Research Value Mapping and the National Cooperative Program for Infertility Research (NCPIR): The Feasibility of a "Capacity Evaluation"

Report Submitted by

Barry Bozeman

Monica Gaughan

Research Value Mapping Program
School of Public Policy
Georgia Institute of Technology
D.M. Smith Building
Atlanta, Georgia 30332

David Roessner
Washington, DC

June, 2002

A Report to the National Institute of Child Health and Human Development,
in fulfillment of a subcontract with IQ Solutions, Inc.

Research Value Mapping and the National Cooperative Program for Infertility Research (NCPIR): The Feasibility of a "Capacity Evaluation"

Summary

As stated in the work scope, this study's purpose is to 1) examine the feasibility of using both qualitative and quantitative Research Value Mapping (RVM)[1] methods to implement a full-scale review of the NCPIR; and 2) prepare an evaluation design and work plan to conduct the review.

Before introducing the NCPIR review questions, we begin with the theory behind the proposed methods. Commonly, evaluations of science programs are based on "output analysis" (e.g. counting publications and citations). The RVM methods suggested here represent a "capacity evaluation." This approach identifies and measures factors that may enhance the *ability* (e.g., capacity) of persons and institutions to contribute to scientific and technical knowledge. A capacity evaluation examines institutions as organizations, in addition to factors on an individual level. When focusing on the individual researcher level, we try to understand how "scientific and technical human capital" (S&T human capital) is developed,and how professional, organizational and social network ties may mitigate the use of the skills.[2]

The questions for the NCPIR review prepared for this feasibility study include:

1. Is the NCPIR serving as a "national resource" for the science community in infertility research? Is it producing information and research products, such as reagents or new potential diagnostics that are useful to other research centers and individual researchers?

2. Are the NCPIR centers making linkages to facilitate translation research goals with institutions supported by the NICHD Reproductive Sciences Branch and other non-NICHD supported institutions? How effective is the NICHD staff in facilitating these linkages and arranging new collaborative projects?

3. As currently configured, what can the NCPIR potentially achieve in research outputs and social impacts? What alternative structures could contribute to additional accomplishments and increased productivity?

4. Does the NCPIR provide adequate opportunities for research fellows pursuing infertility studies with a human application in a supportive and collaborative setting that enhances their research productivity?

[1] See Appendix A for information on the RVM methods developed at the Georgia Institute of Technology (Georgia Tech).

[2] See Appendix B for information on the development of the "Capacity Evaluation" and "S&T Human Capital" concepts used by Georgia Tech in research on science programs.

5. How do NCPIR-supported research fellows compare with other translation- and clinician-scientists (who may or may not have training from other research centers) in terms of their contributions to infertility diagnostics and treatments, as well as their overall research productivity and career paths?

6. What data, beyond that already being collected by the infertility research centers and the NCPIR staff, must be obtained to sustain future evaluations using similar methods? Can some of these data be collected and maintained relatively easy?

While the feasibility analysis that follows considers each of these questions in turn, the basic approaches proposed for the NCPIR review are summarized below:

- Case studies will be undertaken of the two NCPIR centers: Massachusetts General Hospital and the University of Pennsylvania. The case studies will focus on the following dimensions:

    1. Organization structure and design, management style
    2. Funding and resource base
    3. Scientific and technical output profile (composition of outputs in terms of articles, technology, technical assistance)
    4. Educational and human resource activities
    5. Outreach activities
    6. Linkages to other researchers and research institutions (focusing on number, types and functions of linkages).

- The case studies will be used to identify the population of researchers who have been affiliated with the NCPIR centers, past and present. We shall collect curriculum vitae (CVs) from these individuals to study the impacts of center affiliation on their career trajectories and productivity.

- We will conduct comparison case studies, focusing on diverse NICHD centers, and perhaps other NIH centers, examining the same dimensions as above. This will provide us a baseline for comparing the NCPIR centers.

- We will examine the comparison NIH case studies and the NCPIR case studies in light of our existing data on more than 50 science centers, including the National Science Foundation (NSF) Science Centers, the NSF Engineering Research Centers, and the Department of Energy's (DOE) research centers at universities and national multi-program laboratories. This analysis should suggest alternative ways to structure the NCPIR organization and management. (Note: while all centers in our existing data base will be used as a basis for comparison, the level of detail in existing cases varies, some being full cases and some "mini-cases."

We will rely most heavily on the full cases but will use the smaller ones to the extent they are comparable).

- We will develop comparison groups to assess what is happening at the individual researcher level. The primary analysis itself will be based on the review of career trajectories from CV's. Comparison data will come from individuals identified through the case study sites, the SCCPRR centers, and from among clinician scientists identified through the Postdoctoral Individual National Research Service Awards (F32) and the Clinical Investigator Awards (K08). This should permit us to compare these groups and individuals to the NCPIR centers in terms of their ability to enhance the "S&T human capital" and to determine differences in career patterns among the respective groups of researchers. In matching the NICHD-trained researchers to those in the comparison groups, we will rely on random comparisons, rather than seeking to impose matching criteria. However, to ensure some level of comparability, we will match on at least specialization (discipline, field, research, medical specialty, as appropriate) and time-cohort variables.

- We will conduct a survey of relevant NCPIR stakeholders[3] to obtain their individual assessments on the use, or non-use, of the research and other knowledge products created by NCPIR center researchers. These stakeholders will be identified in consultation with NICHD staff, the advisory board for this project, and NCPIR personnel. A questionnaire instrument will be mailed to a sample of the stakeholder group so identified.

Using data from the NCPIR center case studies, the comparative institutional case studies, the NCPIR stakeholder survey, and the CV's of past and present NCPIR researchers in comparison to other researchers, we expect to provide useful answers to each of the research questions posed above. The information collected should be useful to both the NICHD leadership and staff and to the NCPIR center managers by providing:

- A different view of program performance;
- Help in understanding the role of NCPIR in building capacity in the fields of fertility and infertility research as well as related fields;
- Help in identifying, analyzing and assessing NCPIR center approaches to organizing and managing research;
- A number of plausible alternatives (based on the comparison case studies performed here as well as the Georgia Tech RVM Program database of case studies).

These alternatives are a significant feature of the evaluation. They should aid analysis of the performance of the NCPIR centers and provide insight into other approaches to

[3] The definition of NCPIR stakeholders will be approached from the concept of "Knowledge Value Alliances," developed by Rogers and Bozeman (2001) at Georgia Tech as part of the RVM methods. For more background information see Appendix C.

structuring and managing centers, while outlining a set of additional managerial options for NCPIR directors and NICHD officials to consider.

Research Value Mapping and the National Cooperative Program for Infertility Research (NCPIR): The Feasibility of a "Capacity Evaluation"

Organization of the Report

This feasibility study is organized into five sections. Section I reviews of the purposes of the feasibility study and the questions for the NCPIR review. Section II identifies the data sources and summarizes the procedures used for the feasibility study. Section III gives the rationale for a "capacity evaluation." Section IV provides findings about the feasibility of applying RVM concepts and methods to the NCPIR review questions. Section V summarizes the expected usefulness of the RVM concepts and methods for center program design and management.

Section I. Purpose of the Study

The purpose of this study is to examine the feasibility of using both qualitative and quantitative Research Value Mapping (RVM) methods for a full-scale review of the NCPIR and to prepare an evaluation design and work plan to conduct the NCPIR review.[4] The specific NCPIR review questions were developed by the NICHD to guide the feasibility study. They were refined after a January 18, 2002 conference call that included the NCPIR Review Technical Workgroup, the Georgia Tech team, IQ Solutions, and NICHD staff. The revised review questions to be addressed in this feasibility study follow:

1. Is the NCPIR serving as a "national resource" for the science community in infertility research? Is it producing information and research products, such as reagents or new potential diagnostics that are useful to other research centers and individual researchers?

2. Are the NCPIR centers making linkages to facilitate translation research goals with institutions supported by the NICHD Reproductive Sciences Branch and other non-NICHD supported institutions? How effective is the NICHD staff in facilitating these linkages and arranging new collaborative projects?

3. As currently configured, what can the NCPIR potentially achieve in terms of research outputs and social impacts? What alternative program/operational structures could contribute to additional accomplishments and increased productivity?

[4] The Feasibility Study, initiated on November 13, 2001, is the NICHD's first phase in conducting the NCPIR review.

4. Does the NCPIR provide adequate opportunities for research fellows pursuing infertility studies with a human application in a supportive and collaborative setting that enhances their research productivity?

5. How do NCPIR-supported research fellows compare with other translation- and clinician-scientists (who may or may not have training from other research centers) in terms of their contributions to infertility diagnostics and treatments, as well as their overall research productivity and career paths?

6. What data, beyond that already being collected by the infertility research centers and the NCPIR staff, must be obtained to sustain future evaluations using similar methods? Can some of these data be collected and maintained relatively easy?

The intent of the feasibility study was not to provide answers to the questions above but to determine the feasibility of applying evaluation methods to them.

Section II. Feasibility Study Procedures and Data Sources

The feasibility study relied on three data sources:

1. Interviews with NICHD personnel and grantees
2. Public domain documents (e.g. legislation, planning documents)
3. Files and records

The interviews were conducted with a semi-structured protocol in some cases and were open-ended in other cases. This is appropriate for a feasibility study. In the actual research, protocols will be developed for the semi-structure interviews. The objectives of our interviews differed considerably according to the role of the respondent. Without characterizing particular ones, the interviews had the following objectives:

1. To familiarize the authors at Georgia Institute of Technology (Georgia Tech) with the history and current programs of the NICHD and, especially, the NCPIR centers.
2. To help us understand how the interviewees viewed their respective roles.
3. To determine various stakeholders views about the type of information that could prove useful from an evaluation or assessment.
4. To identify readily available data sources and data gaps.
5. To determine any possible difficulties in obtaining information from participants (we found no significant barriers).
6. To obtain an understanding of the institutional and work context of the NCPIR's and their day-to-day work activities.

