An Impact Evaluation Framework for Public-Private Collaborations on Research, Manufacturing, Supply Chain and Early Markets

Gretchen B Jordan, 360 Innovation LLC, Albuquerque New Mexico
Jeffery Dowd, U.S. Department of Energy, Washington, DC
Jonathon Mote, Southern Illinois University - Carbondale

Abstract

Many U.S. government programs are investing in public-private collaborations aimed at accelerating innovation, strengthening U.S. manufacturing and supply chains, and developing early markets for energy technologies. Evaluation of these government initiatives is both mandated and essential for developing an understanding of what works best and what the initiatives achieve over time. An inter-disciplinary expert team is developing a generic framework to guide independent impact and process assessments of this type of partnership initiative. This paper describes that framework. The framework covers the middle part of the product life cycle and interim progress along that cycle before the economic impact of full product adoption occurs. A theory of change model maps the relationships between initiative strategies and intended results, and suggests performance indicators. Guidance will be provided on the use of quasi-experimental study design within the relevant environments. The concepts of technology and manufacturing readiness levels provide common measures for assessing technology and manufacturing status. Netchain analysis (a synthesis of social network analysis and supply chain analysis) is used to assess the connections within and among the multiple networks in innovation: research, development, production, distribution and consumers. The resulting evaluation framework has application to any initiative in public-private collaboration where a similar approach to promotion of innovation and its adoption is taken.

Motivation and Introduction

Many U.S. government programs are investing in public-private collaborations aimed at accelerating innovation, strengthening U.S. manufacturing and supply chains, and developing early markets for energy technologies. There are at least two motivations for this: 1) the need to strengthen the U.S. global position with research in advanced energy technologies and restore the U.S. competitive edge in manufacturing; and 2) the recognition of the complex environment in which innovations and their adoption occur. Investments in public-private collaborations are underpinned by the tacit acknowledgement that quite often there are links between the needs and goals of the two sectors, and that collaborations have the potential to generate synergistic rewards.

U.S. federal government research and development (R&D) agencies are investing in these initiatives primarily to accelerate innovations. Some notable characteristics of such initiatives include 1) a clear goal to accelerate innovation; 2) attention to both R&D and advanced manufacturing and supply chain development, 3) a focus on pre/early-commercial or existing technologies where product, process and other upstream supply chain refinements can give U.S. manufacturers a competitive edge; and 4) time to commercialization and early market adoption of 1-5 years.

Formal evaluation of these types of initiatives is important. The U.S. Congressional and Executive branches expect accountability and demonstration of results, and have written this into legislation and budget expectations.[1] The implication is that programs must demonstrate results, such as the technical milestones and ultimate impacts (e.g., energy savings, reduced emissions, and economic benefits). They are also expected to provide evidence of early stage achievements, including in this case knowledge production, supply chain development, commercialization and early market uptake. Furthermore, evaluations should provide information to program managers on the processes that work to aid with improvements in program planning and implementation. This paper describes a framework for conducting impact and process assessments of Federal government R&D and related investments in public-private collaborations aimed at accelerated innovations. By identifying and properly elucidating the overall systems view of the complex process that attends accelerated innovations, the framework will facilitate the development of appropriate measures and indicators, and provide guidance for monitoring interim progress, while also identifying process improvement actions to guide timely adjustments to program activities.

Use of the generic framework to guide evaluation planning and implementation will lower the cost of evaluation planning, help ensure high quality evaluation, and increase the consistency and therefore credibility of evaluation findings. Further, the use of a consistent approach to evaluation of these initiatives means that it will be possible to synthesize across evaluations, opening up the real possibility of learning that goes beyond any one initiative.

Success in introduction of new advanced technologies involves complicated, non-linear journeys through several key phases, from R&D through commercialization to market uptake. This framework covers the middle part of the R&D product life cycle that existing frameworks on R&D Evaluation (Ruegg and Jordan 2007), Technology Deployment Impact (Reed and Jordan 2007a, 2007b) and Technology Benefit-Cost Evaluation (Ruegg and Jordan 2011) do not cover in depth.

