the consequences of certain government policies, they are focusing on what *is* rather than what *can be*. The ISF and the measurement of connectedness in real time attempts to inform government about network failures or "what could be." With policy reformulations, government could potentially achieve more pay-off from their investments in science than they do presently.

Overall Performance of the Technology Sector. The technical outputs of each arena are a means to an end, locating one of the reasons why a nation may not be achieving its goals. The lack of connectedness can be another reason. But in the final analysis, evaluators need to assess the overall performance of the technological sector. Rather than a single assessment of the overall performance, we suggest the following dimensions:

1. The degree of radicalness on the various dimensions of the product mix
2. The average speed of product dimension development or time to market
3. The commercial success of the product mix in sales and trade balances or the technological position globally

It goes without saying that the level of analysis here is not the single research organization but all of those concerned with producing products in a particular technological sector and therefore the product-mix in the technological sector. First, while we are using economic examples, we would argue that the same logic can be applied to health products, military weapons, and national security. The industrial innovation literature (Hage, 1999, Hage & Meeus, 2006) has tended to focus on product innovation and even radical product innovation in terms of functionality rather than observing that that is one of a set of interrelated dimensions. Second, consistent with the repeated importance of research on qualities, and research on the customization of manufacturing, product mixes at the sector level should be evaluated along a variety of different dimensions to understand correctly the competitive position of the country. Third, it is fairly easy to determine the competitive position of the product mix for policy makers using trade journals and statements of marketing executives who are acutely aware of the relative strengths and weaknesses of their products in a global context.

Also, the average speed of development across the various dimensions involved in the product mix has increased and become a critical factor in maintaining a country's competitive position in the high tech sectors. The Japanese set the bar quite high when they created the hybrid car, a radical product innovation, in just 15 months (Nonaka & Peltokorpi, 2006).

Conclusions

The recent calls in the evaluation literature for theory-led evaluation and better analyses of the systemic obstacles and blockages to innovation, to explain why policy objectives have or have not been reached can reflect a new era in RTD evaluation. These changes in the methodology of evaluation would focus on what could be rather than what is done. If theories-led evaluations can determine what obstacles or blockages are preventing the realization of the objectives of the policy makers, then governments could begin the process of designing better interventions to achieve more effective innovation as indicated by our indicators of sector performance.

But what is the theory or theories that should lead the evaluations? Although some have called for the use of national systems of innovation literature, this literature is largely descriptive and only covers the macro-level. We have suggested that the systemic framework should be a theories-led evaluation, that is, include theories from the macro-, meso-, and micro-levels. One advantage of starting to construct the innovation systems framework with the idea innovation network theory of Hage and Hollingsworth (2000) is that is does help integrate these other literatures at the same time that it is provides a much more complex perspective on the throughput of the innovation processes. Theories from each of the three levels of macro, meso, and micro are essential if one is to understand the blockages and obstacles because they could be located in any part of the system.

Despite pleas from governments that one analyze the entire scientific and technological system, we suggest that instead it is more desirable to evaluate sector by sector. If need be, these separate sector analyses could be combined into a total assessment. Another implication of the call for analysis of blockages and obstacles is that technical advances have to be measured in real time rather than waiting for the appearance of papers, patents and citations to them. If governments are to ever focus on improving policies, they must have quick feedback, not only on whether they are failing to achieve their objectives but more critically on the reasons why. Only then can governments learn how to fashion policies to improve. Only then can evaluators relate short-term accomplishments to long-term sector performances. Then evaluations would truly inform governments and reduce the current cynicism about evaluation that almost always states that research investments are successful.

The ISF for RTD evaluation not only can inform governments but it can inform the theories that have been used to further construct the framework. The efficacy of various theories can be tested across the three levels of analysis, that is, assessing the macro-, the meso-, and micro-contributions to any particular issue. Government interventions, if evaluated this way, could then provide some important answers to questions that have been imposed in these various literatures such as: What are the best kinds of linkages for the transfer of tacit knowledge? Can governments overcome path dependency? How much autonomy do research organizations have from their institutional environment? Answers to questions such as these would considerably advance the sophistication of social science theory.

Acknowledgments

Work presented here was completed for the Office of Basic Energy Sciences in the United States Department of Energy Office of Science by Sandia National Laboratories, Albuquerque, New Mexico, under contract DE-AC04-94AL8500. Sandia is operated by Sandia Corporation, a subsidiary of Lockheed Martin Corporation. Opinions expressed are solely those of the authors.

References

Alter, C., & Hage, J. (1993). *Organizations working together: Coordination in interorganizational networks*. Beverly Hills, CA: Sage.

Archibugi, D., & Pianta, M. (1992). *The technological specialization of advanced countries: A report to the EEC on international science and technology activities*. Boston, MA: Kluwer.

Arnold, E. (2004). Evaluating research and innovation policy: A systems world needs systems evaluations. *Research Evaluation*, *13*(1), 3-17

Borner, K., Sanyal, S., & Vespignani, A. (2006). Network science: A theoretical and practical framework. *Annual Review of Information Science and Technology*, *41*, 537-607.