With respect to the files, after signing a confidentiality agreement, we spent a day at NICHD headquarters examining proposals and proposal reviews. With respect to documents, we solicited documents, CV's and program summaries at the two NCPIR sites.

Section III. The Rationale for "Capacity Evaluation"

The questions above are best answered by a "capacity evaluation" approach to the application of RVM concepts and methods. Most approaches to the evaluation of research programs focus on the program's discrete contributions to scientific and technical knowledge (e.g. articles produced, patents, citations). C*apacity-based research evaluation*, as defined by Bozeman (2001), is "an approach seeking to identify and measure factors presumed to enhance the *ability* of persons, institutions or social agglomerations to contribute to scientific and technical knowledge."

Capacity evaluation is designed for evaluations that have the following characteristics: (1) the focus is less on particular outputs (articles, patents) than on the sustained ability to contribute to a field or to the user community; (2) extra-organizational and inter-organizational relations are as importance as intra-organizational relations; (3) network ties facilitate performance; (4) development of "scientific and technical human capital" (S&T human capital) is among the organization's objectives.

The characteristics of capacity evaluation seem to track well against the NCPIR review questions above.

- The question: *"Is the NCPIR serving as a 'national resource' for the science community in infertility research?"* is clearly a capacity-related question.

- *"Are the NCPIR infertility research centers themselves making appropriate and necessary linkages with institutions in the Specialized Cooperative Centers Program in Reproductive Research (SCCPRR) to facilitate the translational research goals of both programs?"* This question about linkages is fundamentally a network and inter-organizational relations question.

- Similarly, the question: *"Does the NCPIR provide adequate opportunities for clinician-scientist research fellows to pursue infertility studies with a human application in a supportive and collaborative setting that enhances their productivity?"* underscores the importance of capacity building as well as the development of S&T human capital.

- A related S&T human capital question is: *"How do clinician-scientists with NCPIR-supported training compare with clinician-scientists (who may or may not have training from other research centers) in terms of their contributions to infertility diagnostics and treatments, as well as their overall productivity and career paths?"* This question cannot be answered with traditional output-focused approaches to evaluation.

See Appendix B for more details on the concepts of "capacity evaluation" and "S&T human capital" as applied by Georgia Tech to research and evaluation of science programs.

Section IV. Feasibility Study Findings and Recommended Methods

Based upon our experience in developing and applying RVM to diverse science programs and upon the data we have gathered in the feasibility study, we assess each question and offer recommendations about how to proceed for the NCPIR review. For each recommendation we provide a capsule description of the recommended approach, an assessment of the time and resources the approach would likely require, and the degree to which it can be replicated for other center program evaluations. More detailed information on the resources, tasks, and time frame needed to implement the recommendations is provided in the following tables:

- Table 1. NCPIR Review Evaluation Design;
- Table 2. NCPIR Review Work Plan;
- Table 3. NCPIR Review Time Frame; and
- Table 4. NCPIR Review Resource Estimate.

Question 1: Is the NCPIR serving as a "national resource" for the science community in infertility research? Is it producing information and research products, such as reagents or new potential diagnostics that are useful to other research centers and individual researchers?

Feasibility Assessment. This question clearly can be answered, provided there is agreement on the rough definition of the "science community in infertility research." Our interviews indicate that the NCPIR researchers have a high degree of confidence that they can identify the target audience for their work and, we conclude, that it is a sufficiently distinct group that we can identify the members of the population and draw a sample.

Recommended Approach. We recommend a survey of the NCPIR stakeholders, defined by applying the RVM concept of the "knowledge value alliance," (KVA) an approach that is a standard part of RVM methods developed at Georgia Tech.[5] Generally, the KVA is defined as an institutional framework binding together, in a "knowledge covenant," a set of directly interacting individuals, from multiple institutions, each contributing resources to pursue a transcendent knowledge goal (the basis of the covenant). Inherent in the KVA concept is that a key objective of the alliance is to generate multiple uses and multiple types of use (e.g. technology development, skill enhancement, understanding of fundamental phenomena).

The KVA starts with a "knowledge compact," usually through a formal alliance

[5] See Appendix C, Knowledge Value Alliance (KVA) and Organizational Design and Management, for more information on the development of this concept and how it has been applied to research on science programs.

agreement (e.g. a general contract, a Cooperative Research and Development Agreement) and terminates when resources are no longer applied to specific cooperative agreements, or when resources are no longer shared among parties. The KVA is an interactive group but there is no need for each member to interact directly with every member of the alliance. There must be links, however, among the members of the respective institutional representatives (those designated in the alliance agreement). The KVA acts as a selection mechanism, parsing specialized information (e.g. understanding of phenomena, understanding of technologies' product possibilities, skill in equipment operation or processes) for multiple uses by scientists in the immediate and adjacent research fields.[6]

In our judgment, the NCPIR, as an NICHD centers program, fits each of the criteria for a KVA, including the cooperative agreement mechanism (U54) and the program's congressional statute serving as the basis of a "knowledge covenant."

One longstanding RVM method associated with the analysis of the KVA is the "science community user survey." This approach was developed to understand not only the use of "knowledge products" (technologies, scientific research, technical processes), but the user community's awareness, assessment and rationale for not using the new technology.[7]

The KVA user survey involves the following steps:

1. With the close cooperation of the NCPIR center researchers, identify the entire population of individuals who should be construed as the "knowledge value alliance" and the research user community (e.g., bench researchers or clinical researchers who could potentially benefit from the knowledge produced by NCPIR researchers, including public and private funding agencies and their program managers, technology developers), and then, develop a population listing based on these responses.

2. Develop an inventory of the information and research products produced and to be assessed by the KVA user community.

3. Draw a sample from the KVA user community population. Part of this sample should be a random probability sample. But part of the sample should also be

[6] The need to postulate the KVA, rather than continuing to explain knowledge production activities in terms of research teams, projects and programs, arises from the fact that the latter do not capture the diversity and heterogeneity of knowledge production and uses we found in the RVM case studies (Bozeman, et al., 1998). Research projects and programs are often administrative units that reflect the jurisdiction of managers rather than the relevant dynamics of the knowledge production process. For an elaboration of the KVA rationale, see Rogers and Bozeman, 2001.

[7] For examples of the Research Value Mapping program's previous work with this method, see Moon and Bretschneider, 1997; Bobrowski and Bretschneider, 1994; Kingsley, Bozeman and Coker, 1995; Bozeman and Kingsley, 1997.

stratified to capture persons who have cited relevant works (thereby ensuring a sufficient number of individuals who are aware of the work) or who are "once removed" persons cited by the persons citing the NCPIR work. The sample should also include persons shaping the research product, including possible industrial clients and funding providers.

4. Study the use/value of products, determining levels of awareness and use/nonuse of the products, as well as the types of use and assessments of value. This is accomplished through either mailed questionnaires or through telephone interviews, depending upon the appropriate sample size.

Time and Resources. If we assume that the use/value survey entails a mailed questionnaire (or Internet based questionnaire, if suitable), this component of the study will be time- and resource-intensive. Such a study can be conducted in one year, start to finish. In addition to investigator time, the study would require one graduate research assistant devoted solely to the project and two undergraduate students.

Replication. Once the sampling population and survey instruments are developed, the approach can be replicated by others who have some knowledge of survey research. However, the costs will not likely diminish since there are no economies of scale or learning curve efficiencies. Each new survey will still require some tailoring and pilot testing.

1.

Question 2: ***Are the NCPIR centers making linkages to facilitate translation research goals with institutions supported by the NICHD Reproductive Sciences Branch and other non-NICHD supported institutions? How effective is the NICHD staff in facilitating these linkages and arranging new collaborative projects?***

Feasibility Assessment. While the information we collected for the feasibility study indicated relatively little linkage activity, the NICHD staff recommended that the NCPIR review should document such linkages as they exist, probe for why further linkages have not been developed (e.g. resource limitations, mismatched incentives), and suggest ways in which productive linkages might be encouraged.

Recommended Approach. By using both the case studies and the KVA survey it will be possible to document existing linkages.

Time and Resources. The case studies consume considerable resources but are central to the research because other components of the RVM methods are developed from the case study (e.g. identification of the KVA, gathering of CV's). The proposed case studies are not comparable to our previous case studies. These were more numerous and less intensive in terms of the level of detail and, particularly, the range of methods used. Considering only the site visit, narrative, and analysis aspects of the case studies suggested here, these new efforts will generally require four person days preparation, four

person days for the site visit, two person days follow-up, and ten to twenty person days for case study write up and analysis.

Replication. It is possible to use the same techniques we propose here to identify collaborative linkages in center-like research programs. The same interview questions and questionnaire items that serve here can serve in future program reviews or evaluations.

Question 3: (a) As currently configured, what can the NCPIR potentially achieve in research outputs and social impacts? (b) What alternative structures could contribute to additional accomplishments and increased productivity?