The organization of this paper is to first provide an overview of the evaluation approach, followed by a discussion of the theory of change. The paper also includes a brief description of network analysis of supply chains, one of the more unique aspects of this framework. The major questions to be answered by the evaluation are described along with a preliminary menu of performance indicators.

Overview of Approach

The framework rests on a theory of change for public-private collaborations aimed at accelerating or increasing the degree of innovativeness of R&D and advanced manufacturing, and at making associated supply chains in early markets more complete, flexible, and robust. This theory of change is based on existing theories drawn from the academic and evaluation literature, expert advisory reports (Executive Office of the President 2009, 2011), and supply chain models such as that presented in Lowe et al. 2010. A theory of change is often summarized in a visual depiction of the building blocks and relationships between initiative strategies and intended outcome required to bring about that desired outcome.

This theory of change then forms the basis for evaluation planning for public-private collaborations with the objectives outlined in the introduction. The theory of change defines the system that is the scope of the evaluation. A specific collaboration may entail an intervention to influence just a portion of this theory of change, but evaluating the success of that intervention must still consider what happens in the entire system. The theory suggests both the larger questions to be answered by the evaluation as well as the specific questions that may be supported by a particular data collection method. It also suggests measures and indicators of progress and success. For example, the arguments

[1] See for example the GPRA Modernization Act of 2010, www.gpo.gov/fdsys/pkg/PLAW.../pdf/PLAW-111publ352.pdf

often made in support of "open innovation" stem from existing theories backed by some empirical evidence that innovation is spurred by interdisciplinary or cross functional or inter-organizational collaborations. A key question, then, is whether or not these collaborations are present. Social network theory further suggests characteristics of collaborations to look for and measure. Who is in the collaboration network, what is their motivation, do they have shared goals, what is their function (e.g., R&D, production), what is the structure of the network (such as dense or sparse) and what is the nature of their ties (strong or weak)?

Methods for data collection, analysis and reporting will be defined in some detail for the more unusual aspects of this framework. Network analysis of supply chains will be covered in some depth, as will theory of change evaluation, also called theory-based evaluation (Davidson 2000), which compares theory with actual observations. Quasi-experimental designs will be discussed to suggest ways to design a study, such as how to define control groups to compare to program participants when the size of market is small, along with the many, potentially confounding, events occurring in the system environment. The more often used methods of bibliometrics, document review, secondary statistics, and interviews will be discussed only briefly with references for detailed information on each. Another feature of the approach is that attention will be paid identifying effective ways to communicate the evaluation results to program stakeholders, Congress, and the public.

Planning for, and optimizing the likelihood of, a successful investment would hopefully be facilitated by the clarity enabled by the framework and the indicators developed from it for tracking progress. It bears emphasis that a government R&D agency is never the only source of influence on the process of accelerated innovations. Other factors outside the control of an R&D program exert considerable influence on developments in the R&D and the supply chain. These include overall economic conditions, competing or complementary policies, and competing technologies.

Measuring these types of changes within the theory of change enables programs to: 1) document early-stage impacts of investments which will not produce notable energy or environmental impacts for several years; and 2) subject to assessment of attribution, provide evidence the investments are creating desirable changes in network connectivity and accelerating movement through technology and market readiness stages. Measurement will also provide critical information on progress and process to enable a Federal R&D manager to determine if and when mid-course corrections are needed for improving the efficiency of investments.

The Proposed Theory of Change

Overview

The proposed program theory is shown in Figure 1. This generic theory of change is derived from theories in a number of different literatures, including science and innovation policy (Jordan 2010; Mote et al. 2007, Tassey 2007) and network analysis applied to R&D and supply chains (Borgatti and Lin 2009, Kim et al. 2011; Lazzarini, Chaddad and Cook 2001). One of the elements of this theory of change is the supply chain which includes suppliers of materials, components and sub-systems, production or assembly firms, distribution and retail firms, and consumers. Although a supply chain is the movement of hard goods such as parts and products, these companies also engage in the R&D and knowledge production that underlies that. Because this framework is focused on public-private collaborations among supply chain companies and publicly funded R&D, R&D is called out separately as a second element in Figure 1. Knowledge production includes multiple types of R&D. There is research associated with idea generation, design, production, commercialization, distribution, and marketing. This R&D is performed by many different organizations including industry, universities, national laboratories, research institutes, and research consortia.