Browning, L., Beyer, J., & Shetler, J. (1995). Building cooperation in a competitive industry: SEMATEC and the semiconductor industry. *Academy of Management Journal*, *38*, 113-151.

Campbell, J. L. (2004). *Institutional change and globalization*. Princeton, NJ: Princeton University Press.

Campbell, J. L., Hollingsworth, R., & Lindberg, L. (1991). *Governance of the American economy*. New York, NY: Cambridge University Press.

Casper, S. (2006). Exporting the silicon valley to Europe: How useful is comparative institutional theory. In J. Hage & M. T. H. Meeus (Eds.), *Innovation, Science and Institutional Change: A Handbook of Research* (pp. 483-504). Oxford, UK: Oxford University Press.

Coryn, C. L. S. (2006). The use and abuse of citations as indicators of research quality. *Journal of MultiDisciplinary Evaluation*, *3*(4), 115-120.

Dinges, M. (2006). Pressure, deception and peer review. *Plattform: Forschungs und technologiesevaluierung*, *28*(Arpil), 11-22.

Doz, Y. L., & Hamel, G. (1998). *Alliance advantage: The art of creating value through partnering*. Boston, MA: Harvard Business School Press.

Dussauge, P., & Garrette, B. (1999). *Cooperative strategy: Competing successfully through strategic alliances*. Chichester, UK: Wiley.

Edquist, C. (Ed.). (1997). *Systems of innovation: Technologies, institutions and organizations*. London, UK: Pinter.

Edquist, C., & Hommen, L. (1999). Systems of innovation: Theory and policy for the demand side. *Technology in Society*, *21*, 63-79.

Gomes-Casseres, B. (1996). *The alliance revolution: The new shape of business rivalry*. Cambridge, UK: Harvard University Press.

Guerrieri, P., & Tylecote, A. (1997). Interindustry differences in technical change and national patterns of technological accumulation. In C. Edquist (Ed.), *Systems of innovation: Technologies, institutions and organizations* (pp. 107-129). London, UK: Pinter.

Hage, J. (1999). Organizational innovation and organizational change. *Annual Review of Sociology*, *25*, 597-622.

Hage, J. (2006). Introduction, part IV. In J. Hage & M. T. H. Meeus (Eds.), *Innovation, science and institutional change: A handbook of research* (pp. 415-422). Oxford, UK: Oxford University Press.

Hage, J., & Hollingsworth, J. R. (2000). A strategy for analysis of idea innovation networks and institutions. *Organizational Studies*, *5*, 971-1004.

Hage, J., & Meeus, M. (2006). Product and process innovation, scientific research, knowledge dynamics, and institutional change: An introduction. In J. Hage and M. T. H. Meeus (Eds.), *Innovation, science and institutional change: A handbook of research* (pp. 1-19). Oxford, UK: Oxford University Press.

Häkansson, H. (1990). Technological collaboration in industrial networks. *European Management Journal*, *8*, 371-379.

Hall, P., & Soskice, D. (Eds.). (2001). *Varieties of capitalism*. Oxford, UK: Oxford University Press.

Harbison, J. R., & Pekar, P. P. (1998). *Smart alliances: A practical guide to repeatable success*. San Francisco, CA: Jossey-Bass.

Hollingsworth, J. R. (1997). Continuities and changes in social systems of production: The cases of Japan, Germany, and the United States. In J. R. Hollingsworth & R. Boyer (Eds.), *Contemporary capitalism: The embeddedness of institutions* (pp. 256-310). New York, NY: Cambridge University Press.

Hollingsworth, J. R., & Boyer, R. (Eds.). (1997). *Contemporary capitalism: The embeddedness of institutions*. New York, NY: Cambridge University Press.

Jarillo J. C. (1993). *Strategic networks: Creating the borderless organization*. Oxford, UK: Butterworth-Heinemann.

Jordan, G. (2006). Factors influencing advances in basic and applied research: Variation due to diversity in research profiles. In J. Hage & M. T. H. Meeus (Eds.), *Innovation, science and institutional change: A handbook of research* (pp. 173-195). Oxford, UK: Oxford University Press.

Jordan, G. (2005). What is important to R&D workers. *Research Technology Management*, *48*(3), 23-32.

Jordan, G., Hage, J., & Mote, J. (2006, April). *A theory-based framework for evaluating diverse portfolios of scientific work*. Paper presented at the New Frontiers in Evaluation conference, Vienna, Austria.

Jordan, G., Hage, J., & Mote, J. (2007). A theories-based systemic framework for evaluating diverse portfolios of scientific work part two: Macro indicators and policy interventions. *Science and Public Policy*. Forthcoming.

Jordan, G., Streit, L. D., & Binkley, J. S. (2003). Assessing and improving the effectiveness of national research laboratories. *IEEE Transactions in Engineering Management*, *50*(2), 228-235.