Part 3(a): Research Outputs and Social Impacts

Feasibility Assessment. The feasibility of this research question depends upon the definition of "social impacts." When reviewing this research question, one of the researchers we interviewed responded: "*This is the most important question—at the end of the day what did you do to improve the diagnosis and treatment of infertility? Has the science altered thinking about the pathology or physiology, has it provided information that would affect diagnosis? Has it improved the treatment? Has information found its way into application?*"

The above quotation seems to suggest a relatively stringent notion of "social impact": improvements in the quality of life of patients. But the respondent's view of his role in social impact is a relatively narrow one. The key term in the quotation is "altered thinking." This respondent, similar to others we interviewed, sees his role as creating "upstream" knowledge, knowledge potentially useful for clinical purposes. The presumption is that if useful knowledge is created, the roles of translation, application, development and use of the knowledge are the responsibility of others. The "altered thinking" to which this respondent refers is use of NCPIR-created knowledge by clinical researchers, especially those focusing on endocrinology, not by practicing, non-research physicians. If the use of knowledge by clinical research is construed as "the social impact," then it is worth pursuing this line of evaluation and it can be done straightforwardly. The interviewee above noted: "*…there will be examples where (our research has) supported a proposed therapy intervention."* This more limited definition of social impact is measurable, but is quite different from determining the much broader societal impacts that arise from the widespread use of new technology or treatments.

Another interviewee had a more straightforward view about measuring NCPIR's social impacts, a view not really inconsistent with the respondent above This second respondent observed: *"The social impact issue is tough. We disseminate in traditional reviews and research journals. We aren't community advocates, but just the first wave of bench to bedside* [investigators]. *Our audience is other scientists. We do not have the ability to do community work without additional money and maybe that's not appropriate for NIH, they are discovery oriented.* This is a different way of communicating the same

message as the first respondent: that the goal of the NCPIR centers is to affect other researchers.

It is certainly possible to determine "social impact" if what we mean by that is changes in the research community focusing on infertility. If so, this question becomes quite similar to Question 1, and can also be considered in connection with a KVA user survey.

Recommended Approach. We recommend a KVA user survey, as described above in Question 1.

Time and Resources. The time and resources will be essentially the same for all uses of the KVA user survey.

Replication. Survey questionnaires are highly amenable to replication. With just a few necessary revisions, it should be possible to use a questionnaire in perpetuity.

Part 3(b) Alternative Program Structures

Feasibility Assessment. While part (a) of this question targets outputs and impacts, part (b) is prospective about alternative programs structures and their relationships to potential research accomplishments. The approach here must be more interpretive and speculative. Nonetheless, we feel this may be an important part of the study, as it introduces alternative ways of thinking about organizing and managing research and research-related activities.

Recommended Approach. For a basic approach to this question, we recommend case study analysis and case study comparisons. The case study method provides ways to understand context, nuance, and unique features of organizations and institutions. For the NCPIR centers, case studies can help determine how the current center structures operate as well as assess their potential impact on scientific contributions.

Two approaches are envisioned here. First, the evaluators will rely on the NCPIR Advisory Board and NICHD staff to identify several interdisciplinary researcher centers at NIH that are most like the NCPIR centers. Case studies will be conducted on the "near comparison" centers to get a baseline that is as comparable as possible to the NCPIR centers.

To ensure variety in the research center comparisons, we will draw on the more than 50 case studies and "mini-case" studies of centers as a point of comparison for center structures and operations.[8] This type of research center comparison will not be "controlled" since the case studies will not be occurring simultaneously. The tasks envisioned in this approach include the following:

[8]See the Georgia Tech RVM Program website www.rvm.gatech.edu and the Rogers and Bozeman, 2001 paper for more information on prior case studies.

1. Conduct case studies of the Massachusetts General Hospital and the University of Pennsylvania infertility research centers, paying special attention to the relationship of organizational structure to performance.

2. Conduct case studies of several "most comparable" NIH centers (with "most comparable being determined by the evaluators in consultation with NICHD staff and the NCPIR Advisory Board).

3. Examine previous RVM case studies to identify range of variation and possible alternatives in program structure and operations.

4. Identify alternative program structures using the RVM case studies and the comparable NIH case studies, and make recommendations about the relevance of these structures for NCPIR organizational structure, design, and management.

First, the cases will be compared on the dimension of organization structure and design. In Appendix C, we provide examples of organizational models developed from case studies of the Oak Ridge National Laboratory and the University of California-Berkeley. These models can also be used to characterize organizational structure and design of the NCPIR centers and to compare their structures with other centers sponsored by the NIH, and by the Department of Energy and the National Science Foundation centers in the Georgia Tech RVM Program database of case studies.

Secondly, the NCPIR centers can be compared on the following dimensions:

- Funding and resource base;
- Scientific and technical output profile (composition of output in terms of articles, technology, technical assistance);
- Educational and human resource activities;
- Outreach activities; and
- Linkages to other entities in the NCPIR science community, or KVA (focusing on number, types and functions of linkages).

Time and Resources. The case studies can be performed with the existing team of researchers and in less than six months. The comparison phase should take no longer than one month additional.

Replication. From a strict scientific concept of replication, it will be difficult to replicate case studies. That is, there is no ability to provide scientific controls in case studies. However, case studies have a different sort of rigor. Among other uses, it is certainly possible to accumulate case studies that provide systematic comparisons. Indeed, this has been a key research and evaluation strategy that has been used throughout the history of the RVM program. By developing additional case studies, we have tools for continuing to understand the strengths and weaknesses of various organizational and management approaches over time.

Question 4: Does the NCPIR provide adequate opportunities for research fellows pursuing infertility studies with a human application in a supportive and collaborative setting that enhances their research productivity?

Feasibility Assessment. This question is amenable to evaluation and all our interviewees agreed that training is exactly what the programs were established to do, that is what they are doing, and that it can be demonstrated. The latter point—that it can be demonstrated—is quite true, but the chief means by which the NCPIR centers can provide evidence is only through anecdotes and individual case examples. That is appropriate, especially to the extent that it can be systematized. However, it is also useful to take a more systematic approach, one that can provide some quantitative evidence.

This question is fundamentally one pertaining to the capacity of the NCPIR centers to expand S&T human capital. One test is whether the research fellows and trainees supported by the NCPIR centers have been able to flourish in scientific productivity, have had important new career opportunities, and whether their knowledge and skills have been transported into a variety of new settings.

Recommended Approach. We recommend the use of "event history analysis" of CV's, supplemented by interviews. The RVM Program at Georgia Tech has pioneered the use of the CV as a systematic data instrument (e.g., Dietz, et al., 2000; Gaughan and Bozeman, 2001). The CV, provided it is standard, is an excellent means of charting careers and is easy to use with a variety of ancillary techniques such as citation analysis and network studies. The approach was designed specifically for S&T human capital issues similar to the ones implicit in this question. We have already determined that relatively standard CVs are available and that it will be possible to get information about the location of former fellows and trainees. The results will provide a quantitative profile of the career trajectories of fellows and trainees and the interviews will be useful in interpreting those career outcomes.

We suggest using a comparison group for the CV analysis. Determining the exact nature and size of the appropriate comparison group should await early stages of the research. After the focal CV's are collected it will be easier to determine the appropriate comparison group.

The chief limitation of using CV's for quantitative analysis is the method does not provide insight into motives, but this can be accommodated by using the approach in connection with interviews. Similarly, the concepts "adequacy" and "quality" cannot be addressed directly from analysis of CV's as these concepts suggest the need for some interpretation and individual judgment. Again, CV analysis must be connected to interview data, obtained not only from the persons who are trained but from persons who are responsible for training and those who employ or work with the trainees.

While a number of statistical approaches will be used in analyzing the CV data, one approach will be "event history analysis." Originally developed by demographers, event history analysis is a time series approach that determines "hazard rates" (likelihood of occurrence) for an event or set of events (e.g. tenure, receiving the first RO1, developing a patent) and how particular variables contribute to the hazard rate. These rates are especially useful for comparing persons who are not part of the same time cohort but who are subject to the same critical events. (For more information, see Allison, 1984).

Time and Resources. Gathering the CV data is straightforward, but coding it is quite time consuming, requiring 20-45 minutes per CV. However, this can be done by supervised and trained undergraduates, and thus, does not involve a major cost element. It is necessary, however, to have a graduate student available for the training and coordination. Given the relatively small number of CV's involved, this part of the study can be completed, including the interview elements, in about six-nine months, depending on the number of fellows and trainees analyzed and the size and nature of the comparison group. Currently, our expectation is that comparison CV's can be developed from individuals at the SCCPRR centers as well as from the clinician scientists supported through the Postdoctoral Individual National Research Service Award mechanism (F32) and the Clinical Investigator awards (K08).

Replication. This approach requires some training. But after the training it is easy to replicate the approach and it has the advantages of being almost entirely unobtrusive, employing readily available data. Very few evaluation techniques for assessing research programs can make this claim. The chief problem for replication is the time required for coding the CVs; however, the CV coding problem is potentially amenable to automation.

Question 5: How do NCPIR-supported research fellows compare with other translation- and clinician-scientists (who may or may not have training from other research centers) in terms of their contributions to infertility diagnostics and treatments, as well as their overall research productivity and career paths?

Feasibility Assessment. There is one major obstacle to evaluating the clinical-scientists with NCPIR-supported training to others: the number of subjects available for comparison is not high. We have not yet precisely measured the number of "clinician-scientists with NCPIR-supported training," but our rough estimate is that the number is less than 30. This implies that a comparison based on inferential statistics is probably not appropriate. More optimistically, due to the low number of trainees, it may be possible to develop information about the entire population (or a very high percentage of the entire population).

Recommended Approach. In addition to developing information about as much of the population as possible, it is also useful to identify appropriate comparison groups. We plan to consult with NICHD staff and the NCPIR Advisory Board, but we expect to draw useful comparison groups from the SCCPRR centers as well as from the clinician

scientists supported through the Postdoctoral Individual National Research Service Award mechanism (F32) and the Clinical Investigator awards (K08).

Time and Resources. In examining the CV's and career trajectories from these comparison groups and from NCPIR researchers and trainees, we provide a relatively unobtrusive and relatively low cost (except for coding) approach to analysis.

