
1

Manufacturing Partnerships: Evaluation in the Context of Government Reform

**Philip Shapira,
Gordon Kingsley, and
Jan Youtie**

Philip Shapira is an Associate Professor in the School of Public Policy at Georgia Tech in Atlanta, Georgia. Gordon Kingsley is an Assistant Professor in the School of Public Policy at Georgia Tech. Jan Youtie is a Senior Research Associate in the Economic Development Institute at Georgia Tech.

Abstract

Current efforts to reform government have promoted the development of inter-governmental and inter-sectoral partnerships to implement public polices and more effectively deliver services. This paper probes the additional complexities posed for program evaluation by the development of partnerships, drawing on a study of the U.S. Manufacturing Extension Partnership. The highlighted challenges include diverse partner objectives, information deficiencies, variances in partner contributions and outcome effects, and difficulties in tracking a full set of benefits and costs, including learning benefits, institutional improvements, and transaction costs. Contrasting evaluation strategies to address the challenges of evaluating partnerships are considered.

Introduction

America's federalist form of government is intrinsically messy and disjointed: by both original design and subsequent development, various elements of authority, responsibility and oversight are divided between different levels and branches of government. At the same time, almost unceasing efforts are made to reshape the relationships between these constituent governmental components, to alter their roles, and to refocus or relocate power and control among national, state, and local units. We are presently in a period where such reform endeavors are exceptionally vigorous. From diverse points along the political spectrum, calls have been made to "reinvent" government (Gore, 1993), rewrite the "contract" as to the federal role (Gingrich & Armey 1994), and promote state and local "laboratories" of policy leadership and program innovation (Osborne, 1988).

One particularly prominent element of many recent reform proposals and initiatives is the promotion of “partnership” among different levels of government and between the public and private sectors (Gore 1993; Osborne and Gaebler, 1993). Partnership involves mutual cooperation between different agencies and organizations to find better ways to design and implement policies and programs. In itself, the concept of partnership is not new, and, in the modern era, there have been periodic efforts to increase the degree of inter-governmental and inter-sectoral collaboration in the United States (Graham, 1976; Jackson, 1995). As in the past, current partnership initiatives seek to improve coordination, increase efficiency and reduce duplication; they also hope to result in programs that more flexibly respond to local conditions and needs.

In addition to these “customary” motivations, fresh attention is being paid to the quality, value and outcomes of publicly-sponsored policy actions. Efforts to deliver services through partnerships are thus increasingly geared to more systematically combine the different capabilities of public and private organizations, leverage further resources and activities, stimulate shared learning and innovation, and result in measurable impacts for people, businesses and communities (Pilcher, 1991; Osborne and Gaebler, 1993).

New public-private partnerships ascribed with such characteristics have been promoted in fields as diverse as education and training, environmental protection, neighborhood revitalization, welfare reform, and technology development (Alliance for Redesigning Government, 1996). In the latter field alone, ten federal agencies joined with states and other local organizations to spend \$3.1 billion on hundreds of partnered technology programs in fiscal year 1994. Private sector participants were collaborators in most of these schemes, contributing further matching private resources (Coburn & Berglund, 1995).

As the scale and scope of government-backed partnerships increases, issues of evaluation arise. It is hard enough, in most cases, to evaluate single-agency programs, but how can evaluation take place in multi-organizational contexts where partners join together for sometimes quite varied reasons, can contribute in different ways, and have different organizational perspectives on how to judge success? Much of the existing literature on partnerships focuses on such factors as why organizations participate in inter-organizational relationships and how these new relationships are structured and evolve (Rogers and Mulford, 1982; O’Toole, 1989; Oliver, 1990; & Provan and Milward, 1991). How multi-organizational groups subsequently influence policy has also been studied (Mazmanian and Sabatier, 1989). But, there is less evidence about the outcomes of partnerships in achieving desired objectives of efficiency, resource leverage, and economic impact. In the latest round of partnership promotion, a body of anecdotal case study material is emerging (see, for example, Alliance for Redesigning Government, 1996). Often it is assumed that partnerships self-evidently work, otherwise they would break up and other organizational forms would emerge. This is a tendency found in transaction cost theories of organizations, which seek economically rational reasons for the development of organizational forms (Moe, 1990). However, in the public policy arena, institutional arrangements are promoted and protected for a variety of reasons (including ones of political as well as economic rationality). We thus should not assume, on the face of it and without further investigation, that partnerships, however fashionable, will result in improved and more effective delivery of needed services.

