2009

Quality Assurance and Program Evaluation: Terms, Models, and Applications

M. J. Leahy

V. A. Thielsen

Michael J. Millington
Utah State University

B. Austin

A. Fleming

Recommended Citation

Quality Assurance and Program Evaluation: Terms, Models and Applications in Rehabilitation Administration

Michael J. Leahy
Virginia A. Thielsen
Michael J. Millington
Bryan Austin
Allison Fleming

Abstract

Quality assurance and program evaluation have become essential elements in continuous improvement efforts within the public rehabilitation program and community-based rehabilitation organizations. This article is a review of quality assurance and program evaluation models and approaches that appear the most promising in helping rehabilitation organizations improve the quality of services and outcomes of individuals receiving vocational rehabilitation services in the public and nonprofit practice settings. Examples of applications will be presented, and a discussion of key elements that appear to be critical in the design and utilization of these methods in rehabilitation administration will be discussed.

Over the past 35 years, quality assurance and program evaluation systems and methodologies have become essential elements for continuous improvement efforts within the public rehabilitation program and community-based rehabilitation organizations in this country. While these concepts and methods are certainly not new to the rehabilitation field, they have taken on increased importance in recent years in relation to higher levels of accountability, the need for service delivery system improvements, and competition for financial resources to adequately fund these programs.

The initial drive within the rehabilitation movement for program evaluation began with the 1973 Rehabilitation Act, and explicit mandates for the public rehabilitation program to gather and analyze data on the effectiveness of the services provided to citizens with disabilities in order to assess the impact of these services within each state public vocational rehabilitation agency (Rubin & Roessler, 2008; Walls, Misra & Majumder, 2002; Wright, 1980). This significant accountability mandate resulted in a large-scale effort to expand program evaluation activities within these state agencies. Although initially these efforts were limited by a lack of knowledge and methodology in relation to program evaluation, in the years that followed new models and evaluation methods were greatly improved (Rubin & Roessler). At about the same time, the Commission on Accreditation of Rehabilitation Facilities (CARF) began to stress the value of program evaluation, and developed specific standards for community-based rehabilitation organizations to guide the development of program evaluation systems within these nonprofit facilities to assess the effectiveness and efficiency of services provided to the individuals with disabilities they served. More recently, specific performance indicators and targeted performance expectations have been established through the Rehabilitation Act Amendments of 1992 and 1998 for all state agencies, and CARF has increased, through its ac-

creation standards, the requirements for community-based rehabilitation organizations to address their effectiveness through ongoing outcome evaluations (Lewis, Armstrong, & Karpf, 2005).

While these federal mandates from the Rehabilitation Services Administration (RSA) for public rehabilitation programs and CARF accreditation standards for nonprofit organizations initially focused on the design of these program evaluation systems and the collection, analysis and reporting of data on the impact of services, the focus has evolved over the years to emphasize how these data are used for continuous improvement in the delivery of services and the outcomes achieved. In recent years there has also been an increased amount of attention on the concepts of quality assurance systems or plans, evidenced-based practice, and knowledge translation efforts within the healthcare, disability and rehabilitation arenas. All three of these initiatives are associated with the use of quantitative and qualitative data obtained through research and evaluation efforts to inform policy, practice and continuous improvement strategies for organizations that serve people with disabilities.

Mounting pressure to demonstrate accountability through the development and use of evidence-based practices has led to an increased focus on research dedicated to improving practice and the translation, dissemination, and utilization of that knowledge throughout the profession. The National Institute of Disability and Rehabilitation Research's (NIDRR) emphasis on knowledge translation into practice only highlights the paucity of any real evidence based-practice and set the agenda for the work to be done. As indicated by Pruett, Swett, Chan, Rosenthal and Lee (2008), the powerful question posed by Paul (1967) years ago, "*What* treatment, by *whom*, is most effective for *this* individual with *that* specific problem, and under *which* set of circumstances?" (p.111) remains for the most part unanswered within rehabilitation settings.

The purpose of this article is to provide a foundation of terms, models and applications upon which the science of the practice can be proved, and improved. To this end we will introduce Quality Assurance as the broader, more strategic concept that contains and interprets Program Evaluation from a quality perspective. Program Evaluation is defined within this context and differentiated from QA as much as possible. Select and representative models of PE are described, revealing the wide latitude taken with the term. Two current applications of quality-driven program evaluation provide examples of how the profession is addressing the issue. We will conclude with a synthesis of key factors for facilitating QA and PE implementation, drawn from both literature and professional experience.

Quality Assurance

Quality Assurance (QA) is a borrowed term that originates in industrial manufacturing. It is an artifact of the quality movement in business and reflects the values of the marketplace, that is, the centrality of the profit motive and the primacy of customer satisfaction. Quality Assurance evolved in business (Shewhart, 1980), was adopted by the medical field (Laffel & Blumenthal, 1989), and has found a foothold in Vocational Rehabilitation. As it spreads, it evolves to the particulars of the context, but remains fixed at its core: QA is a systematic process designed to identify, analyze and eliminate variation (defects) in processes and outcomes. It assumes a metric set of standards and best practices (Donabedian, 1988). Processes, not people, are the subject of evaluation and the object of continuous improvement. The needs and expectations of the client are the ultimate arbiter of meaning (Laffel & Blumenthal, 1989) and success.