Replication. The ability to replicate these analyses are similar to question 4, the chief difference being the value of developing a data base that can serve as a continuing baseline as the number of fellows and trainees increases.

Question 6: What data, beyond that already being collected by the infertility research centers and the NCPIR staff, must be obtained to sustain future evaluations using similar methods? Can some of these data be collected and maintained relatively easy?

Feasibility Assessment. Currently, little systematic data is collected by the NCPIR centers, and thus, any sustained evaluation approach requires some attention to additional data gathering. One category of data can be collected and maintained with great ease: the CV data. Among the many advantages of using CV's for research evaluation, one of the most practical is that it is difficult to envision a means of directly gathering data from subjects that requires less time or fits better with the subjects' current record keeping practices and standards. But there are other data that could be useful for strategic analysis, planning and evaluation. During the evaluation, we shall focus on possibilities of developing data that are useful but, at the same time, minimally burdensome.

Recommended Approach. We recommend that any review of an NICHD centers program include use of CV analysis. To be sure, this will have the disadvantage of requiring relatively difficult coding, but the coding can be outsourced and is routine (at least after initial training). Survey questionnaires require somewhat more staff time and effort but seem, nonetheless, well within the bounds of reasonable expectation. Several other approaches to record keeping and data analysis may prove useful, but our recommendations on needed data are best provided at the end of the evaluation.

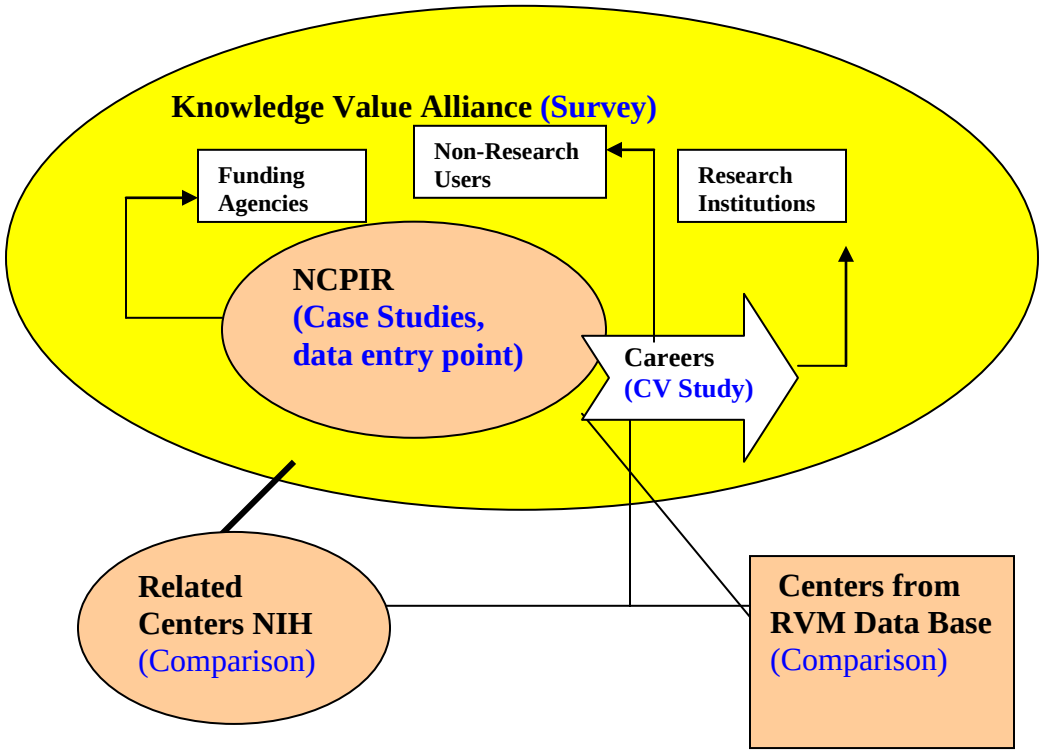
Summary: Relationships Among Recommended Approaches

Given the variety of evaluation methods recommended here, it seems useful to summarize their range and applicability. The focus of the S&T human capital study will be on the NCPIR centers themselves, as well as on the persons who have been trained at the NCPIR centers and gone on to other positions. The focus on the NCPIR science community (or KVA) user survey, the survey analysis, will be much broader. While the definition of the NCPIR science community will be identified as part of the evaluation, it is expected to include researchers and clinicians as the targets of NCPIR research, potential industrial users (e.g. pharmaceutical companies), and public and private funding agency personnel. Our study will be at several levels of analysis. For example, in the

case studies, the level of analysis will be the center itself. In the KVA study, we will work at both the individual and group level of analysis. The CV study will be at the individual level of analysis. Our study will seek to link the various levels of analysis to give multi-layered (individual, group, organization, institution, network) insights into the structure and performance of NCPIR's and the persons associated with these centers.

The inter-relationships among the various methods to be employed is diagrammed in Figure 1.

Figure 1. Research Approaches



Summary: Overall Time and Resources. While the list of tasks seems formidable, the researchers are prepared to perform all or some agreed upon set of them. To perform all the tasks above, the research period should be 1.5 years. It will be possible to compress the time period of the work by adding staff. We expect that performance of all tasks will require 5-6 person months of senior research time, six months for two half-time graduate students, and 6 months for two undergraduate students (chiefly involved in coding). While we have had considerable experience doing centers reviews for other clients, these are not an entirely valid gauge for the current study. Previous efforts have examined many more centers but in much less detail and with fewer methods.

Section V. Usefulness of RVM Concepts and Methods.

We expect the information generated in this application of RVM concepts and methods to the NCPIR centers to be of use to the NICHD leadership and staff and the NCPIR centers in a number of ways.

- By providing a view of performance, it will be possible to systematically document the scientific and technical impacts of the research produced at the NCPIR centers. In this respect, our evaluation resembles traditional output evaluations. A difference, of course, is that the output measures will be related to the question of developing research capacity.

- The research will provide an understanding of the role of NCPIR centers in building capacity in the fields of fertility and infertility research as well as related reproductive scientific fields.

- The CV-based studies will help NICHD staff and the NCPIR centers to understand the impacts of the NCPIR support for research career trajectories.

- The analysis of alternative organization structures and design for center program management should serve as resource for recommending improvements to the planning and management of fertility and infertility research.

- The surveys we conduct, based on the NCPIR science community as a KVA, will provide information about the uses and value of NCPIR research and knowledge products, as assessed by a wide array of potential users.

Table 1. NCPIR Review Evaluation Design

(1) Evaluation Question	(2) Variables/or Concepts	(3) Data Sources
1) Is the NCPIR serving as a "national resource" for the science community in infertility research?	- Indicators of use by stakeholders (KVA): citations, influence in research, influence in conceptualization of research, influence in technology development -Indicators of KVA awareness of NCPIR research and knowledge products -KVA respondents' assessments of quality of NCPIR research and knowledge products	-KVA survey questionnaires -Interviews from case studies -Documented instances of use of NCPIR research and knowledge products (e.g. awards, citations, licenses)
2) Are the NCPIR infertility research centers making linkages with the SCCPRR to facilitate translational research?	-Qualitative accounts of linkages (specifying types of linkages, duration, content and motives, network relations) -Sociometric models of linkage (if justified by number and intensity of linkage -Identifying obstacles to linkage, including resource limitations, role strain, incentive structures.	-Interviews -Case studies -Evidence from CV's (e.g. organizational linkages through career trajectories)
3a) As currently configured, what can the NCPIR potentially achieve in research capacity and scientific and technical human capital?	-Qualitative profiles of students, trainees, and "graduates" in terms of their career choices and productivity -Career trajectories -Quantitative measures of productivity in terms of publications, especially compared to comparable researchers,	-Case studies and interviews -CV analysis, comparison of CV data for NCPIR researchers and trainees in relation to F32, K08,and RO1 recipients not involved with NCPIR -Case analysis of scientific and technical human capital development

3b) What alternative structures could contribute to additional accomplishments and increased productivity?	-Chief dimensions for variables include formal organization design and structure (e.g. organization chart, formal authority relations), observed organizational structure and patterns of communication and authority (e.g. vertical differentiation, role specialization, communication and collaboration patterns)	-Case studies and interviews -Formal documents and records -Comparison to RVM cases database -CV analysis (especially in connection with assessment of educational activities)
4) Does the NCPIR provide adequate opportunities for clinician-scientist research fellows to pursue infertility studies with a human application in a supportive and collaborative setting that enhances their productivity?	-Patterns of collaboration (quality, periodicity, perceived opportunities, perceived effectiveness, rationales) -Productivity as measured by research outcome indicators (e.g. publication, citation, uses and applications)	-Case studies and interviews -Analysis of CV's and career trajectories
5) How do clinician-scientists with NCPIR-supported training compare with clinician-scientists (who may or may not have training from other research centers) in terms of their contributions to infertility diagnostics and treatments, as well as their overall productivity and career paths?	-Qualitative profiles of students, trainees, and "graduates" in terms of their career choices and productivity -Career trajectories -Quantitative measures of productivity in terms of publications, especially compared to comparable researchers, -Case analysis of scientific and technical human capital development	-Addressed by same methods as Question 4 above.
6) What data, beyond that already being collected by the infertility research centers and the NCPIR staff, must be obtained to sustain future evaluations using similar methods? Can some of these data be collected and maintained relatively easy?	-Indicators of center administrative data useful for strategic analysis, planning, and evaluation -Indicators of data available to measure scientific and technical capital	Case study analysis and interviews to assess feasibility of ongoing CV record keeping and data analysis procedures