In this study, we examine the dilemmas of evaluating partnerships through a program that has incorporated many of the design principles of the latest wave of inter-governmental thinking - the Manufacturing

Extension Partnership (MEP). Sponsored by the National Institute of Standards and Technology of the U.S. Department of Commerce, this program supports the cooperative delivery of technology assistance to small and mid-sized manufacturers. Using findings and observations from a recent assessment of collaborative service delivery in the MEP, the paper examines the complexities of assessing effectiveness and impacts in a decentralized, collaborative program. We begin with a discussion of the development of the MEP and how this program has been shaped by the deliberate promotion of partnership and collaboration. This is followed by our observations on the effects of partnership promotion on service delivery and how this approach to service delivery poses additional difficulties for evaluation. We then discuss possible strategies for addressing these challenges.

The MEP: Development, Organizational Arrangements and Service Delivery

The Manufacturing Extension Partnership (MEP) is a network of technology assistance and service providers which aims to upgrade the performance and competitiveness of small and medium-sized manufacturers in the United States.¹ MEP is a collaborative initiative between federal and state governments which also involves non-profit organizations, academic institutions, and industry groups. The National Institute of Standards and Technology of the U.S. Department of Commerce is the MEP's federal sponsor (National Institute of Standards and Technology, 1996).

The provision of technology and related business assistance to small and mid-size firms in the United States is not a new activity. In the 1950s and 1960s, state industrial extension and technology assistance programs were established in several states, including Georgia, North

Carolina, and Pennsylvania. These early programs diffused technical information and used professional engineers and other technical specialists to help local firms improve their use of technology (Shapira, 1990). Increased concerns about industrial competitiveness and regional economic development over the next two decades prompted the development of similar programs in other states. By 1990, manufacturing extension and technology transfer programs had been established in 28 states (National Governors' Association, 1990 & 1991). However, despite this growth, most programs had relatively few resources, services were patchily available, and a number of states had no programs at all. A review at that time found that only 13 of the state programs had field staff who were able to work on-site with firms - a factor critical in being able to address specific manufacturing problems on the shop floor of small firms (National Institute of Standards and Technology, 1990).

The federal government started its own direct support of industrial extension programs in the late 1980s and early 1990s. Under the auspices of the 1988 Omnibus Trade and Competitiveness Act, the National Institute of Standards and Technology set up a handful of manufacturing technology centers, working in an initial collaboration with selected states (U.S. Congress, 1990; National Research Council, 1993). But the most dramatic expansion of federal sponsorship came with the Technology Reinvestment Program, first announced in 1993 and implemented in 1994 and 1995 (Advanced Research Projects Agency, 1993). Although this multi-billion dollar program was targeted towards the post-cold war conversion and restructuring of America's defense-industrial base, it did make significant funds available for the upgrading and deployment of technology in civilian industries, including small and mid-sized manufacturers. The National Institute of Standards and Technology was allocated a share of these resources to increase the

number of manufacturing technology centers and to organize, through what was then the embryonic Manufacturing Extension Partnership, a collaboratively-delivered set of industrial extension services to manufacturers throughout the country.

The Technology Reinvestment Program had a major impact on the structure and character of the MEP. Funding was awarded through a competitive review process and states and other applicants generally had to match one-half of proposed costs with their own or private funds. Additionally, applicants were guided to form partnerships of service providers. There was an explicit requirement that proposals for funding address a criterion entitled "Coordination and Elimination of Duplication." This criterion required the proposer to understand and link with related service providers in the service region, be consistent with existing state strategies, and not duplicate existing resources or services. Proposers' partnerships were judged in terms of the number, diversity, and skills of constituent service providers, geographic scope and coverage, cohesiveness, organization, and management structure (National Institute of Standards and Technology, 1994).

Boosted by the Technology Reinvestment Project and with subsequent support from the civilian budget through the U.S. Department of Commerce, the MEP has now grown to a network of more than 72 centers in 50 states. A variety of partnership elements are fostered. The federal sponsor works with states and local centers in program management and development and in backing joint working groups, staff training, information and communications systems, and common tools. Here, MEP's federal funding of nearly \$90 million in fiscal 1996 is supplemented at least a further \$100 million in (mostly) state and (some) private funds. In almost every case, individual centers operate with and through local networks of associated service providers. Across the MEP system, there are more than

750 formally-affiliated organizations which deliver or support the delivery of services to small and medium-sized manufacturers, and many more which do so informally or on a flexible basis. These organizations are quite varied in that they include non-profit technology or business assistance centers, economic development groups, universities and community colleges, private consultants, utilities, federal laboratories, and industry associations. The average center has 19 organizational affiliations, although some centers have only a few such relationships, whereas at least one center has more than 100 (Shapira & Youtie, 1996). Generally all of the centers have governing or advisory boards which include local public- and private-sector representatives.