Kline, S., & Rosenberg, N. (1986). An overview of innovation. In R. Landau & N. Rosenberg (Eds.), *The positive sum strategy*. Washington, DC: National Academy Press.

Kogut, B., Shan, W., & Walker, G. (1993). Knowledge in the network and the network as knowledge: The structuring of new industries. In G. Grabher (Ed.), *The embedded firm: On the socioeconomics of industrial networks* (pp. 67-94). London, UK: Routledge.

Lundvall, B. A. (1992). *National systems of innovation: Towards a theory of innovation and interactive learning*. London, UK: Pinter.

Lundvall, B. A. (1993). Explaining interfirm cooperation and innovation limits of the transaction-cost approach. In G. Grabher (Ed.), *The embedded firm: On the socioeconomics of industrial networks*. London, UK: Routledge.

Malerba, F., & Orsenigo, L. (1993). Technological regimes and firm behavior. *Industrial and Corporate Change*, *2*, 45-71.

Malerba, F., & Orsenigo, L. (1996). Technological regimes and sectoral patterns of innovative activities. *Industrial and Corporate Change*, *6*, 83-117.

Marburger, J. H. (2005). *Statement of Dr. John H. Marburger, III," President's Science Adviser and Director, Office of Science and Technology Policy, to the United States House of Representatives, Committee on Science, Fiscal Year 2006 Federal R&D Budget, February 16, 2005*. Retrieved June 6, 2007 from www.house.gov/science/hearings.

Molas-Gallart, J., & Davies, A. (2006). Toward a theory-led evaluation: The experience of European science, technology and innovation policies. *American Journal of Evaluation, 27*(1), 64-82.

Mohrman, S., Galbraith, J., & Monge, P. (2004). *Network attributes impacting the generation and flow of knowledge within and from the basic science community*. University of Southern California, The Center for Effective Organizations.

Mohrman, S., Galbraith, J., & Monge, P. (2006). Network attributes impacting the generation and flow of knowledge within and from the basic science community. In J. Hage & M. T. H. Meeus (Eds.), *Innovation, science and institutional change: A handbook of research* (pp. 196-216). Oxford, UK: Oxford University Press.

Mote, J. E. (2005). R&D ecology: Using 2-mode network analysis to explore complexity in R&D environments. *Journal of Engineering and Technology Management*, *22*, 93-111.

Mote, J. E., Jordan, G., & Hage, J. (2006, February). *Measuring radical innovation in real time*. Paper presented at the meeting of the American Association for the Advancement of Science, St. Louis, MO.

Nieminen, M., & Kaukonen, E. (1999). University research in innovation systems: Reflections based on the Finnish case. In G. Schienstock and O. Kuusi (Eds.), *Transformation towards a learning economy—The challenge for the Finnish innovation system*. Helsinki: SITRA.

Nelson, R. R. (Ed.). (1993). *National innovation systems: A comparative study*. Oxford, UK: Oxford University Press.

Nonaka, I., & Peltokorpi, V. (2006). Knowledge-based view of radical innovation: Toyota Prius case. In J. Hage & M. T. H. Meeus (Eds.), *Innovation, science and institutional change: A handbook of research* (pp. 88-104). Oxford, UK: Oxford University Press.

North, D. (1990). *Institutions, institutional change, and economic performance*. Cambridge, UK: Cambridge University Press.

O'Doherty, D. (Ed.). (1995). *Globalisation, networking, and small firm innovation*. London, UK: Graham & Trotman.

Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy*, *13*, 343-373.

Picher, R. (2006). The ERC and policy makers' expectations: Evaluation as a change agent. *Plattform: Forschungs-und technologiesevaluierung*, *28*, 4-10.

Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institute.

Wagner, C. (2005). Six case studies of international collaboration in science. *Scientometrics*, *62*(1), 3–36.

Van de Ven, A. H., & Polley, D. (1992). Learning while innovating. *Organization Science*, *3*, 92-116.

van Waarden, F., & Oosterwijk, H. (2001). Turning tracks? Path dependence, technological paradigm shifts, and organizational and institutional change. In F. van Waarden, B. Unger, G. Schienstock, P. Tulkki, R. Kaiser, E. Grande, & H. Oosterwijk (Eds.), *Building bridges between ideas and markets, part I: Some summary findings and part II, annex 1, country-specific reports* (Final Report of the TSER-Funded Project). Utrecht: Utrecht University.

Weiss, C. (1997). Theory-based evaluation: Past, present and future. *New Directions for Program Evaluation*, *76*, 41-55.

Whitley, R. (1992a). *Business systems in East Asia: Firms, markets and societies*. London, UK: Sage.

Whitley, R. (1992b). *European business systems: Firms and markets in their national contexts.* London, UK: Sage.

GRETCHEN JORDAN is Principal Member of Technical Staff, Sandia National Laboratories.

JERALD HAGE is Director, Center for Innovation, Department of Sociology, University of Maryland.

JONATHON MOTE is Assistant Research Scientist, Center for Innovation, Department of Sociology, University of Maryland.