Table 2. NCPIR Review Work Plan

1. Finalize research plan in consultation with NICHD staff and advisory board.
2. Arrange times and interviews for NCPIR case studies.
3. Develop interview protocol.
4. Develop data gathering protocol (documents, records curriculum vitae).
5. Conduct interviews at NCPIR centers and affiliates.
6. Identify population for "knowledge value alliance" (stakeholder) survey.
7. Acquire curriculum vitae (CVs) from NCPIR researchers and trainees on site, develop lists of former researchers and trainees, along with addresses and email addresses.
8. Choose several diverse NIH centers for comparative case studies.
9. Conduct comparative cases (repeating steps 5,7 above).
10. Select preliminary sample for KVA survey.
11. Develop KVA survey questionnaire.
12. Develop CV coding protocol (including inter-coder reliability approaches) and codebook.
13. Pre-test KVA questionnaire.
14. Examine results of KVA questionnaire pre-test for response patterns, response bias, selection effects; revise questionnaire.
15. Code CV's.
16. Write case study reports for five case studies (2 NCPIR centers and 3 comparative NIH centers).
17. Develop KVA survey codebook and database management approaches.
18. Finalize sample for KVA survey.
19. Send alert letters on KVA survey.
20. Send KVA questionnaire to sample.
21. Code questionnaire data.
22. Send follow-up reminder letter for KVA questionnaire.
23. Analyze CV data.
24. Analyze survey data.
25. Develop management analysis from case studies, focusing both on the five for current project and comparison with previous RVM case studies.
26. Develop management and institutional designs typologies in connection with the case studies.
27. Write results from CV analysis.
28. Write results from survey analysis.
29. Integrate three chief approaches to analysis, showing relationship among each and contributions to models and hypotheses.
30. Write draft final report.
31. Submit draft final report for comment.
32. Revise draft final report.
33. Present final report.
34. Disseminate results through publications and presentations.

Table 3. NCPIR Review Time Frame

Research Activity/Month	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Finalize research plan	x													
Conduct NCPIR case studies	x	x												
Identify population KVA stakeholders for Survey	x	x												
Acquire CV's	x	x	x	x										
Chose three NIH Centers for comparative case study		x												
Conduct comparative cases		x	x	x										
Develop KVA survey Questionnaire				x	x									
Develop CV coding protocol and codebook				x										
Code CV's				x	x	x	x	x	x	x				
Write case study reports					x	x	x	x	x					
Pre-test KVA survey Instrument						x								
Revise KVA survey Instrument and mail							x							
Develop survey codebook and database							x							
Code survey data							x	x	x	x				
Analyze CV data									x	x	x			
Analyze survey data										x	x	x		
Develop management analysis from case studies, both from project and previous data base					x	x	x	x	x	x	x	x		
Integrate analysis											x	x	x	
Write final report													x	x
Present final report and disseminate results														x

Table 4. NCPIR Review Resource Estimate

Core Activity	Professional Time Required	Other Resources Required and Cost	Comment
Conduct Case Studies of NCPIR's (2) and NIH Comparison Cases (3)	(This assumes a two-person team for each of the five cases) 20 person days preparation for five case studies, 20 person days for five site visits, 30 person days writing five cases, 10 person days for comparative analysis and integration of cases, 15 person days developing management typologies, interpreting and developing implications. **Total for five cases: 95 professional days.**	Travel, two persons for five cases, approximately $8,500 travel expenses for cases.	Almost all other research operations flow from the cases. The cases will be used to identify the population for the KVA survey, begin collection of CV's and for the analysis of the institutional design and performance of the centers
CV analysis	Collection time is during case studies (no additional time required); but 3 additional professional days devoted to acquiring data from persons no longer working at the centers. Coding of CV's will require 7 professional days supervision and database management, 12 non-professional days for data entry, 12 professional days for analysis of CV and comparison with existing CV data base,10 professional days for reporting the results of analysis. **Total: 32 professional days, 12 non-professional days.**	No other resources required excepting labor.	The major difficulty for this analysis is the coding and database creation.
KVA users survey	The case studies will be used for the initial identification of the KVA (the population of respondents) but 5 professional days will be required to finalize the population and draw a sample; 15 professional days required for survey instrument design; 10 professional days required for pre-test and redesign, 15 professional days required for analysis, 10 professional days required for reporting, 220 non-professional days required for survey administration, tracking, and data entry. **Total: 55 professional days, 220 non-professional days.**	Printing for mailed questionnaires, cost of mailing.	

Coordination with NICHD and Project Management (including integration of research and writing final report)	**Total: 20 professional days.**	Travel expenses for six person trips to NICHD, for person trips for consultant to travel to meet with other project researchers. Cost for production and distribution of final report.	
Grand Total	**Approximately 202 professional days; 232 non-professional days.**		

Appendix A. Research Value Mapping (RVM) Program

Georgia Tech's Research Value Mapping (RVM) program was formalized in 1996, building upon research evaluation work centered at Georgia Tech's School of Public Policy. The mission of the RVM program is to develop and apply innovative approaches to evaluating research and research-related activities, such as technology transfer. A primary objective includes the training of doctoral-level policy analysts in the skills of R&D evaluation.

The RVM program's early support was from the Department of Energy's (DOE) Office of Science. The Office of Science provided money for the developing and applying of new ways to evaluate public-sponsored R&D, especially basic research. The general RVM approach was developed in the early years of the program. That approach involved combining case study-based, qualitative methods with a variety of quantitative techniques (e.g. portfolio analysis). While the core DOE Office of Science funding continues, new sponsors have included the National Science Foundation (with two separate grants), a DOE grant to study contractor competition at the national laboratories, the Small Business Administration, and grants from Sandia National Laboratories to develop performance evaluation approaches. Recent program thrusts have focused on "capacity evaluation" and analyzing scientific and technical human capital, studies of scientific career trajectories, research on gender and productivity among university researchers, the impacts of grants on scientific careers, and scientific collaboration. Many RVM studies have been conducted with international partners.

The current director of the Research Value Mapping program is Juan Rogers. Barry Bozeman served as founding director of the program. The RVM program has supported more than forty graduate students and has included twelve faculty researchers and three postdoctoral researchers. The program is multidisciplinary and includes persons trained in political science, economics, engineering, sociology, environmental science, and management. Research supported by RVM has resulted in more than 100 publications in major journals in a wide variety of social science and engineering fields. RVM faculty are active as consultant and advisors to federal policy-makers, state technology-based economic development programs, as well as science policy managers in France, Canada, New Zealand, Japan, Argentina, Denmark, Sweden, and Finland.

In 2001, the program moved its operation from the Georgia Advanced Telecommunications Technology Building (off campus) to the D.M. Smith Building at Georgia Tech, integrating its operations with the School of Public Policy.

Appendix B: Capacity Evaluation and S&T Human Capital

S&T Human Capital: The Conceptual Starting Point for Capacity-Based Evaluation

S&T human capital takes stock of scientists' capacity—the amalgam of the inherent and social learning and the skills that permit scientists to both create and disseminate knowledge (Bozeman, Dietz and Gaughan, 2001). S&T human capital includes not only the formal educational endowments usually encompassed in traditional human capital concepts, but also the skills, know-how, "tacit knowledge," and experiential knowledge embodied in individual scientists. Thus, scientific and technical capital is the reservoir of knowledge, both technical and social, that scientists bring to their work.

Much of this capital, especially the aspect that is interpersonal and social, is embedded in social and professional networks. These networks integrate and shape scientific work, providing knowledge of scientists' and engineers' work activity, helping with job opportunities and job mobility, and indicating possible applications for scientific and technical work products. In these broader networks, S&T human capital includes actors in the technical enterprise who are users and developers of science and technology rather than creators themselves, as well as individuals in firms who appropriate knowledge and bring it to the marketplace. An important assumption of S&T human capital, one different from most concepts of scientists' and engineers' social contexts, is that the network in which knowledge is created must not be separated from the context of its use.

To summarize, S&T human capital is the sum total of scientific and technical and social knowledge and skills embodied in a particular individual. It is the unique set of resources that the individual brings to his or her own work and to collaborative efforts. Since the production of scientific knowledge is by definition social, many of the skills are more social or political than cognitive. Thus, knowledge of how to manage a team of junior researchers, post-docs and graduate students is part of S&T human capital. Knowledge of the expertise of other scientists (and their degree of willingness to share it) is part of S&T human capital. An increasingly important aspect of S&T human capital is knowledge of the workings of the funding institutions that may provide resources for one's work. Let us emphasize that none of this discounts the more traditional aspects of individual scientists' talents, such as the ability to conduct computer simulations of geological fracture patterns or the ability to draw from knowledge of surface chemistry to predict chemical reactions in new ceramic materials. The concept simply recognizes that, in modern science, being scientifically brilliant is only necessary, not sufficient. In most fields, a brilliant scientist who cannot recruit, work with, or communicate with colleagues or who cannot attract resources or manage them once obtained, is not a heroic figure but a tenure casualty or one or another variety of underachiever.

The S&T human capital framework for capacity evaluation assumes:

- Science, technology, innovation, and the commercial and social value produced by these activities depends upon the conjoining of equipment, material resources (including funding), organizational and institutional arrangements for work and the unique S&T human capital embodied in individuals.

- While the production function of groups is not purely an additive function of the S&T human capital and attendant non-unique elements (e.g., equipment), it closely resembles one. (The "missing ingredient" is how well the elements fit together and complement the production objectives at hand.)

- Most important, the S&T human capital model of effectiveness is: enhancing the ability and capacity of R&D groups and collectives to produce knowledge. Thus, the object of evaluation is best viewed in terms of capacity, not discrete products.