The MEP thus incorporates many of the design principles articulated in government reform proposals and advanced by advocates of new governmental partnerships. First, the programs seeks a cooperative relationship between the public and private sectors. The private sector is involved not just as a recipient, but also as a service partner and an advisor. Second, the program is decentralized and flexible, with individual centers able to develop strategies and program services which are appropriate to state and local conditions. Third, MEP seeks not to duplicate existing resources. Rather than provide services directly from the federal level, MEP awards are designed to get existing service providers, whether they be consulting firms, non-profit organizations, academic institutions, public agencies or trade associations, to cooperate and coordinate in their efforts to assist local manufacturers.

Collaborative Service Delivery and Evaluation Complexity

In aggregate, the MEP has encouraged the formation of partnerships linking different sectors, levels of government, and service providers. Of course, this raises a further and most critical

set of questions about how these partnerships operate, now that they have been established, and how they can be evaluated. In this section, we address these issues, drawing upon detailed case studies of partnership and service coordination for six centers within the MEP system (Shapira and Youtie, 1996). All of the centers studied had extensive partnership arrangements; they were chosen, with the aid of an external review panel, to represent organizations with among the most well-developed partnership arrangements in the MEP system. With one exception, the centers pre-dated the Technology Reinvestment Program. All centers had used the resources of the Technology Reinvestment Program and the MEP to establish or expand their partnership linkages.²

The case studies, which were undertaken in 1995, involved a series of on-site interviews with MEP program managers, staff, service providers, small business customers, and state program sponsors. These were supplemented by reviews of program documents from the centers and their affiliates and information obtained through the MEP reporting system. The topics we investigated included the history and organizational development of extension services in the case study area, the role of the Technology Reinvestment Project and the Manufacturing Extension Partnership in promoting change in service arrangements, and the process of selecting, organizing and managing different service providers and relationships. We examined planning systems, financial and contractual arrangements, marketing and outreach methods, and partnership coordination, including joint projects, sharing of program resources, referrals, and personnel training. Finally, we explored measures and examples of program impact and probed how the effectiveness of partnership linkages and the center as a whole was being judged.

Our study confirmed that the Technology Reinvestment Project and the MEP had stimulated state and local service

providers to work together in more coherent ways than they had previously. This was mainly due to the stimulation and support from the federal level. Indeed, we did not find that state governments consistently required providers of manufacturing assistance within their jurisdiction to coordinate their efforts. Moreover, while some of the pre-existing case study centers had, at their own initiative, developed a wide range of informal (and in some instances, formal) alliances prior to affiliation with the MEP, it really took the intervention of the MEP and the incentive of funding by the Technology Reinvestment Program for serious attention to be paid to forming multi-organizational partnerships. We also found that enhanced service coordination through local partnerships between different service providers made available a wider range of expertise to firms and, in many instances, a more systematic approach to providing assistance. Involving other partners allowed MEP centers to maintain flexibility and centers to expand their services quickly by leveraging existing resources.

While we discerned that federal efforts to promote partnerships had, in the case of the MEP and with the incentive of available federal funds, so resulted in partnerships being formed, this is not an ultimate indicator of success. Obviously, we need to know whether desired outcome objectives were achieved, and at what cost. Our study, which focused on how the MEP had affected processes of partnership formation and development, did not (and could not) analyze final impacts for firms and local economies. However, we did gain an understanding of how partnership modes of service delivery made efforts to conduct comprehensive evaluations more complex. We identified several issues here, as the following section describes.

Diversity in partnership objectives

In classic text-book evaluation designs, the aim is to match objectives with

results. Yet, we found that the sponsors, service providers, and customers of particular service partnerships often had quite different objectives and so varied in how they judged success. Among the sponsors, state and federal aims in co-funding partnerships did not necessarily coincide. A principal federal interest in promoting manufacturing extension is the improvement of industrial competitiveness and productivity for the nation as a whole. States, on the other hand, are particularly concerned about what happens within their own boundaries (and rightfully so). They prize new jobs (even though enhanced productivity may reduce employment in some instances) and have few qualms about supporting “their” businesses, even at the cost of jobs in another state. (For example, by assisting low-wage small firms, some extension efforts may help the shift of high-wage jobs from larger manufacturers in other states - a process which may be zero-sum or worse from a national perspective.) Similarly, from the viewpoint of individual partner organizations and assisted firms, each has different stakeholders and interests which may or may not consider MEP evaluative metrics as important benchmarks of performance or outcomes.

Diversity in objectives is, of course, found in many policies and programs whether operated through partnerships or not. However, it is apparent that as partnerships grow, so does the multiplicity of stakeholders and differential interests, thereby adding to the complexity of what and whose objectives evaluation should measure. Moreover, it was also apparent to us that some partners were added to bolster local political support, head off potential “turf battles” and augment outreach efforts, rather than provide distinct service delivery capabilities. Evaluative results may show that these partners contribute little on standard measures, but underestimating the role of these partners could impair a federal sponsor’s ability to maintain ongoing in-state

political support for inter-governmental partnership initiatives such as MEP.