Figure 1 shows areas of knowledge production, with fundamental understanding and proof of concept coming from basic and applied research; design and prototypes demonstrated at increasing scale in development research; ability to produce at desired volume, rate and cost (which includes refining products as signaled by the market) in manufacturing research; and knowledge of desired product and process characteristics from market research. Although these areas of knowledge production are shown as linear from left to right, the process is, of course, non-linear.

Publicly funded research and public-private R&D collaborations can be an integral part of the introduction of products and the development of supply chains for those products. In addition to funding the education of scientists and engineers and funding basic and applied research, government has an acknowledged role in establishing technical infrastructure that is used by all the private sector. Two examples are development of technical standards all must follow and provision of specialized facilities for testing and benchmarking that companies would not be able to afford otherwise. Research consortia often work on pre-commercial development issues of mutual interest. Public-private partnerships might engage in research at any point along the supply chain.

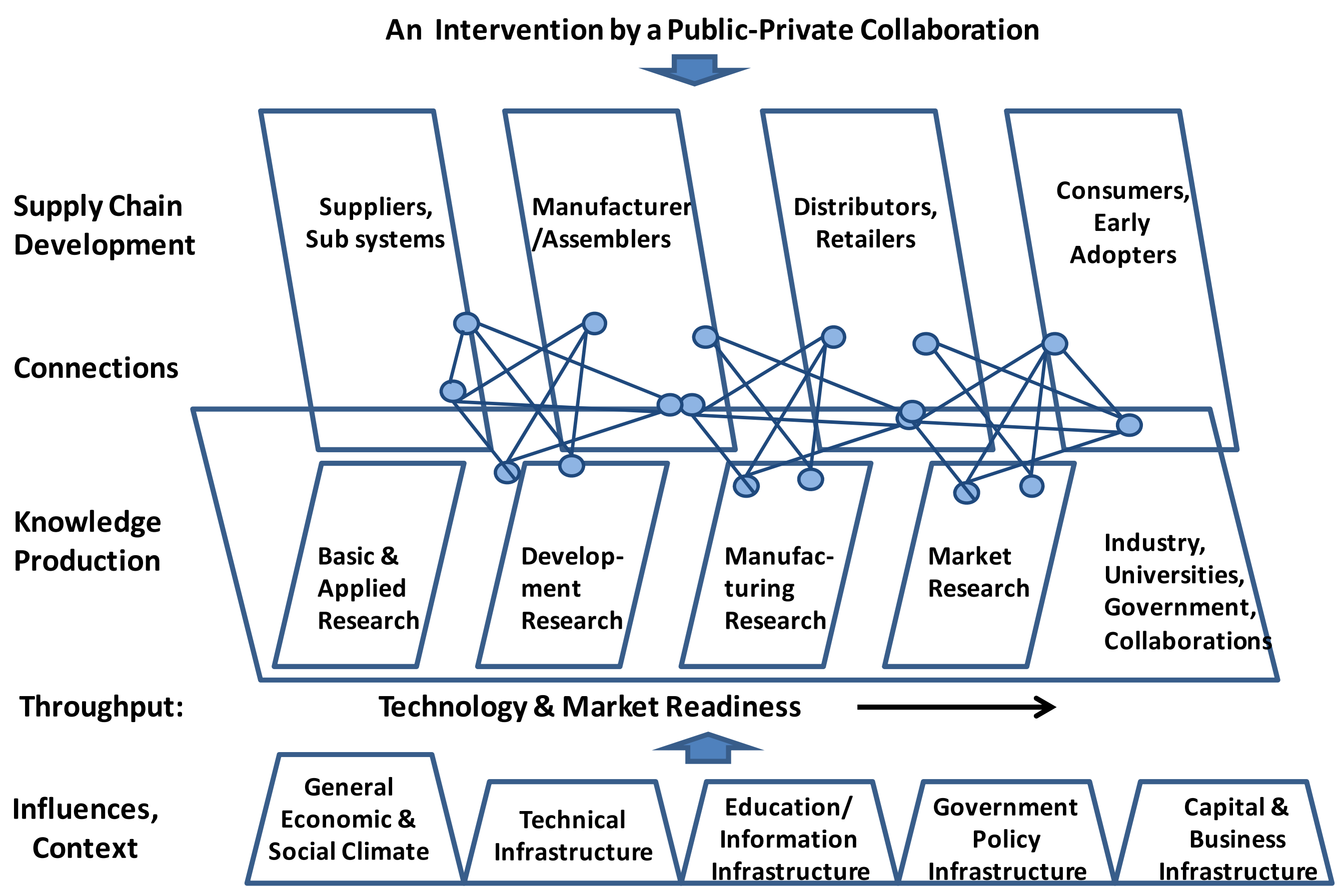

Figure 1. Theory of change for R&D collaborations to accelerate innovation along a supply chain

A third element of the theory of change is the myriad of connections between the other elements, knowledge production and the supply chain, as well as connections within each of those. It is worth noting that these same connections can be described as connections between the development of a new product or technology and development of the market in which it will be used. Theory suggests that connections among actors and their functions can create, accelerate, or expand innovation. Here "innovation" includes invention or adaptation of a product or process design and its development, commercialization, production, distribution and adoption. In Figure 1 connections are shown notionally as networks where nodes are dark circles representing companies or other institutions and lines between nodes show connections. For instance, suppliers bring parts to manufacturers, and research institutes bring product prototypes or manufacturing designs to manufacturers. The position of the notional network is intended to show that there are connecting relationships within a part of the supply chain or R&D, between parts of the supply chain, and between parts of the supply chain and areas of R&D. The extent and positioning of connections will differ for different situations and change over time as networks are primarily self forming.