The structure of standards and best practices is provided by logic models. Bringing actual performance in line with standards and practice is accomplished through an iterative process of continuous improvement, most notably, the original "Plan-Do-Check-Act" (Teague, 2005).

Logic Models

Continuous improvement requires well-defined and defended objective measures and reasoned arguments for investing limited resources in innovation. Logic models provide the structure for evaluation (McLaughlin & Jordan, 2004) and thus the argument for program relevancy (Reed & Brown, 2001). Logic models are particularly helpful to organizations as they begin to evaluate and understand what a program or process is designed to address, how the program is uniquely qualified to address the problem, and identify the expected performance outcomes (e.g., how the program knows that it has been successful).

In the logic model, evaluators interpret practice in the context of the client and consider the potential impact of service on short, intermediate, and long term outcomes. Resource/Input refers to all inputs into the system, including the unintended and situational. Activities refer to all system throughputs, including mediating factors (McLaughlin & Jordan, 2004). The logic model differentiates between output (the service rendered) and outcomes (the consequences of outputs) thereby expanding the measure of service impact to the spectrum of client needs and expectations. Short, intermediate and long term outcomes are tracked for both client (satisfaction) and agency (efficiency and effectiveness).

The logic model helps program managers navigate the issue of "merit" (program effectiveness in producing outputs) versus "worth" (broader, long term impact) of the program (McLaughlin & Jordan, 2004). Developing logic models provides the opportunity to develop an agency-wide understanding of the program, its intended impact, and how the work of each staff person contributes to the success. Misconceptions can be identified and corrected. Shared knowledge and participation in development can lead to more meaningful staff buy-in. The logic model increases capacity for improved data collection and linkages between program resources, activities, outputs, customers reached, and outcomes (McLaughlin & Jordan, 2004). Logic models help to organize staff training by moving the focus from client activities within service, to how clients are likely to change as a result of participating (Reed & Brown, 2001)

Building a logic model takes considerable time and effort. Once built it must be maintained. One could argue that mapping out and standardizing the processes of rehabilitation takes resources away from service provision. One could further argue that enforcing standardization at any level diminishes the counselor's professional and clinical autonomy. After all, counseling is a complex, not a linear, function; all decisions are unique to the situation and values of the client. The success of QA in the medical field undercuts all of these arguments (Donabedian, 1988). The assumption of the professional role, embraced by rehabilitation, requires the existence of a body of specialized knowledge that can be taught, applied, and measured for its effect. Logic models are not optional for any profession. They provide the profession's rationale for existence and the current benchmark upon which all improvements in theory and practice will be made. Logic models provide the structure upon which professional growth can happen (McLaughlin & Jordan, 2004).

PDCA Model

Continuous improvement implies a well-defined set of values and an iterative process for successive approximations of perfection. The original and most pervasive example of such a process is called Plan-Do-Check-Act (PDCA: see Teague, 2005).

1. Plan. Identify targets for improvement and prioritize according to estimated best return on investment. Implement a strategic planning process based on best practice that provides a reasoned solution or innovation. The planned innovation is clearly stated and includes (a) the anticipated outcomes, (b) measurable criteria for determining success, (c) measures for monitoring level of improvement, and (d) time frames for evaluating the results.

2. Do. Implement the plan, in graduated steps if possible. Document experiences. Gather data on identified criteria at specified and meaningful junctures.

3. Check. Evaluate the results: This is the critical step in the PDCA cycle. After you have implemented the process or procedure for a short time you must determine how it is working. Does early feedback indicate it is leading to the improvement you expected to see? Are there minor changes that need to be made now? Are the measures you are using to monitor the process sufficient?

4. Act. Decide to adopt, abandon, or revise the innovation. Prepare to plan the next improvement.

The PDCA cycle (see Figure 1) is a dynamic QA model whereby the end of one cycle is the beginning of the next. What is learned in each cycle is incorporated and the process reevaluated resulting in a continuous improvement process.

Program Evaluation

Program Evaluation is the appropriate, timely, and systematic collection, analysis, and reporting of data to facilitate stakeholder judgment concerning program worth in regards to its design, demands, size and type of effect, match between effect and need, cost effectiveness, strength of causal connections, and utility (Green & Attkisson, 1984).