Disparate external impacts

Efforts to develop standardized evaluation metrics among manufacturing extension programs have emphasized business outcomes, such as increased sales or reduced costs, and economic development impacts, particularly effects on jobs (Oldsman, 1996; Shapira, Youtie, & Roessner, 1996). But differences in the specialized services that various partners offer present challenges to the use of narrowly fixed measures across the whole partnership. Our study noted the diverse organizations involved in MEP programs, including universities, colleges, economic development organizations, technology centers, federal labs, utilities, and different kinds of private sector consultants. Some partners offered services that were more likely to have quantifiable effects than others. Some produced short-term gains, whereas others promised deeper, lasting effects on an industry or regional economy. For example, partners which provided human resource and training services or helped firms develop supplier linkages and inter-industry networks frequently had less tangible effects on a firm’s bottom-line than organizations which showed firms how to make fast savings in energy consumption or fix a particular machine problem. The risk here is that partners whose activities have deeper long-term impacts may be unfavorably compared to service providers whose results are more readily measurable now.

In a sense, there is a tension that goes beyond differences in partner roles and to the basic expectations ascribed to partnerships themselves. On the one hand, partnerships are promoted as innovative mechanisms, able to address structural, long-running problems and needs better than previously fragmented approaches and services. On the other hand, partnerships are also expected to efficiently and quickly

produce visible results. There are undeniably possibilities for partnerships to simultaneously make progress on both of these fronts, but realistically, as we found in our cases, tensions may be apparent between what can be measured now and what is of long-term value.

Information base for evaluation

At the center of many program evaluation systems is a method of tracking program activity, customer contacts, and services delivered to measure the nature of the program intervention. Separate measures can then sought on the impact of these interventions, in this case on the firms who are the programs customers (ideally compared with non-customer controls). But information and service tracking in a multi-organizational partnership is not easily accomplished nor standardized. Our study found that partnerships tracked information with varying degrees of intensity and with dissimilar means. One center had established an electronic information system shared among the program's 17 partner organizations. But in most of the other centers, individual partners maintained their own tracking systems suited to the needs of their principal sponsors, not necessarily the partnership. Moreover, these partnerships used informal, if any, means of communicating program activity information among the partners. For example, although all the centers referred manufacturing customers to organizations within or outside of the partnership where appropriate, none of the centers had systematic information about customers' valuation of referrals across the program.

Additionally, although programs reported service counts ("outputs" - in that they add-up items like the number of firms served, but do not fully capture the "outcomes" of those services) and financial data to the national MEP program, none of the centers studied had or were in the process of implementing their own evaluation plans

across all organizations in the partnership. From a pragmatic perspective, multi-organizational information systems and evaluations were expensive; in the context of limited public resources, their payback may not appear worthwhile, certainly from the view of many state sponsors who are frequently satisfied with output (rather than outcome) measures and program-solicited company testimonials.

Accounting for internal learning benefits

Partnerships are, or at least should be, learning organizations which adjust to change and promote improvement among constituent service organizations. In fact, in our MEP cases, we did observe ongoing changes in affiliations and operations. Some centers reduced or dropped individual partners that were ineffective, although others still had difficulty in dealing with less active partners. In a few cases, organizational linkages which were inactive for years became more central to the center's service delivery strategy as manufacturing needs changed. Other examples were observed where MEP centers worked with "problem" partners to help them improve their performance. This kind of adjustment and learning is inevitable, and indeed very desirable, as partnerships develop. Yet funders often encourage the partnerships they sponsor to propose that they already have all the strategic "pieces" in place, rather than admit - in writing - that some of those pieces have weaknesses.

Verbally, however, most MEP managers openly discussed the realities of their partner links, their strengths and weaknesses, and how they saw those links evolving. In the early years of these partnerships, considerable attention was paid to mechanisms for creating awareness and understanding between organizations such as various forms of meetings, presentations, and cross-training. This is followed by an ongoing process of learning and partner adjustment, in response to changes in

industrial needs, institutional capabilities, personnel, and funding. While there are “churning” costs here (as we note below), there is also a great deal of internal institutional learning and improvement, part of which is reflected in the improved quality and seamlessness of services we observed being provided to firms and part of which goes to improvements in the organization itself. Evaluation impact measures which purely emphasized outcomes would not fully capture the often intangible and qualitative benefits of such internal learning.

Accounting for internal program costs

While the MEP cases showed definite benefits from increased partnership coordination, it was also apparent that these improvements had some direct (i.e., MEP) and indirect (i.e., non-MEP) costs. For example, MEP programs actively engaged in service coordination incurred various transactions costs, including the expense of information exchange, contracting, consulting, and monitoring. In some cases, the minimization of inter-organizational tensions required the expenditure of “political” capital. It was also noted that while there was a reduction of inefficient program overlaps in several cases, the other program resources “leveraged” by MEP centers were not “free” in that they had to be paid for by other public or private sources. It is possible to estimate some of these costs, but again others may be intangible. Our main point, though, is to emphasize that partnerships are not a “free ride” and that evaluation has to find ways to consider the full set of costs incurred in their development and operation, as well as a full set of benefits.