The fourth set of elements in the theory of change is the general context and the specific infrastructure for any specific technology-market pair which influences the creation, acceleration, or expansion of the innovation. These influences may be positive or negative, that is, they may speed or hinder the innovation. Some will be barriers and solutions to barriers where public-private collaborations choose to engage. The others then are important aspects of the context for the intervention and its success, and as such are important alternative explanations of the success or lack of success of the collaborative intervention. The general context includes the economic climate and social and cultural norms. The specific context of the technology-market pair is grouped into four categories of infrastructure: technical, information/education, government, and business. These are consistent with typical descriptions of elements of innovation policy. The definition of each will become more clear in the table of indicators. This representation of technology-market infrastructure follows that used previously in an Impact Evaluation Framework (Reed and Jordan 2007) with modifications based on Greg Tassey's work (Tassey 2007) as incorporated into Jordan's innovation system model (Jordan 2010).

The theory of change, which covers initiative strategies for achieving intended outcomes, suggests four major areas of interim outcomes to be evaluated:

- Supply chain development (particularly domestic U.S.): Entry of firms in development, pre-production supply, production, and distribution, and enhancement of desired supply chain characteristics such as improving its robustness and flexibility.
- Connections: The extent and characteristics of connections within knowledge production, within the supply chain, and among these two.
- Knowledge production: Knowledge developed along the supply chain, which includes creation/adaptation, acceleration or expansion of inventions, designs, prototypes, validation, products or processes, and increased information about markets and aspects related to adoption of an innovation.
- Throughput: Progress of technology development for a particular application and operating environment and progress of that market toward producing, modifying as needed, and selling the technology. Throughput embodies the concepts of technology, manufacturing and market readiness levels.

Note that the focus of this framework is the middle part of the product life cycle, that is, higher technology readiness levels (TRL) beginning with product or process design (TRL 5-8), which overlap with manufacturing readiness levels (MRL 5-9) (Fernandez 2010). These levels would include early market development such as niche markets. TRL 9/MRL 10 represent more mature markets. These readiness levels are defined briefly in Figure 2.