The differences between QA and PE arise from their origins; they differ in perspective, not intent. QA came out of the marketplace, where competition stimulates innovation. PE, as it is experienced in rehabilitation comes out of government, where legislation promotes compliance. PE tends to focus on scientific methodologies more than QA (Brollier, 1985). PE is cyclical in emphasis rather than de-

Figure 1. Shewhart Cycle for Quality Assurance
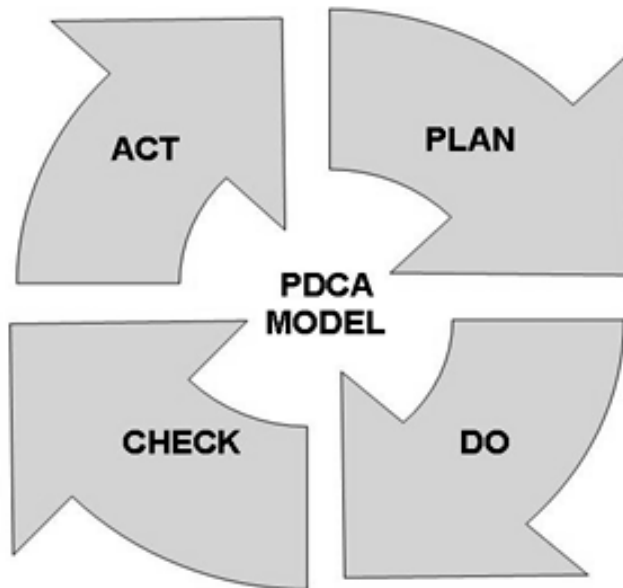
velopmental. It tends to describe a compartmentalized approach to statistical control in a program rather than part of a quality approach to management with implications for policy, planning, leadership, structure, and staffing throughout the organization. But there is no solid boundary between the two terms. QA and PE have evolved on separate courses that have converged (Baker, 1983). The differences between the models grow less distinct and relevant over time.

Program evaluation, like quality assurance, is a multifaceted construct. Unlike QA, PE is operationalized through a sizeable range of models. A review of literature spanning 20 years revealed more than 20 different models (Langbein & Felbinger, 2006; Patton, 2000; Rossi, Lipsey & Freeman, 2004; Smith & Brandon, 2008; Stufflebeam & Shinkfield, 2007; Wholey, Hartry & Newcomer, 2004) of varying interests and import. Some are familiar to rehabilitation agencies and CRPs. Others, one could argue, should be. All of them offer at least a modicum of new perspectives that could inform the profession. For the purposes of this paper we chose to illustrate the central facets of program evaluation through a selected review of the most relevant. Considering their merits in terms of practicality, simplicity, validity, and feasibility in the context of rehabilitation, we distilled these 20 models down to the most representative three: Utilization-focused Evaluation Model (Patton, 2000), the CIPP Model (Stufflebeam & Shinkfield, 2007), and the Input-Intervention- Output Model (Walls, 2001).

Utilization-Focused Evaluation (UFE)

In the UFE model, the foundation of PE is the partnership between the evaluators and intended users (Patton, 2000). There is no prescribed methodology; methodology arises out of partner negotiation.

Patton (2000) described the evaluator's role in this collaborative relationship as active, reactive, and adaptive. The evaluator is *active* in identifying intended users, and focusing useful questions. The evaluator is *reactive* in listening to the intended users and responding to what they learn about a particular situation as the evaluation progresses. The evaluator is *adaptive* in altering evaluation questions and designs as there is greater understanding of the situation or in response to changing conditions.

There is a good cultural fit for UFE in the rehabilitation profession. It makes sense to counselors in the field, promotes buy-in, and takes advantage of existing skills. While UFE is effective, it is limited in scope. The qualitative and relativistic nature of the process makes it difficult to quantify and standardize. UFE's approach is challenged by user turnover. Evaluators rely heavily on information from stakeholders on intended use, so if users are lost from the planning group, there can be serious consequences to the success of the evaluation. Competing interests may complicate partnership negotiations. Success is dependent upon stakeholder commitment (time, talent, or treasure) to the process, political advocacy, evaluator credibility, and ethical implementation.

Kuipers and Quinn (2003) utilized these evaluation principles in work with community rehabilitation services agencies. The authors involved multiple stakeholders (e.g., funding sources, staff, consumers), and the evaluation was focused on three domains- *people* (e.g., consumers, families, service providers), *program* (e.g., actual services), and *perspective* (e.g., context of services). The authors suggest a discussion template for community rehabilitation programs that wish to do participant oriented program evaluation, in this case, using guided discussion questions to elicit input. The evaluators and stakeholders identified and agreed on focus points, determined corresponding goals, and prioritized each according to overall objectives. The evaluators drafted a framework as a result of stakeholder input, and then invited comments from the work group for revisions. Based on the input and feedback from the stakeholders, measurable outcomes were developed that corresponded with the stated goals. As in any situation with a variety of users, there were some conflicts to negotiate during the evaluation planning process. It was important in this instance that all stakeholders had an

opportunity to influence the planning, and that the framework is developed in a comprehensive, contextually appropriate way (Kuipers & Quinn).

CIPP Model

CIPP is a program evaluation acronym for *context*, *inputs*, *processes*, and *products*. Context evaluations assess strengths, weaknesses, and changes that can be made to produce better outcomes. Input evaluations assess different ways of providing services, personnel needs, and financial resources necessary to meet targeted goals. Process evaluations assess how programs are running and provide support to help staff carry out planned activities. Product evaluations assess and summarize outcomes achieved (Stufflebeam & Shinkfield, 2007). The CIPP is more formal than the UFE, but it is still value driven. The organization's values lead to goals, which lead to plans, which lead to actions and ultimately outcomes. Throughout the process of the evaluation, the CIPP model is guided by values of equity and fairness and facilitates participation by individuals receiving services and others affected by system changes (Stufflebeam & Shinkfield, 2007).