Strategies for Evaluating Partnerships

The central purposes of evaluation for partnerships are, of course, no different than for all other kinds of program evaluations. These include developing an understanding of program operations and impacts, guiding the activities of program managers, influencing external policy decisions, testing whether resources and services are being appropriately committed, and providing feedback opportunities to learn about the elements the program seeks to effect (Goldenberg, 1983; Feller, Glasmeier, and Mark, 1996). Similarly, evaluations of partnerships have to face the “normal” problems facing all program evaluations, which include deriving suitable operational measures, determining relationships between interventions and outcomes, and determining the general applicability and reliability of any findings (Sylvia, Meier, & Gunn, 1991). But, in addition, we suggest that partnerships pose extra challenges and hazards for evaluation. While most of these can be construed to fit within standard frameworks of the internal and external threats to evaluation designs, it is useful to highlight them separately. As seen in the partnerships we studied, these challenges revolve around diverse partner objectives, information deficiencies, variances in partner contributions and outcome effects, and difficulties in tracking a full set of benefits and costs, including learning benefits, institutional improvements, and transaction costs.

Why do these extra challenges arise? In part, they are a function of the greater organizational complexity inevitably associated with partnerships. Multiple organizations across different levels and sectors bring diverse interests and institutional structures into play; chains of command are diffused and intervention points increase. It becomes more difficult to

identify standard goals against which to benchmark program performance, common metrics, and even a set of program services which do not change over the course of the evaluation period. The challenges perhaps also reflect the management style of the program's sponsor, which in the situation we studied encouraged a considerable degree of flexibility, experimentation, and innovation as to how partnerships were put together and operated. Finally, there are the tensions embodied in the current concept of partnership itself, where sometimes hard to reconcile expectations are combined together in a single promotional package.³

These challenges notwithstanding, programs such as the MEP must still meet demands to demonstrate their effectiveness and provide policy-relevant information to guide future development, including resource allocations and service approaches. How might this be done? One strategy could be to encourage each constituent organization within a partnership to develop and administer separate evaluation plans. The individual organizations would develop their own metrics based on their stakeholders' demands and the characteristics of their own services. These metrics would be implemented with surveys and other methods to assess service outcomes. A "benefit" of this strategy is that partner organizations are not bound to standard performance metrics.

However, there are several difficulties with this approach. First, performing organizational-specific evaluations creates system-wide inefficiencies. Each organization must pay for the development of an evaluation plan, including conceptualization of performance measures, administration of information systems and data collection, analysis, and reporting. The demand for evaluating performance may not be high enough to support such a costly way to demonstrate results.

Second, having individual organizations conduct evaluations imposes burdens on program customers, in this case

small and medium-sized manufacturers. For example, one customer could get different, multiple information requests from the various organizations participating in a joint project. The more these customers are burdened with estimating the results of services provided, the more costly to the firm those services become. These firms at some point will react against this, either through information non-response or program non-participation. Regardless of which organization is taxing these customers, the overall partnership will likely get the blame and customer relationships will be damaged.

Third, some organizational partners may have more excess funds and internal capabilities for conducting evaluations than others. Wide variations in the depth and methodological quality of existing evaluation approaches will occur and produce different results, making it difficult for the managing partner and the partnership sponsors to obtain and present credible and reliable information (Shapira, Youtie & Roessner, 1996).

An approach at the other end of the spectrum is to centralize program evaluation responsibilities with a managing partner (the MEP center), a principal sponsor (the lead federal agency, in the MEP context), or an independent oversight agency. If the managing partner took responsibility for evaluation, including evaluation plan development, data collection, analysis, and reporting, it would reduce the variation of individual partner evaluations, generate intra-organizational efficiencies, minimize the potential for over-burdening customers, and increase the resource-base for evaluation. However, our findings indicate that this strategy is difficult to implement within a multi-organizational environment. It does not recognize the different demands of partner organization stakeholders. It is hard to adjust a centralized evaluation plan to constant changes in partnership composition. And different goals and visions among partner organizations cannot be taken into account. If partner organizations have to

bend to the requirements and burdens of a centrally administered evaluation function, they may be less willing to participate in the partnership. The evaluation could potentially damage the relationships between the center and its partner organizations, and important leveraging and service quality benefits could be lost. Moreover, while partnership-level evaluations can provide important internal information for program learning and improvement, there are likely to be great differences between the approaches and qualities of evaluations between different partnerships. This would make national comparisons less reliable.