TRL 5, MRL 5	TRL 6, MRL 6	TRL 7, MRL 7-8	TRL 8, MRL 9	TRL 9, MRL 10
Component validation –semi integrated; Produce in lab	System model demonstrated; Produce manufactured prototype	System prototype demonstrated; Pilot manf. line capability demonstrated	Actual system demonstrated; Limited production demonstrated	System proven in operations; Full rate production

Figure 2. Technology and Manufacturing Readiness Levels

Supply Chain and Network Analysis Considerations

The supply chain in the generic theory of change model uses a model of supply “netchains” proposed by Lazzarini et al. (2001). Supply chain studies speak of three networks that cut across the netchain. The movement of ideas, knowledge and information occurs in a development network (DVN) which is called Knowledge Production in this theory of change. The goal of a DVN is a new product design. Suppliers get involved in early supplier-involvement projects for new products, and often they collaborate with one another under the manufacturer’s leadership. Sometimes, manufacturers involve distributors for design issues so distributors can be prepared to innovate where needed. Consumers of course get involved in development to provide the voice of customer.

In addition to the development network a production network (PDN) exists upstream to the manufacturer/assembler (called the original equipment manufacturer or OEM). A PDN is made up of integrators and component manufacturers. There are multiple tiers of suppliers, not just the first-tier suppliers that deliver goods directly to OEM. There is also a distribution network (DBN) that lies downstream to the OEM, made up of distribution centers and retailers. There is a close link between upstream supply chain flexibility and downstream forecasting of level and location of consumer demand.

Flexibility is just one measure to be taken of a supply chain. Two others are robustness and effectiveness, which go hand in hand. A "robust" supply chain is one that can withstand changes. For example, if a company outsources everything to a top-tier supplier, things get simpler for that company as it has to manage just one supplier. However, that company does not have a robust supply chain because if that particular relationship ends it loses the entire supply chain. On the other hand, if a company takes a multi-tier approach and manages a select group of second and third-tier suppliers, that company has a more robust supply chain. Even if it loses its top-tier supplier, it still retains control over the key parts of its supply chain, that is, the supply chain is effectively delivering its product.

Effectiveness for our purposes focuses on both mobilization (building the number of actors working towards a common goal) and innovation (actors bringing in new ideas, processes, etc). Obviously, the number of actors is the initial measure, and more actors in this context is a good thing. Other measures can help understand what is happening in the network and the roles of individual actors. It is useful to know who are the central actors (hubs) and the brokers and the boundary spanners (actors who link different hubs together). From a network management perspective (to the extent government plays an active role in network management), this analysis is necessary to determine whether new actors need to be recruited for the network, as well as what new linkages might be needed.

Network structure is key. The primary driving force in emergent networks is “birds of a feather flock together”[2]. But proximity (in geographic terms) is also important, as in the saying “those close by, form a tie”. These two forces create dense networks, which is important for mobilization, but limit the

[2] "homophily"

ability for new information to enter. In addition, dense networks have several paths between any two nodes, which is important for maintaining information flows even when nodes or links are removed or leave the network. Important in a dense network are central actors with many direct connections to others that can disperse and direct the flow of information.

Diversity is also important, as it is advantageous that there be some actors that are not heavily embedded in the dense cluster. Actors that are loosely connected to the dense cluster allow for non-redundant ties and new information coming into the network. These diverse actors can be identified in two primary ways, as brokers (connecting disconnected parts of the network) or boundary spanners (connecting two or more clusters or connected to clusters outside of the main network). Over time, peripheral actors are drawn into the dense core due to the driving forces highlighted above. In general, efforts to facilitate networks must take into account these two forces.

Questions and Indicators Addressed in the Framework

Major Questions Addressed

Interim Impact questions

1. For the technology-market assessed, what has changed over time in (a) the technology, product or process, (b) the supply chain, (c) the formulated networks of researchers, manufacturers and supply chain agents, and market actors, and (d) elements of infrastructure and intervening factors? What are quantifiable movements in the indicators and measures?

2. What, if any, spillover impacts (not intended goals of the initiatives, such as use in another technology application) on R&D, supply chain and network members and others (e.g., non-partnership entities) have occurred?