The CIPP model typically takes on two roles for evaluation purposes. First, in a formative role, the evaluator follows progress over time, with repeated measures of identified criteria. Questions such as, "What needs to be done? Is it being done? How should it be done? Is it succeeding?" (Stufflebeam & Shinkfield, 2007; p. 327). Answering these questions during the evaluation process give key stakeholders (i.e., policymakers, administrators, and staff) the opportunity to provide feedback to help modify existing practices and resources to areas that will have the biggest impact. Second, in a summative role, the evaluator turns to outputs and outcomes. Questions include, "Were important needs addressed? Was the effort guided by a defensible design and budget? Was the service design executed competently and modified as needed? Did the effort succeed, and why or why not?" (p. 327).

The answers to these questions help keep staff engaged on targeted goals and also provide important information to stakeholders about outcomes (Stufflebeam & Shinkfield, 2007). One of the strengths unique to the CIPP model is that it is guided by ethical and professional principles (i.e., program evaluation standards) as established by the Joint Committee on Standards for Educational Evaluation (The Joint Committee on Standards for Educational Evaluation, 1994). Finally, this model includes and elicits feedback from all stakeholders using or being affected by the program services during the evaluation process. This may have an empowering effect on all who are involved which may lead to increased motivation, and therefore, better outcomes.

A limitation of the CIPP model may be its apparent complexity at first glance. The thought of integrating multiple evaluations and implementing this model in its true form may seem intimidating for some agencies. Further, it appears that training and considerable research expertise is necessary to live up to the high standards stipulated by the Joint Committee on Standards for Educational Evaluation. However, picking and choosing aspects of the CIPP model appears feasible for any agency and may be best used to further complement a more eclectic and individualized approach to program evaluation.

Input-Intervention-Output

The Input-Intervention-Output Model (IIOM; Walls, 2001) is a more quantitative and longitudinal variant of the input, processes, and products model introduced in the CIPP. IIOM focuses on the organizational systems. Input variables track client demographic and status (e.g., medical insurance coverage, primary support, previous employment status, and days from application to eligibility) data from the beginning of the rehabilitation process to the conclusion. Intervention variables track services and interventions rendered during the rehabilitation process (e.g., restoration, assessment, counseling, training, job placement). Outcome variables track specific and expected outcomes fol-

lowing case closure (e.g., work status, hours worked, primary support), criteria of value to client or agency.

IIOM provides a useful framework to organize variables for display, and to guide the analysis of outcomes, particularly in relation to large data sets such as the RSA 911 database. IIOM has been extensively used in rehabilitation program evaluation designs and in reporting results of the public rehabilitation program. It is the appropriate structure for statistical quality control and applied research. Using this type of analytic model allows for descriptive research on correlations among these variables and studying the impact of client characteristics on processes and outcomes.

Exemplary Practice

QA and PE are virtually indistinguishable in practice, and rightly so. They are different aspects of a singular function. It is more informative to observe how they integrate than it is to sustain definitional boundaries. What follows is a description of two current initiatives that demonstrate how QA and PE have been applied in emerging best practice.

Project Excellence

A unique model and approach that has shown positive results over an extended period of time is a synergistic model of program evaluation between a public rehabilitation agency and a university rehabilitation counseling program (Leahy, Thielsen, Groomes, Shader-Patterson, & Shamsiddeen, 2006). This long-term partnership between the Michigan Department of Energy, Labor and Economic Growth-Rehabilitation Services (MRS) and the Office of Rehabilitation and Disability Studies at Michigan State University is called "Project Excellence." This collaborative partnership operates from a utilization focused evaluation model (Patton, 2000) that values input and participation by all levels of MRS staff and university faculty, with additional input from doctoral students pursuing their studies in rehabilitation counselor education. The mutually endorsed mission that guides the work of Project Excellence states, "Project Excellence will further develop and improve the application of qualitative and quantitative data analysis with respect to Michigan Rehabilitation Services culture and programs while prioritizing regulatory responsibilities and program excellence values." (p. 14). This mission serves as the foundation for conducting research and evaluative projects that are implemented to increase the quality of services of customers with disabilities on a continuous basis (Leahy et al., 2006).

MRS now has a system in place to evaluate data to assist with making necessary policy and practice decisions to improve customer services and outcomes. MRS is also able to meet the training needs of their staff on a consistent basis and able to track the effectiveness of their training efforts. Further, over the years since the start of Project Excellence, MRS has improved its agency performance in achieving key federal indicators related to the Rehabilitation Services Administration (RSA) 911 data set. As depicted in Figure 2, one of the unique features of this ongoing project is the active involvement of staff at all levels with the program evaluation initiatives. Ideas for projects are requested each year from throughout the organization. Once these projects are identified and prioritized, Project Excellence staff work directly with groups of MRS staff and other stakeholders to design the research questions and methods to address the issue. Once data (qualitative or quantitative) have been collected and analyzed by Project Excellence, it again reviewed through work groups of MRS staff before finalization of the report and recommendations. Once this occurs the report goes to an implementation committee of MRS staff to plan for the possible implementation of the recommendations. If these activities involve a change or modification in policy or practice, Project Excellence will continue to monitor the impact of the intervention or innovation. This allows the program evaluation process to go full circle from the identification of issues to the implementation of findings and ongoing monitoring of these collaborative efforts.