Modeling S&T Human Capital for Capacity Evaluation

Perhaps the best approach to fleshing out a S&T human capital model for research evaluation is to develop schematics for analysis. At the *individual* S&T human capital level, the model includes the endowments (human capital) of the individual researcher (cognitive, knowledge-based, skills-based) and the researcher's social ties, both direct and indirect (i.e., social capital). At the *project* S&T human capital level, the focus is on the aggregate of all project participants' endowments and social ties, as well as the physical and economic resources available to a project. Beyond the project and program levels, one may consider S&T human capital in virtually any social aggregation including a scientific field, sub-field, informal network or discipline. Each level of analysis is dynamic and functions in response to a set of distinct drivers and events. Evaluation of research programs may focus on either level, examining the capacity (i.e., S&T human capital + physical and economic resources) of the individual, the project (or a similar organizational unit), and the higher levels of social organization (e.g., knowledge value collective, network, scientific field).

Figure A provides a model of the individual's S&T human capital, showing, within the "box" (i.e., the individual researcher) unspecified dimensions of cognitive skills, scientific and technical substantive knowledge, and work-related or craft skills.

Let us begin by considering the "internal resources" of the scientist or technologist. To represent those internal resources, we have assumed that any individual's scientific capabilities can be classified into one or more of three, presumably, overlapping internal resources categories:

1. Cognitive skills.
2. Substantive scientific and technical knowledge.
3. Contextual skills.

The exact ways in which these capabilities relate to one another is, to us, an open empirical question, though studies in the psychology of science have begun to point the

way. (e.g., Prpic 1994; Simonton 1997). We are less concerned with the detailed specification of these internal resources than with their recognition as a component of S&T human capital.

Figure A depicts not only the internal resources of the scientist but also those external resources directly relevant to the production of knowledge and technology—social capital and embedded network ties. The different shapes of the nodes mean the convenience of recognizing qualitatively different types of linkages. Those differences may be based on the institutional setting of the network partner (e.g., industrial, academic) or the role (e.g., entrepreneur, funding agent, scientific colleague). We are less concerned at this point with the strength of ties or the density of networks, research on scientists and technologists and their networks (e.g., Meyer-Krahmer, 1998; Pickering and King, 1995). Scientists employ a wide variety of network-mediated resources to enable their work and these resources—this scientific, technical and commercial social capital—is uniquely configured for any particular scientist.

In Figure A, the broken line and shaded area represents the intersection of the research project with the individual's S&T social capital (network ties) and internal resources. The focus is on evaluation of research projects and, thus, the social organization compass points toward the ways in which the individual's S&T human capital tracks against the project's boundaries. But any social configuration can be mapped against the individual's S&T human capital to depict how the resources are deployed. A similar map could be drawn for a research program, a single research study, a laboratory or virtually any social organization or set of social interactions.

A unique aspect of the S&T human capital approach to program evaluation is that it recognizes how the careers of individual scientists evolve over time. Figure A represents a part of the individual scientist's productive life cycle, focusing specifically on a scientific project as a time anchor. In Figure A, at time *t-1* (pre-project) the individual, at least in this example, has fewer network ties and fewer dimensions of knowledge, skills, and craft. But in time *t+1* the individual has more dimensions of knowledge, skills, and craft, and a greater number of social ties. In this case, the task for the evaluator would be to determine the relationship between shifts in S&T human capital and participation in the project or program.

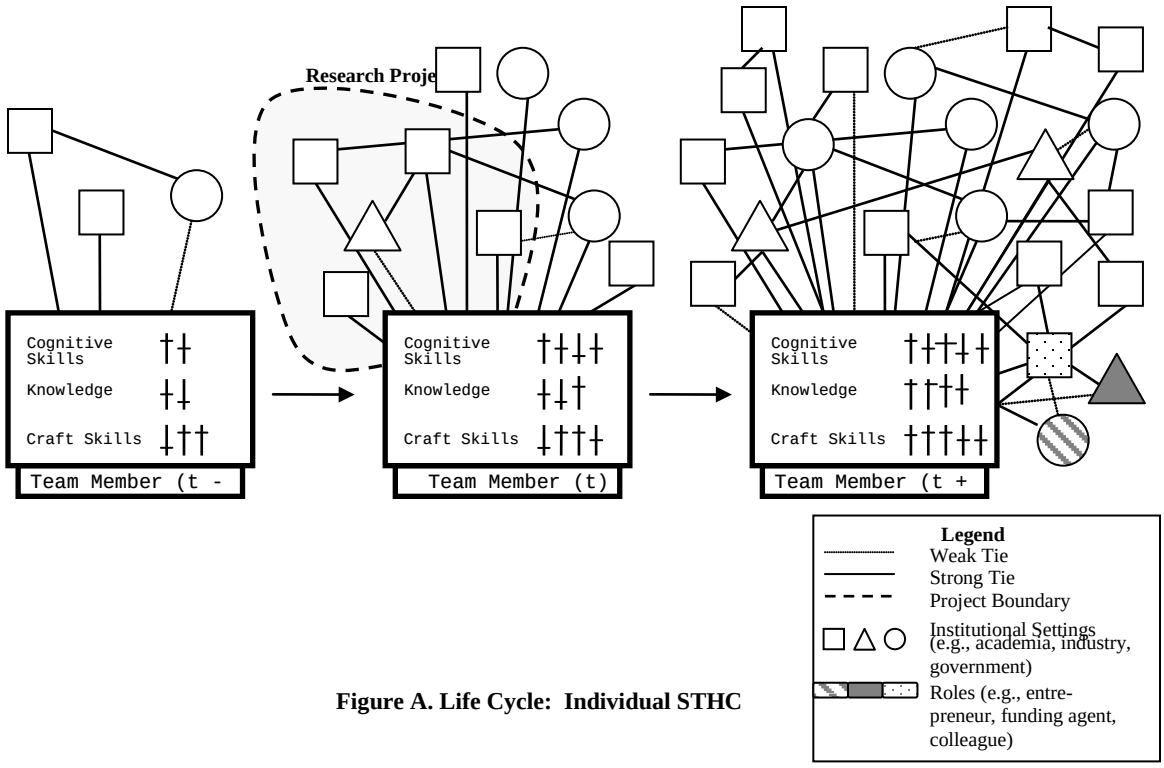


Figure A. Life Cycle: Individual STHC

S&T Human Capital as a "Capacity Evaluation Problem

The "evaluation problem" is to determine the extent to which a project or program has enhanced the S&T human capital of participants. As a result of the project, are the participants better able to contribute to future scientific and technical endeavors? Has their S&T human capital increased, has it increased in ways for which there is likely a future demand, and has it increased *because* of participation in the project or program?

One important implication of the S&T human capital model for capacity evaluation is its implications for program management. At the individual level, the management task is to properly assess the individual's S&T human capital and then to ensure that it is deployed in a way that maximizes the project's (organization's, program's) goals. Thus, the internal resource dimensions must be tapped and social resources must be exploited effectively. A beginning point, then, is a good knowledge of the individual's unique resources as represented in the S&T human capital model. Then more "generic" management activities become important—providing incentives, aligning individual and project goals, providing funds, equipment and other resources needed to fully exploit S&T human capital.

S&T Human Capital Evaluation: Data Sources and Measurement Issues

We have recently begun to apply a S&T human capital model for actual capacity-based evaluation (e.g., Dietz, et al., 2000; Gaughan and Bozeman, 2001) and multiple

evaluation methods are required with this approach. In the first place, case studies are important at the outset of capacity evaluation. RVM project case studies have underscored how much is missed by focusing on "the products" or even on sharp boundaries of projects. For example, in one case (Bozeman, 1998) breakthroughs in super-conducting materials are best accounted for by the ongoing relations between a team of multidisciplinary scientists held together by an entrepreneurial science manager. The development of management, political, and network-based skills in the project were just as important to its outcome as the educational or cognitive endowments of the parties involved. In another case, one involving development of state-of-the-art software (Bozeman and Gaughan, 1998), a work group's productivity could only be understood in terms of the entry/exit patterns of laboratory personnel and the specific talents gained and lost. These sorts of findings cannot be derived in traditional evaluation but, instead, require an approach to evaluation that: (1) is longitudinal, (2) examines networks or some other conceptual apparatus implying social connection, and (3) is capacity-oriented rather than product-oriented.

Case studies are only the beginning for a capacity evaluation. Among the research approaches we have used in previous studies is the scientist's CV or resume (see Dietz et. al., 2000; Gaughan and Bozeman, 2001). The CV is a reasonably standard means of recording career guideposts and accomplishments. It provides an excellent source of information pertinent to career trajectories and, when accompanied by probing questionnaires or interviews, can give an account of both the "what" questions and "why" questions as well. Most important, the CV is readily available. Many scientists provide them online in web pages, but even if not already publicly available, scientists customarily provide their ready-to-distribute CVs without any "tailoring" or additional burden beyond putting them in an envelope or an email attachment. Nor is their much need for tailoring in most cases. The information in a CV is exactly what one would wish in an analysis of S&T human capital and scientific careers. It may not be utterly complete. For example, it typically says little about the acquisition of tacit knowledge or about particular interactions with commercial users. But it is almost always a good starting point for gathering information about how the scientist has developed S&T human capital. It also provides some information relevant to network analysis, by listing collaborators and student advisees.

CV data generally are not adequate for "stand alone" evaluation. To measure capacity in projects, groups, networks, and knowledge value collectives, one must examine ties or linkages between the researchers. These are revealed to a limited extent though the unobtrusive measures of citation and patent analysis; however, many vitally important ties are not reflected in formally discernible collaboration patterns. Thus, interviews and questionnaires are generally an indispensable aspect of S&T human capital evaluation and have been employed by RVM researchers as a follow up to analysis of CV's (e.g., Bozeman, 2002).