3. What influence on the observed changes and spillovers can be attributed to the government R&D collaboration?

Process questions

1. What were the goals and strategies for achieving these (the theory of change) for the program(s) and to what extent was the program implemented as planned or modified? Does data show this theory or its modification were accurate?

2. How have networks formed and how effective have these networks been? Are some networks proven to be more effective than others and why? What program investments are working best to foster effective networks and integration of R&D and the supply chain?

3. What are the lessons about the influence of intervening factors, including factors that delayed or inhibited desirable change in technology and/or market readiness, or connections?

Interim Outcome, Progress and Process Indicators by Area of Impact

The major broad questions about the status of the desired interim outcomes, the program activities and outputs, and the mediating non-program influences on the outcomes, all must be specified, and measures identified for each aspect that the evaluation can assess. Some of these will have quantitative indicators while others will have qualitative indicators. Most often there will be multiple, partial indicators for a single variable. For example, performance of an automobile could include size, weight, turning radius, horse power, fuel efficiency, maintenance costs, and more. The theory of change guides the selection of indicators. Ideally by collecting data on what is expected according to the theory, as well as looking for other factors and explanations, the evaluation can build a case for both progress and program influence on that progress. Table 1 provides a partial menu of indicators.

Table 1. A Menu of Indicators by Question Area -- Specifics Vary by Program Focus

<table>
<tr><td colspan="3">Interim Results</td></tr>
<tr><td>Technology Readiness</td><td>Connections</td><td>Market Readiness</td></tr>
<tr><td>Development/production progress, validation/ demonstration;

Accelerated invention or adaptation, product design & commercialization (Time to commercialization)

Changes in product/process functionality & costs</td><td>Actors present in various functional areas; their relationships with each other, such as types of partnerships

Effectiveness of supply chain and its parts; robustness, flexibility

Public-private cost share</td><td>Production/distribution progress

Extent to which product characteristics meet desires of early adopters

"Optimization" of production & operations; Percent of supply chain that is domestic U.S. firms

Production volume; sales</td></tr>
<tr><td colspan="3">General context and specific technology-market infrastructure:
What exists, what changed, and extent to which it is favorable?
Education and workforce size and skill; Information: amount, accessibility, credibility; Government Policies (regulation, subsidies, procurement, etc.), Business Infrastructure (availability of venture capital, distribution channels, etc.), Technical Infrastructure (standards, test facilities, etc.)
Intervening factors: What did the government R&D program do? What did others do? Extent the government influenced progress and outcomes? Economic and social climate?</td></tr>
<tr><td colspan="3">Menu of progress and process indicators (levels and trends)</td></tr>
<tr><td>Level of funding for R&D

Movement thru TRLs, MRLs

Improved design processes; Patents, patent citations</td><td>Formation of partnerships, consortia, teams

Geographic clustering

Measures of network structure, nature of ties</td><td>Incentives for firms to enter; Existence of firms

Investment in components, manufacturing, distribution

Manufacturing cost, flexibility</td></tr>
</table>

Summary and Conclusions

Many U.S. government programs are investing in public-private collaborations aimed at accelerating innovation, strengthening U.S. manufacturing and supply chains, and developing early markets for energy technologies. Evaluation of these government initiatives is both mandated and essential for developing an understanding of what works best and what the initiatives achieve over time. To this end, the generic framework described in this paper can guide independent impact and process assessments of this type of partnership initiative. The framework covers the middle part of the product life cycle and interim progress along that cycle before the economic impact of full product adoption can be assessed. A theory of change model maps the relationships between initiative strategies and intended results, and suggests performance indicators. Guidance will be provided on the use of quasi-experimental study design within the relevant environments. The concepts of technology and manufacturing readiness levels provide common measures for assessing technology and manufacturing status. Netchain analysis, a synthesis of social network analysis and supply chain analysis, is used to assess the connections within and among the multiple networks in innovation: research, development, production, distribution and consumers.