Figure 2. Project Excellence Work Flow
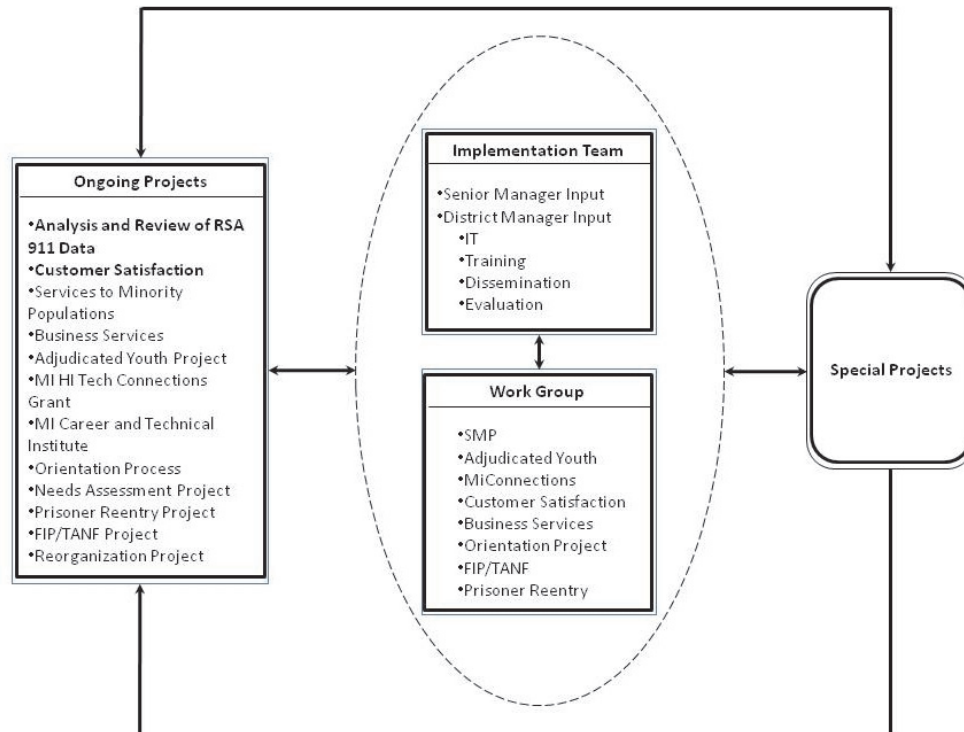
A notable strength that makes this partnership model successful is the long-standing trusting relationship that has been built over many years. Further, creative leadership, common values, and shared goals are "success factors" that continue to be significant contributors to Project Excellence's on-going impact (Leahy et al., 2006).

Peckham and the Walker Model

Staff from Peckham, Inc., a large organization in the Lansing, Michigan, were interviewed in April 2009, and asked to describe their process of evaluation and continuous improvement. This organization is CARF-Accredited, and runs a variety of business services, production, and employment programs. They utilize a modified Walker Model (Walker, 1981) approach as the overall framework of their program evaluation activities, which is fairly typical in community-based rehabilitation organizations accredited by CARF, but have greatly expanded this process with an empowerment model that runs throughout the organization in terms of continuous improvement and innovation. The Walker Model typically requires the development of program objectives, primary measures, goals and expectations in relation to effectiveness, efficiency, and consumer satisfaction. Each year, staff participates in three separate aspects of evaluation. Two aspects, the overall business plan and goals related to supporting the overall mission are for all staff regardless of whether or not service delivery is part of the responsibility of their department. The service departments also have goals and outcomes evaluations.

In the service goals and outcomes evaluation, each department is asked to come up with five goals that they feel are most important to focus on this year. Four of the goal areas are prescribed by CARF standards (customer satisfaction, service access, efficiency, and effectiveness) and the program staff selects the additional one. Often, goals will be in line with or exceed those set by external funding sources. For example, the community placement services department might set a number of placements for the goal related to effectiveness, and could either use the goal set by their referral sources, or select a goal that is higher. Evaluation personnel work with staff from these programs to craft goals that are measurable and provide support in identifying how progress will be measured. Departments submit quarterly reports, and identify goals that are not being met so they may come up with a corrective action plan to explain and address the issue. On a yearly basis, department managers present their plan and outcomes to managers from all service departments.

Individuals who receive services from this organization also have an opportunity to participate in the quality improvement process. Town hall meetings are held in a variety of formats to accommodate different learning styles, communication levels, and languages. These meetings are utilized to ask consumers what could be done to improve services. During the meetings, staff also updates consumers on implementation of past recommendations. Consumer satisfaction surveys are sent out to both external professionals who refer customers and to the customers themselves at regular intervals for longer-term programs and after completion for short-term programs. Ratings and comments are compiled and shared with department staff and as part of the overall yearly reporting. Any issues that seem to require staff attention are directed to the manager of the department, and they create an action plan to address or explain the issue.