Raising the level of evaluation to yet a higher level, that of the federal sponsor, will almost certainly lead to greater standardization in evaluation measures and will probably allow the commitment of more resources to evaluation. National resources are especially useful in assembling valid comparative information and databases of unserved firms to provide controls. The federal sponsor can undertake aspects of evaluation itself or support independent outside professionals to independently verify program results. Also, federal oversight agencies can, and indeed do, conduct evaluations, using their own resources (see, for example, U.S. General Accounting Office, 1995). One issue with oversight reviews, however, is that they tend to be broad in scope and relatively infrequent (for example, the General Accounting Office has only undertaken two reviews of the MEP and its federal predecessors since 1988) and thus do not assist in the “real-time” management of the program (although they have an important role in the continuance of federal funding).

Interestingly, some manufacturing extension managers are strong advocates of a standard, national evaluation approach: in part, they seek definite targets to “aim” at; there is also desire to have the responsibility for evaluation lifted off their shoulders. Other managers, however, greatly resist the idea of having to match up to central

measures (Shapira, Youtie, and Roessner, 1996). Particularly relevant to partnerships, there are concerns that standardized measures run up against the desired flexibility of local programs and have a “chilling effect” on experimentation, risk-taking, and the customization of services to regional circumstances and needs. That said, and as an inescapable by-product of receiving federal money, the MEP partnership programs are subject to a standardized set of measures related to service inputs and outputs (including numbers of projects undertaken, individuals trained, and fee revenues generated) and project outcomes (including changes in business sales, costs, investment, and jobs). It is our observation that such measures do not unduly squash the flexibility found in local programs, but nor does the data collected appear to have a great influence on the fortunes of individual programs. In the periodic external reviews that local MEP programs have to undergo to receive continued federal funding, more attention is often paid to issues of management, organization, strategy, partnership composition, and finance than to quantitative impact measures.⁴

Of the evaluation approaches so far discussed, all have applicability and, indeed, all are practically applied in one way or another. Yet there is still a sense in which in more is not necessarily better, given the weaknesses and incompatibilities already noted. This is not simply a question of methodology, for example the choice between quantitative or qualitative measures. Rather, it reflects the structural constraints felt by any one element of a partnership when trying either to evaluate itself in isolation or to assess the partnership as a whole. However, there is also a further possibility, which is to allocate responsibility for evaluation not to any particular partner, level, or sponsor but to the system as a whole. This involves the promotion of discursive processes whereby program sponsors, service providers, customers, and other stakeholders are

encouraged to engage with one another to assess and review program practices, performance standards, the results of studies, and best practices.

The institutional architecture of this system-wide approach to evaluation has been likened, by Sabel (1996) to the continuous improvement user groups and quality benchmarking methods now increasingly applied in the private sector. The role of program sponsors, at both local and national levels, is to support mechanisms whereby within-program and between-program reviews and comparisons can be made, to stimulate stakeholders to set their own benchmarks based on such comparisons, and to assist in the dissemination of program learning, analysis, and best practice. It is an approach which is particularly supportive of the decentralized, innovative, learning, and self-improvement features expected of new inter-governmental partnerships. It, of course, runs up against many entrenched aspects of hierarchical program management and budgetary politics. Nonetheless, these kinds of evaluative and learning processes are at work both within the individual MEP partnerships and at the system level as a whole, where there are now many working groups, professional meetings, reviews, and other methods of information exchange and comparison. To a significant degree, discursive evaluation is an inevitable and healthy by-product of partnership promotion, as different interests and groups build new relationships and compare and learn from one another. But the support of key program sponsors and managers is also critical, if the learning so obtained is to be institutionalized and generalized.

Conclusions

In the context of current efforts to reform public policy implementation through the promotion of inter-governmental and inter-sectoral partnerships, this paper has

probed how such initiatives increase the complexity of evaluation though a study of the U.S. Manufacturing Extension Partnership. We noted that the MEP's emphasis on collaborative local relationships among different local service providers and sectors did not necessarily require a common program goal agreement, led to diverse and sometimes non-comparable services being provided, and posed additional challenges for tracking service information and accounting for a full set of costs and benefits. Conventional strategies of program evaluation, which allocated responsibilities to a single constituent program stakeholder or to an external oversight agency, each had particular strengths and weaknesses in dealing with partnerships. We suggested, however, that system-wide approaches, involving varied partnership stakeholders in discursive yet structured comparative evaluation processes, had a particularly important role.