Having outlined a S&T human capital model for evaluation, let us consider more systematically the ways in which it differs from related models. Figure B contrasts two models of evaluation and two models for the study of science. One may infer from this

table that the evaluation methods flowing from an S&T human capital model are not radically different from other approaches but their implications are.

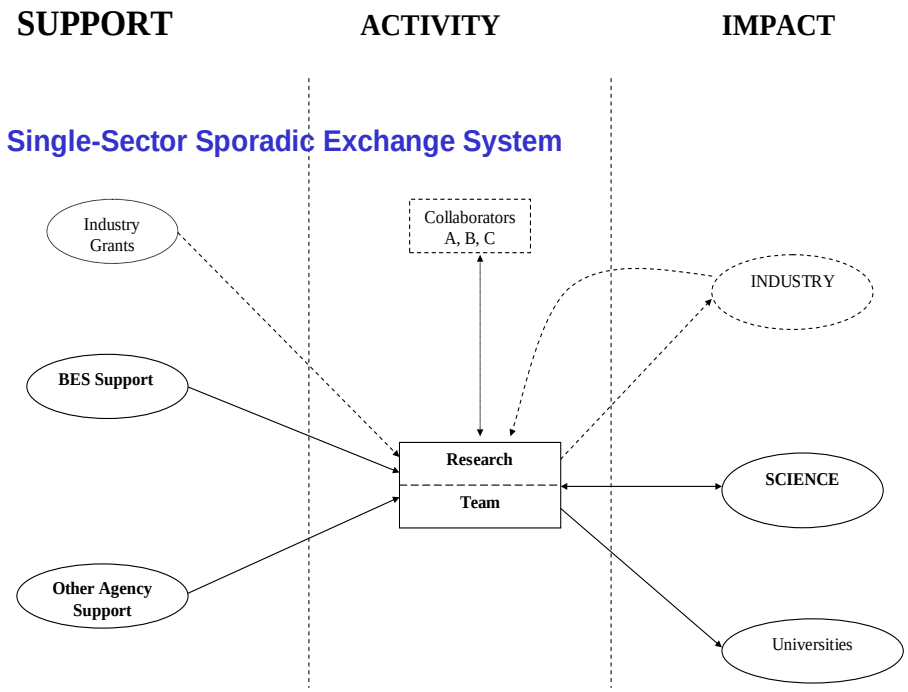
Figure B. Contrasting Models for Analysis of Scientific Productivity

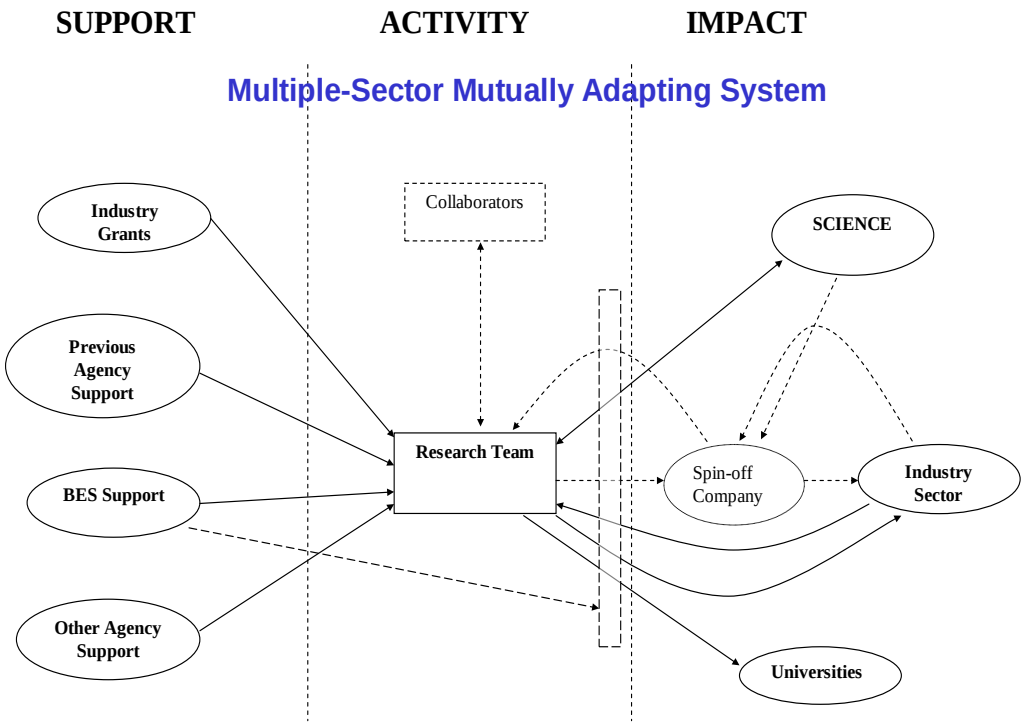
Character-istics of Model:	Individualistic: Human Capital	Individualistic: Output Evaluation	Social: Network Analysis of Scientists	Social: S&T human capital
Focal Dependent Variables	Income or status as derived from formal education and training	Scientific outputs such as publications, patents, algorithms	Citations, communication interactions	Productivity of (groups, collectives and networks) in terms of ability to produce knowledge and new applications of knowledge
Preferred Analytical Techniques	Econometrics/production function (e.g. Cobb-Douglas)	Varied, including benefit-cost analysis, case studies, citations	Sociometrics, citation analysis	Multiple, including sociometrics, citation analysis, supplemented by case studies and life course studies
Illustrative Studies	Becker	Kingsley, Bozeman and Coker (1995)	Allen (1977) Crane (1972)	Rogers and Bozeman (2001)
Chief Limitations	Limited applicability to post-Ph.D. scientists; income as productivity measure	De-emphasizes social fluidity of science, longitudinal component difficult to account for	Limited utility for policy evaluation, minimal institutional components, normative criteria often unclear	Costly; difficult to identify boundary rules
Chief Advantages	Precision, formalization	Conforms to policymakers' and managers' evaluation expectations	Provides good explanations of the social dynamics of science	Useful for capacity-oriented analysis, conforms to contemporary social milieu of science, enables longitudinal analysis

Appendix C. Types of Knowledge Value Alliances (KVA)

Presented below are two examples of the types of Knowledge Value Alliances (KVAs) identified by the Research Value Mapping Program at Georgia Tech (See Roger and Bozeman, 2001).

The first example is the **"Single-Sector Sporadic Exchange System,"** characteristic of mostly small university-based projects. The organizational structure is relatively simple and contains few inter-institutional relations. Yet, based on the research questions pursued by these KVAs, the structures can be complex and involve several scientific disciplines.





The second example of a KVA type, called the **"Multiple-Sector Mutually Adapting System,"** has a more complex organizational structure in contrast to the structure of the first example. A good illustration of this type KVA , found by Roger and Bozeman, is the R&D companies formed in the last decade to exploit the research results in such fields of molecular biology and software engineering.

Appendix D. "Lessons Learned" from the Feasibility Study

The feasibility study was not intended to make any conclusive inferences about the operations and contributions of the NCPIR centers. Nevertheless, we have some initial impressions and NICHD staff indicated that it would be useful, even at this preliminary point, to share some of the most fundamental "lessons learned." These are provided succinctly in this appendix, with the caveat that these are preliminary observations and may well be revised in light of further inquiry.

1. **The NCPIR centers focus chiefly on providing applied research rather than on linkage and translation.** When asked about linkages and translation work, respondents at both interview sites explained that their work activity focused chiefly on providing research to other researchers. Rather than linking directly to clinicians, the research output of the centers is, for the most part, either basic research or applied research aimed at clinical researchers.

2. **Within its niche, the NCPIR researchers are extremely productive.** By most any standard, the NCPIR researchers are highly productive researchers, producing both high quality and a considerable quantity of research relevant to fertility and infertility.

3. **Many of the NCPIR researchers can directly identify the contribution of NCPIR funding to their work.** In our previous studies of NSF and DOE science centers, we generally found that researchers had difficulty determining the contribution of particular funding sources to their research and technical productivity, chiefly due to the number of sources and the inter-relationships among projects. But many of the NCPIR researchers had specific lines of research that were supported largely or exclusively by NCPIR funds, and thus, could provide a better account of the impacts of NCPIR funding. Incidentally, we take no position on the desirability of "segregating" projects by research funding, we simply note that the ability to do so enhances valid evaluation.

4. **Compared to other research centers we have studied, the NCPIR centers have relatively simple, undifferentiated organizational structures.** Many of the centers we have previously studied had unusual or even unique organizational structures.[9] Both NCPIR centers seemed to us to have relatively limited direct interorganizational relations, generally with organizations similar to their own, simple hierarchical decision patterns, and the type of loose-coupling of work tasks found in most academic units. In a recent paper, we (Rogers and Bozeman, 2001) refer to this structure as "single-sector sporadic exchange system." This implies limited environmental exchange and when they do, it is chiefly within one sector). This is not surprising; the structure of the NCPIR centers seems to relate closely

[9] We use the term "organizational structure" in the sociological sense: that is, we are not referring to the organization chart but the formal structure but actual patterns of authority flow, communication, environmental relations, and resource development and allocation.

to their relatively small size (a fraction of most of the research centers we have examined) and their organization around principal investigators. Nor do we imply that the structure in inherently good or bad. Structure should conform to mission and effectiveness criteria. But we do note that if the NCPIR centers are to engage in more linkage activity, outreach, translation work, clinical services, or institutional partnering, their structures will need to change and, indeed, will inexorably change.