The organization-wide methods of evaluation and continuous improvement include the Organization Business Plan, and Key Indicators that support the organizational mission. Each department is responsible for developing its part of the business plan, which is reviewed quarterly. To develop the yearly Key Indicators, leadership staff select employees who have not been previously involved in designing these goals. The group brainstorms and comes up with the indicators that will be focused on that year. Examples from past years have included initiatives like health and wellness, infusing technology, continuous learning and advancement, and financial health. All staff have a vested interest in supporting these goals, as at the end of the year if the Board of Directors judges that the goals are met then all staff receive a bonus. Seminars and in-service training programs are convened to support the key indicators, such as financial planning workshops or online training programs that are open to all staff and consumers. An example of a goal that was set for one of the indicators was that all staff would complete five paid hours of online training. Managers are motivated to allow staff to complete the five hours because of the organization-wide support for the goal and by the offer of the staff-wide bonus if all goals are met.

Another form of organizational improvement lies with the staff. Staff are asked to come up with ideas to improve their work, and put these ideas into practice. Evaluation staff manage a database and the documentation of the ideas submitted and put into practice are sent to an administrative representative and the person's supervisor. If a staff member has completed 10 of these ideas put into practice by the time he or she has a yearly review, an additional half percent is added to the yearly raise. Consumers of the agency are also asked to share any ideas they have for improvement on an ongoing basis. A separate database is kept to hold ideas that have not been put into practice as more of a "suggestion box."

Evaluation staff noted many benefits to the work of the agency as a result of these evaluation practices. However, they acknowledged that the complete evaluation plan took years to develop and the process of getting all staff involved was a challenge. Among the greatest challenges were the time and effort that this commitment to continuous improvement requires, the support that non-evaluation

staff need to understand and develop goals for their departments, and making these practices part of the repertoire for every employee. Often staff do not see how their work is related to the evaluation goals, and it takes vigilance to keep them engaged. Clear benefits have been observed as well. As a result of the consistent focus on improvement and eliciting ideas from staff and consumers, an attitude of, "we are doing well, but we can always do better" is present in the organizational culture. Because each service department has yearly goals and documentation to support the progress towards goals, it is easy to prepare for CARF reviewers and other external evaluators for the agency. Evaluation staff have also found that the information gathered from the service goals is useful in marketing. Department managers have more specific information about the success of their programs to share with potential business partners.

Key Elements of Evaluation

In 2003, the United States General Accounting Office (GAO) conducted a study of agency capacity for evaluation and concluded that the key elements of evaluation were an evaluation culture –a commitment to self-examination, data quality, analytic expertise, and collaborative partnerships. Although the agencies studied by the GAO were large national organizations (e.g., National Science Foundation, the Administration for Children and Families), and not specifically related to rehabilitation or disability, the findings of this report are very consistent with our own assessment of key factors that lead to success in designing, implementing and maintaining quality program evaluation systems for the purposes of accountability and continuous improvement. Therefore we will use the framework identified by the GAO, along with additional elements we have identified through our review and experience, to discuss key elements and factors for program evaluation within the public rehabilitation program and community-based rehabilitation organizations.

Evaluation Culture – A Commitment to Self-examination

Organizations demonstrate an evaluation culture through commitment to self-examination, learning through experimentation, and the use of data-driven policy and practice decisions (GAO, 2003; Leahy et al, 2006; Lewis et al, 2005). In this type of culture, innovation and adherence to core institutional values are at the center of ongoing interest in evaluating performance and identifying areas where additional resources or training are needed to meet the mission of the agency. In organizations where this type of evaluation culture is not present, program evaluation and quality assurance activities are viewed from merely a compliance perspective, and typically do not produce the type of meaningful results and improvements that are possible for organizations that fully embrace and engage themselves within a culture where evaluation is highly valued and resourced. One of the clear challenges in today's rehabilitation environments to the maintenance of this type of evaluation culture is frequent turnover of leadership personnel who set the expectations and rewards for this type of culture within organizations. In addition, poor economic times and scarce resources may also impact the amount of emphasis given as a result of financial allocation issues.

Demand for Data Quality

In rehabilitation organizations program evaluation activities, at a minimum, result in the collection of data (quantitative and qualitative), organized in relation to specific evaluation or research questions that include measures related to effectiveness, efficiency, consumer satisfaction, and needs assessment (Lewis et al, 2005). Data quality or data integrity (Leahy et al, 2006) is absolutely critical in drawing any conclusions about the results obtained through evaluation, and organizations should take steps through monitoring and evaluation efforts to ensure the credibility, reliability and consistency of data (GAO, 2003) used in the studies and reports they generate.