Where does this leave the program evaluator? We propose two principal sets of implications. First, it is clear that there is a growing need for robust evaluations of partnership initiatives. The early evidence, as our case studies of the MEP showed, is positive as to their impacts on service delivery and quality for customers and users. Yet, this is not conclusive, and more rigorous and longer-term analyses are needed, particularly using a range of outcome measures. As such studies are constructed, the findings from our cases would indicate that evaluators should develop a good understanding of the particular characteristics of the partnerships they are examining, so as to become attuned to the mix of different objectives, services, and activities being undertaken. Outcome measures will be influenced by this review. In addition, efforts to discover and fully assess partnership costs and benefits, including learning benefits, institutional changes, transactional costs, and modifications in service delivery are important. While quantitative assessments

will be vital, we suspect that carefully structured case studies will prove particularly useful in uncovering these aspects, determining the partnership's net contribution, and linking partnership activities to customer outcomes (along the lines of the approach advocated by Yin, 1994).

The second implication is rather different, for it suggests a modification, if not a change, in the program evaluator's traditional perspective. In promoting structured comparative measures and processes, evaluators can assist in facilitation - in helping partnerships document and define what they do, better understand how their actions lead to program and customer impacts, and assist in discerning viable program improvements. The key change here is not so much in analytical techniques, but in the key questions posed. Conventional program evaluation efforts tend to be geared towards program justification and to external interests, focusing inquiries on such questions as: Does partnership intervention produce results and are these results worthy of continued funding? The need for this type of evaluation remains, as we have noted. But the new partnership initiatives prompt further questions geared towards ongoing program improvement and innovation, for example: How does this partnership produce results? What practices will lead to further gains in results achieved relative to resources and needs? And, how can those practices be implemented? The answers here are ones that program managers, partnership members and customers will find very relevant. They will also be most useful in the broader debate about how best to reshape the delivery mechanisms of public policy.

Acknowledgment

This paper draws upon research supported by the U.S. Department of Commerce, National Institute of Standards

and Technology (NIST) under award 50SBNBC8305. The views expressed are the authors' and do not necessarily reflect those of the research sponsor.

Endnotes

1. Small and medium-sized manufacturers are commonly defined as industrial firms with fewer than 500 employees. There are about 415,000 small and medium-sized manufacturing enterprises in the United States, representing 99 percent of all manufacturing enterprises and almost two-fifths of U.S. manufacturing jobs (Small Business Administration, 1992). Evidence about the technology and industrial challenges facing small and mid-sized manufacturers is not reviewed here, but can be found in several recent studies, for example: U.S. Congress, Office of Technology Assessment, 1990; National Research Council, 1993; Organization for Economic Cooperation and Development, 1993; and the Modernization Forum, 1994.

2. The six partnership cases comprising the study (with their service areas) were the Chicago Manufacturing Center (Chicago, Illinois area), the Georgia Manufacturing Extension Alliance (state of Georgia), the Great Lakes Manufacturing Technology Center (Cleveland, Ohio area), the Manufacturing Extension Partnership of Southwest Pennsylvania (Pittsburgh, Pennsylvania area), the Minnesota Manufacturing Technology Center (state of Minnesota), and Oklahoma Alliance for Manufacturing Excellence (state of Oklahoma). For a full discussion of these cases, see Shapira and Youtie (1996).

3. One further challenge, which is particular to the MEP, is the issue of controls. The program is now active in more than four-fifths of the states, and includes almost all of the major state-level manufacturing extension service providers as partners. It is thus hard to compare the

performance of partnership-based extension services with non-partnership based ones (the latter is now a small and diminishing group as the MEP continues to expand). Of course, it is still possible to conduct more important controlled studies of the performance of MEP assisted firms compared with those who have not been assisted, and several such studies have been undertaken (for references, see Shapira, Youtie, and Roessner, 1996).

4. One of the authors of this article has been a panelist on seven external federal reviews of manufacturing extension centers or partnerships and a participant in two further cases where programs were reviewed.

References

Advanced Research Projects Agency (1993). *Technology Reinvestment Project*. Program Information Package. Arlington, VA: U.S. Department of Defense.

Alliance for Redesigning Government (1996). *Public innovator learning network*. Washington, DC: National Academy of Public Administration (World-Wide Web site <http://www.clearlake.ibm.com/Alliance/>).

Coburn, C. & Berglund, D. (1995). *Partnerships: A compendium of state and federal cooperative technology programs*. Columbus, OH: Batelle Memorial Institute.

Feller, I., Glasmeier, A., & Mark, M. (1996). Issues and perspectives on evaluating manufacturing modernization programs. *Research Policy*, 25, 309-319.

Gingrich, N. & Armey, D. (1994). *Contract with America: The bold plan by Rep. Newt Gingrich, Rep. Dick Armey, & the House Republicans to change the nation*. New York: Random.

Goldenberg, E. (1993). The three faces of evaluation. *Journal of Policy Analysis and Management*, 2, 515-525.

Graham, O.L. Jr. (1976). *Toward a planned society: From Roosevelt to Nixon*. New York: Oxford University Press.