5. **The NCPIR centers currently collect little data that may be used in evaluation or planning.** Research organizations vary a great deal in the extent to which they collect or retain data useful for evaluation and planning. As far as we can determine, the NCPIR centers collect no data for this express purpose, but, of course, some of the data they collect for other purposes provides evaluation and planning opportunities. One reason for the informal approach to management and planning data is that the programs are small enough that the principals serve as the institutional memory. For example, we asked if records were kept on the placement of trainees and postdoctoral researchers and the response was that there were no such records, but none were needed since the numbers were small and all the cases could be recalled.

6. **The NCPIR centers are "lean" organizations.** The NCPIR centers resemble traditional academic departments in their low "administrative intensity" (defined in terms of the ratio of production [research] workers to support staff). Many of the DOE Engineering Research Centers and NSF Science Centers, almost all of which are much larger than the NCPIR centers, have considerable support staff. More administrative support will likely be required if the NCPIR's have even a modest change toward more translation work, outreach or inter-organizational partnering.

7. **The NCPIR centers are "soft money" operations.** We were not surprised at the extent to which NCPIR researchers are dependent on grants funding for their continued employment—that is common in medical and health research operations throughout the United States. We think it is worth noting that the situation is different at other types of research centers. It does not seem to us that the NCPIR researchers have a tentative or highly uncertain livelihood. Funding for their work is sufficiently available that, even though competitive, career survival on soft money is a viable option. However, the dependence on soft money, even when widely available, may have some potential (though we have not yet seen any evidence) of undermining research stability. If individuals are largely dependent on soft money they are infinitely malleable and will follow the available research dollars. To be sure, that may happen in the case of the NCPIR centers, in general, as persons interested in genetics find these centers' application-oriented, research agenda complementary with their own interests.

8. **NCPIR center directors and other research leaders have an appropriately strategic view of the NCPIR center funding.** Many of the individuals we interviewed showed a keen awareness of the role of stable center funding as opposed to RO1 funds. The center funding permits a great deal more flexibility and program stability than would be possible under research programs based entirely on R01 funds.

References

Allison, Paul D. (1984). *Event History Analysis: Regression for Longitudinal Event D*ata, Newbury Park, CA: Sage Publications.

Allen, T. (1977) *Managing the flow of technology*. Cambridge: MIT Press.

Becker, G. (1993) *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. University of Chicago Press.

Becker, G. (1962), 'Investment in human capital: A theoretical analysis,' *Journal of Political Economics* 70, No.5, pp.S9-S49.

Bobrowski, P. and S. Bretschneider, (1994) "Internal and external interorganizational relationships and their impact on the adoption of new technology: An exploratory study," *Technological Forecasting and Social Change*, Vol. 46, pp. 197-211 (1994).

Boix, C. and Posner, D.N. (1998), 'Social capital: explaining its origins and effects on government performance,' *British Journal of Political Science* 28, pp.686-693

Bozeman, B. et. al. (1998) *Research Value Mapping and Basic Research.* Final Report to the Department of Energy, Office of Science. Atlanta: Georgia Tech, Research Value Mapping Program.

Bozeman, B. and Melkers, J. (1993), " *Evaluating R&D impacts: Methods and practice*," Boston, Kluwer Academic Publishers.

Bozeman, B (2001) "A Capacity-Based Approach to R&D Evaluation," keynote speech to the International R&D Management Conference, Wellington, New Zealand, January.

Bozeman, Barry and Jim Dietz. "Strategic Research Partnerships: Constructing Policy-Relevant Indicators," *Journal of Technology Transfer,* 26, 4 (October, 2001), pp. 385-393.

Bozeman, B., Deitz, and Gaughan, (2001) "Scientific and technical human capital: an alternative model for research evaluation," International Journal of Technology Management, 22, 7/8, pp. 331-354.

Bozeman, B. and M. Gaughan. (1998) "Software Development at the Idaho National Engineering Laboratory." Atlanta: Research Value Mapping Program.

Bozeman, B. and M. Gaughan (under review) "Scientific and Technical Human Capital Accumulation: An Event History Model," *Research Evaluation.*

Bozeman, B. and M. Gaughan (2000) “Career Trajectories and Scientific and Technical Human Capital,” Invited paper, NSF Workshop on Research Evaluation, Bad Herrenalb, Germany, September 21-23.

Bozeman, B. and G. Kingsley, (1997) “The Research Value Mapping Approach to R&D Assessment,” *Journal of Technology Transfer,* 22, 2, 33-42.

Bozeman, B. and J. Rogers (2001) "Strategic Management of Government-Sponsored R&D Portfolios: Project Outputs and ‘Scientific and Technical Human Capital’" *Government and Policy [Environment and Planning C]*, Vol. 19, pp. 413-442.

Bozeman, B. and J. Rogers (in press), “A Churn Model of Scientific Knowledge Value: Internet Researchers as a Knowledge Value Collective,” *Research Policy* (in press).

Bozeman, B., J. Rogers, D. Roessner, H. Klein, and J. Park. (1996) *The R&D Value Mapping Project: Final Report*. Report to the Department of Energy, Office of Basic Energy Sciences. Atlanta, GA: Georgia Institute of Technology.

Bozeman, B. and et al. (1999), *The Research Value Mapping Project: Qualitative--quantitative case studies of research projects funded by the Office of Basic Energy Sciences. Final report to the Department of Energy, Office of Basic Energy Sciences.* Atlanta, Georgia Institute of Technology.

Callon, M. and Law, J. (1989), “On construction of sociotechnical networks: Content and context revisited,” *Knowledge and Society: Studies in the Sociology of Science Past and Present* 8, pp.57-85.

Callon, M. (1991), “Techno-economic networks and irreversibility,” In Law, J. ed., *A sociology of monsters: Essays on power, technology and domination*

Callon, M. (1992), “The dynamics of techno-economic networks,” In Coombs, R., Saviotti, P., and Walsh, V. eds., *Technological change and company strategies* London, Academic Press.

Coleman (1990) *Foundations of Social Theory.* Harvard University Press.

Coleman, J.C. (1988), “Social capital in the creation human capital,” *American Journal of Sociology* 94, pp.S95-S120.

Crane, D. (1972) *The Invisible College*. University of Chicago Press.

Crow, M. and B. Bozeman (1998) *Limited By Design: R&D Laboratories in the U.S. National Innovation System*. Columbia University Press.

Debackere, K., Clarysse, B., Wijnberg, N.M., and Rappa, M.A. (1994), “Science and Industry--A theory of networks and paradigms,” *Technology Analysis & Strategic Management* 6, No.1, pp.21-37.

DeBresson, C. and Amesse, F. (1991), “Networks of innovators: A review and introduction to the issue,” *Research Policy* 20, 123-130,

J. Dietz, I. Chompolov, B. Bozeman, E. Lane and J. Park (2000) “Using the C.V. as a Research Tool in Social Studies of Science,” *Scientometrics*, Vol. 49, No. 3, pp. 419-442.

Gaughan, M. and B. Bozeman (2001) “Impacts of Research Grants and Institutional Change on Scientists’ Careers: Comparing Center Funding with “Small Science” Grants.” Presented to the Annual Meeting of the Society for the Social Study of Science, Boston, MA, November.

Jordan, Gretchen. (2000) “Recognizing and Assessing Competing Values in S&T Organizations and Implications for Evaluation,” paper presented at “U.S. European Workshop on Learning from Science and Technology Policy Evaluation,” Bad Herrenalb, Germany, September 11-14.

Kingsley, G., B. Bozeman and K. Coker. (1996). “Technology Transfer and Absorption: An R&D Value Mapping Approach,” *Research Policy*, 25. 967-995.

Liyanage, S. (1995), “Breeding innovation clusters through collaborative research networks,” *Technovation* 15, No.9, pp.553-567.

Meyer-Krahmer, F. (1998), “Science-based technologies: University-industry interactions in four fields,” *Research Policy* 27, No.8, pp.835-851.

Moon, J. and S. Bretschneider, (1997)"Can State Government Action Affect Innovation and Diffusion," *Technology Forecasting and Social Change*. Vol. 54, No. 1.

Persson, O. and Beckmann, M. (1995), “Locating the network of interacting authors in scientific specialties,” *Scientometrics* 33, No.3, pp.351-366.

Pickering, J.M. and King, J.L. (1995), 'Hardwiring Weak Ties - Interorganizational Computer-Mediated Communication, Occupational Communities, And Organizational-Change,' *Organization Science* 6, No.4, pp.479-486.

Prpic, K. (1994), “The sociocognitive frameworks of scientific productivity,” *Scientometrics* 31, No.3, pp.293-311.

Quinn, R. and Kim Cameron , eds. (1988). *Paradox and Transformation: Toward a Theory of Change in Organization and Management.* Cambridge, MA: Ballinger.

Roy, U. (1997), “Economic growth with negative externalities in innovation,” *Journal of Macroeconomics* 19, No.1, pp.155-173.

Rogers, J. (2001), “Software’s ‘Functional Coding’ and Personnel Mobility in Technology Transfer: Linkage Fields between Industry and Publicly Funded Research,” *International Journal of Technology Management* Vol 22 No. 8.

Rogers, J. and B. Bozeman (2001), “Knowledge Value Alliances: An Alternative to R&D Project Evaluation,” *Science, Technology and Human Values*, Vol. 26, No. 1, pp. 23-55.

Rogers, J. Barry Bozeman and Ivan Chompalov (2001), “Obstacles and Opportunities in the Application of Network Analysis to the Evaluation of R&D," *Research Evaluation* Vol. 10, No. 3.

Simonton, D.K. (1992), “The social-context of career success and course fro 2,026 scientists and inventors,” *Personality and Social Psychology Bulletin* 18, No.4, pp.452-463.

Veum, J.R. (1999), “Training, wages, and the human capital model,” *Southern Economic Journal* 65, No.3, pp.526-538.