Although there has been considerable amount of data collected and analyzed through program evaluation, most of these efforts attempted to determine relationships among existing data and employment outcomes, rather than a purposeful approach to define what types of intervention or services appear to work best with what specific populations, under what specific conditions (Paul, 1967). This is an area of weakness and limitation in regard to our existing program evaluation and research on employment outcomes that needs to be addressed in future evaluation and research initiatives in order to establish evidence-based practices.

As indicated by Saunders et al (2006), in their study of employment outcomes in rehabilitation over the past 25 years, "we need more intervention related studies (pre-experimental, quasi-experimental and experimental) and less ad hoc studies regarding employment outcomes. Research efforts need to be theoretically based, with clear definitions of variables, populations and interventions. Data, such as effect size, need to be routinely included in research reports to maximize our ability to use meta-analysis and other statistical approaches. Researchers need to adhere to guidelines and recommendations regarding statistical inference, such as those drafted by the APA Task Force on Statistical Inference and reported by Wilkinson (1999), and others (Parker & Szymanski, 1999) in order to improve the quality and strength of our empirical efforts. Finally, in order to meaningfully address knowledge translation and the development of evidence-based practices, we need more replication and extensions of previous research in order to build coherent, rigorous lines of research that serve to inform policy and practice in the future" (p. 17).

Analytic Experience

There is a great deal of variability in relation to the analytic experience of program evaluators within the public rehabilitation program and community-based rehabilitation organizations (GAO, 2003). There are also different needs in relation to the complexity of the program evaluation design and differences based on size of the agency and the number of programs and services offered. Regardless of these factors all individuals responsible for program evaluation should receive the level of training and support required to accomplish the tasks required and to ensure that the results are systematic, credible and objective. In addition to hiring and training qualified staff to perform evaluation roles within rehabilitation organizations, there are some unique partnerships that have been developed and maintained with university level rehabilitation programs where research expertise has been effectively used to either supplement existing agency-based evaluation programs or to assume this function for the state agency (Leahy et al, 2006).

Collaborative Partnerships. Agencies engage in collective partnerships for the purpose of leveraging resources and expertise (GAO, 2003). In rehabilitation, collaborations and partnerships have become critically important in the delivery of services to individuals within a community. This is a recognized strength of the profession. It is, however, less likely that these same collaborations and partnerships have impacted program evaluation functions for the respective agencies involved.

There are however opportunities for this level of collaboration where each agency participates and contributes toward a desired evaluation outcome. For example, when conducting a comprehensive statewide needs assessment, it makes intuitive sense to involve multiple agencies and organizations in this task rather than have individual agencies perform separate needs assessments to inform resource allocation.

Leadership Responsibilities

One of the most critical issues that affects the quality and usefulness of program evaluation efforts is the degree to which the leadership of the organization or agency values these evaluation efforts. This is similar to the development of an evaluation culture. If the State Director in the public rehabilitation

programs or CEO or Executive Director of Community-based rehabilitation organizations does not value and invest in program evaluation efforts, those activities will end up minimally effective and only serve to satisfy regulatory or accreditation standards compliance.

Utilization and Continuous Improvement

The ultimate worth and value of evaluation can be judged in terms of their utility and usefulness in informing policy and practice and assisting the rehabilitation agency or organization in continuous improvement efforts. In thinking about utilization, the conventional three-way classification of the manner in which evaluations are applied is helpful (Rossi, Lipsey & Freeman, 2004; Leviton & Hughes, 1981; Weiss, 1988). This classification includes (1) direct (instrumental) utilization, where there is documented use of evaluation findings by decision makers and other stakeholders; (2) conceptual utilization, where the use of evaluations is to influence thinking about issues in general; and (3) persuasive utilization, where evaluation results are used to either support or negate a particular position on an issue (Rossi et al.). The variables that appear to affect evaluation utilization include: relevance, communication between researchers and users, information processing by users, plausibility of research results, and user involvement or advocacy (Leviton & Hughes).

Although the goal in rehabilitation management may be direct utilization of evaluation findings, this is difficult to achieve and potentially a very complex process to navigate. Data and findings are but one consideration in the overall decision process. It is more likely that utilization occurs over time, and in smaller increments than one would initially expect. However, even with these challenges, the goal should remain direct and conceptual utilization of evaluation finding to assist agencies and organizations make continuous improvements and creative innovations to the services they offer to individuals with disabilities.

Conclusion

Evaluation theory, of which QA and PE are intertwined threads, has evolved towards more sophisticated models and applications over time (Shadish, Cook, & Leviton, 1995). It has shifted the focus of evaluation in rehabilitation from compliance to continuous improvement and in doing so has opened new paths to the development of evidence-based practice and knowledge translation, dissemination, and utilization. Thus endowed, the profession is challenged to address Paul's (1967) fundamental question in terms of evaluation as it has been similarly challenged in service. To paraphrase, "*What* evaluation strategy, by *whom*, is most effective for *this* program with *those* specific issues, and under *which* set of circumstances?" The evidence base for program evaluation in rehabilitation is rudimentary. We have included key factors for optimizing the organizational utility of an evaluation plan, but we cannot offer grounded empirical advice on how to craft its component parts. If this evolution to greater complexity is to serve the evaluation needs of the rehabilitation profession, we need to study its application with the same scientific rigor that quality PE would apply to service.