Gore, A. (1993). *Creating a government that works better and costs less: A report of the National Performance Review*. New York: Plume.

Jackson, G. (1995) *Government partnerships*. Washington, DC: National Academy of Public Administration (electronic document available at <http://192.156.133.18/Alliance/clusters/ig/overview.11.html>).

Mazmanian, D. & Sabatier, P.A. (1989). *Implementation and public policy*. Lanham, MD: University Press of America, Inc.

Modernization Forum (1994). *The role and importance of the Manufacturing Extension Partnership centers*. Dearborn, MI: The Modernization Forum.

Moe, T. (1990). The politics of structural choice: Toward a theory of public bureaucracy. In O. Williamson (ed.) *Organization Theory*. New York: Oxford University Press.

National Research Council (1993). *Learning to change: Opportunities to improve the performance of smaller manufacturers*. Commission on Engineering and Technical Systems, Manufacturing Studies Board. Washington, DC: National Academy Press.

National Governors' Association (1990). *Promoting technological excellence: The role of state and federal extension activities*. Washington, DC: National Governors' Association.

National Governors' Association (1991). *Increasing the competitiveness of America's*

manufacturers: A review of state industrial extension programs. Washington, DC: National Governors' Association.

National Institute of Standards and Technology (1990). *Technology-based economic development: A study of state and federal technical extension services.* Special Publication 786. Gaithersburg, MD: National Institute of Standards and Technology.

National Institute of Standards and Technology (1994). Technology Reinvestment Project deployment activity areas: Lessons learned. Materials presented at workshops in Oakland, CA and Atlanta, GA, March 28, 1994. Gaithersburg, MD: National Institute of Standards and Technology.

National Institute of Standards and Technology (1996). *Manufacturing Extension Partnership Source.*TM Gaithersburg, MD: National Institute of Standards and Technology (World-Wide Web site <http://www.mep.nist.gov/>).

Oldsman, E. (1996). Does manufacturing extension matter? An evaluation of the Industrial Technology Extension Service in New York. *Research Policy*, 25(2), 215-232.

Oliver, C. (1990). Determinants of interorganizational relationships: Integration and future direction. *Academy of Management Review*, 15, 2, 241-255.

O'Toole, L. Jr. (1989). Goal multiplicity in the implementation setting: Subtle impacts and the case of waste-water privatization. *Policy Studies Journal*, 18, 1, 1-20.

Organization for Economic Cooperation and Development (1993). *Small and medium-sized enterprises: Technology and competitiveness.* Paris.

Osborne, D. (1988). *Laboratories of democracy.* Boston: Harvard Business School Press.

Osborne, D. & Gaebler, T. (1993). *Reinventing government: How the entrepreneurial spirit is transforming the public sector.* New York: Plume.

Pilcher, D. (1991). The third wave of economic development. *State Legislatures*, November 1991, 34-37.

Provan, K.G. and Milward, H.B. (1991). Institutional-level norms and organizational involvement in a service-implementation network. *Journal of Public Administration Research and Theory*, 1, 4, 391-417.

Rogers, D.L. and Mulford, C.L. (1982). The historical development. In D.L. Rogers, & D.A. Whetten (Eds.) *Interorganizational Coordination: Theory, Research, and Implementation.* Ames, IA: Iowa State University Press.

Sabel, C. F. (1996). A measure of federalism: Assessing manufacturing technology centers. *Research Policy*, 25, 281-307.

Shapira, P. (1990). Modern times: Learning from new state initiatives in industrial extension and technology transfer, *Economic Development Quarterly*, 4(3), 186-202.

Shapira, P. & Youtie, J. (1996). *Coordinating industrial modernization services: Impacts and insights from the U.S. Manufacturing Extension Partnership.* Atlanta, GA: Georgia Institute of Technology, School of Public Policy.

Shapira, P., Youtie, J., & Roessner, D. (1996). Current practices in the evaluation of U.S. industrial modernization programs. *Research Policy*, 27, 185-214.

Small Business Administration (1992). Small business data base, USEEM file,

version 8, 1991, reported in *The state of small business: A report to the President*. Washington, DC: U.S. Government Printing Office. Tables A.7 and A.22.

Sylvia, R., Meier, K., & Gunn, E. (1991). *Program planning and evaluation for the public manager*. Prospect Heights, IL: Waveland Press.

U.S. Congress, Office of Technology Assessment (1990). *Making things better: Competing in manufacturing*. OTA-ITE-443, Washington, DC: U.S. Government Printing Office.

U.S. General Accounting Office (1995). *Manufacturing Extension Programs: Manufacturers' Views of Services*. Briefing Report GAO/GGD-95-216BR. Washington, DC.

Yin, R. (1994). *Case study research*. Second Edition. Thousand Oaks, CA: Sage.