Measuring these types of changes within R&D, the supply chain, and the R&D-supply chain networks enables programs to: 1) document early-stage impacts of investments which will not produce notable energy or environmental impacts for several years; and 2) subject to assessment of attribution, provide evidence the investments are creating desirable changes in network connectivity, and accelerating movement through technology and market readiness stages. Measurement will also provide critical information on progress and process to enable a Federal R&D manager to determine if and when mid-course corrections are needed for improving the efficiency of its investments.

The resulting evaluation framework has application to any initiatives in public-private collaboration where a similar approach to promotion of innovation and its adoption is taken. As this paper was written, the development of the framework was only partially complete. Comments and suggestions about the framework would be welcomed by the authors.

Acknowledgements

The authors acknowledge and appreciate advice provided by other members of the framework project team: Thomas Choi of Arizona State University, Rosalie Ruegg of TIA Consulting, Angela Becker-Dippmann of Pacific Northwest National Laboratory, and Yaw Agyeman of Lawrence Berkeley National Laboratory. Any errors and omissions are strictly those of the authors.

References

Borgatti, Stephen P and Xun Li. 2009. On Social Network Analysis in a Supply Chain Context. *Journal of Supply Chain Management*. Wiley Online Library. Volume: 45, Issue: 2, Pages: 5-22.

Davidson, E. Jane. 2000. "Ascertaining Causality in Theory-Based Evaluation", *New Directions for Evaluation*, No. 87, Fall, 17-26.

Executive Office of the President. 2009. *A Framework for Revitalizing American Manufacturing*, December. http://www.whitehouse.gov/sites/default/files/microsites/20091216-maunfacturing-framework.pdf

Executive Office of the President, President's Council of Advisors on Science and Technology. 2011. *Report to the President On Ensuring American Leadership in Advanced Manufacturing.* June. available from http://www1.eere.energy.gov/industry/amp/pcastreport.html

Fernandez, Joseph A. 2010. *Contextual Role of TRLs and MRLs in Technology Management*, Sandia Report SAND2010-7595, Sandia National Laboratories, Albuquerque, NM.

Jordan, Gretchen B. 2010. A Theory-Based Logic Model for Innovation Policy and Evaluation, *Research Evaluation*, 19(4), October 2010, 263-274.

Kim, Yusoon, Thomas Y Choi, Tingting Yan, and Kevin Dooley. 2011. "Structural investigation of supply networks: A social network analysis approach". *Journal of Operations Management,* 29: 194-211.

Lazzarini, Sedrgio G., Fabio R. Chaddad, and Michael L. Cook. 2001. Integrating supply chain and network analysis: The study of netchains. *Chain and network science*. Wageningen Publishers. Volume: 1, Issue: 1, Pages: 7-22.

Lowe, Marcy, Saori Tokuoka, Tali Trigg, and Gary Gereffi. 2010. *Lithium-ion Batteries for Electric Vehicles: The U.S. Value Chain*. Center on Globalization Governance & Competitiveness. October. http://www.cggc.duke.edu/whoweare/db_peopledetail.php?p=marcy_lowe.

Mote, Jonathon, G.B. Jordan, J. Hage, and Y. Whitestone. 2007. New Directions in the Use of Network Analysis in Research and Product Development Evaluation. *Research Evaluation*, Volume 16 Number 3 September.

Reed, John H. and Gretchen Jordan. 2007a. *Impact Evaluation Framework for Technology Deployment Programs*, U.S. DOE, July.

Reed, John H. and Gretchen Jordan. 2007b. Using Systems Theory and Logic Models to Define Integrated Outcomes and Performance Measures in Multi-program Settings, *Research Evaluation,* Volume 16 Number 3 September.

Ruegg, Rosalie, and Gretchen Jordan, 2007. *Overview of Evaluation Methods for R&D Programs*, U.S. Department of Energy, March.

Ruegg, Rosalie and Gretchen B. Jordan, 2011, *Guide for Conducting Benefit-Cost Evaluation of Realized Impacts of Public R&D Programs,* Prepared for U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Revised working draft, August.

Tassey, Greg. 2007. *The Technology Imperative*, Edward Elgar.