References

Baker, F. (1983). Quality assurance and program evaluation in health care. *Evaluation and the Health Professions*, *6*(2), 149-160.

Brollier, C. (1985). Program evaluation research: An administrative tool. *Work-Related Programs in Occupational Therapy*, *2*(4), 105-111.

Deming, E. W. (1986). *Out of the crisis*. Boston, MA: MIT Center for Advanced Engineering Study.

Donabedian, A. (1988). The quality of care: Can it be assessed? *Journal of the American Medical Association, 260*, 1743-1748.

Green, R. S., & Attkisson, C.C. (1984). Quality assurance and program evaluation: Similarities and differences. *Behavioral Scientist*, *27*, 552-582.

Kuipers, P., & Quinn, R. (2003). The template: A cooperative approach to evaluating community rehabilitation services. *Journal of Rehabilitation, 69*(1), 4-14.

Laffel, G. & Blumenthal, D. (1989). The case for using industrial quality management science in healthcare organizations. *Journal of the American Medical Association*, *262*(2), 2869-2873.

Langbein, L. & Felbinger, C.L. (2006). *Public program evaluation: A statistical guide*. Amonk, NY: M.E. Sharpe.

Leahy, M. J., Thielsen, V., Groomes, D. A. G., Shader-Patterson, L., & Shamsiddeen, J. (2006). A synergistic program evaluation model partnership between a state agency and a university program. *Journal of Rehabilitation, 72*(3), 12-18.

Leviton, I.C., & Hughes, E.F. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, *5*(4), 525-548.

Lewis A., Armstrong, A.J., & Karpf, A.S. (2005). Using data to improve outcomes in rehabilitation practice. *Journal of Rehabilitation Administration*, *29*, 107-120.

McLaughlin, J.A., & Jordan, G.B. (2004). Using logic models. In J.S. Wholey, H.P. Hatry, & K.E. Newcomber (Eds.), *Handbook of practical program evaluation* (2nd ed., pp. 7-32). Hoboken, NJ: John Wiley & Sons.

National Conference of State Legislatures. (2006, June). *Peer review: Florida's office of program policy analysis and government accountability.* Retrieved April 14, 2009, from http://www.wmich.edu/evalctr/jc/

Parker, R.M., & Szymanski, E.M. (1999). From the editors: Recommendations of the APA Task Force on Statistical Inference. *Rehabilitation Counseling Bulletin*, *43*, 3-4.

Patton, M.Q. (2000). Utilization-focused evaluation. In D.L. Stufflebeam, G.F. Madaus, & T. Kelligan (Eds.), *Evaluation models* (pp. 425-439). Boston: Kluwer Academic Publishers.

Paul, G.L. (1967). Strategy of outcome research in psychotherapy. *Journal of Counseling Psychology*, *31*,109-118

Pruett, S.R., Swett E.A., Chan F., Rosenthal, D.A., & Lee, G.K. (2008). Empirical evidence supporting the effectiveness of vocational rehabilitation. *Journal of Rehabilitation*, *74*(1), 56-63.

Reed, C.S., & Brown, R.E. (2001). Outcome-asset impact model: Linking outcomes and assets. *Evaluation and Program Planning, 24*, 287-295.

Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks: Sage.

Rubin, S.E., & Roessler, R.T. (2008). *Foundations of the vocational rehabilitation process* (6th ed.). Austin, TX: Pro-Ed.

Saunders, J.L., Leahy, M.J., McGlynn, C., & Estrada-Hernandez, N. (2006). Predictors of employment outcomes for persons with disabilities: An integrative review of potential evidenced-based factors. *Journal of Applied Rehabilitation Counseling*, *37*(2), 3-20.

Shadish, W.R., Cook, T.D., & Leviton, L.C. (1995*). Foundations of program evaluation: Theories of practice.* Thousand Oaks, CA: Sage.

Shewhart, W. A. (1980). *Economic control of quality of manufactured product/50th anniversary commemorative issue*. Milwaukee, WI: American Society for Quality.

Smith, N.L., & Brandon, P.R. (2008). *Fundamental issues in evaluation*. New York: Guilford.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). CIPP model for evaluation: An improvement/accountability approach. In D. Stufflebeam (Ed.), *Evaluation theory, models, and applications* (pp. 325-365). San Francisco: Jossey-Bass.

Tague, N, R. (2005). *Quality toolbox* (2nd ed). Milwaukee, WI: ASQ Quality Press.

The Joint Committee on Standards for Educational Evaluation. (1994). The *program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.

U.S. General Accounting Office. (2003). Program evaluation: An evaluation culture and collaborative partnerships help build agency capacity. *United States General Accounting Office Report to Congressional Committees. GAO-03-454. Washington, D.C.: U.S. Government Printing Office.*

W.K. Kellogg Foundation. (2004). *Using logic models to bring together planning, evaluation, and action: Logic model development guide*. Battle Creek, MI.

Wakler, R. (1974). The ninth panacea: Program evaluation. In E.C. Bennett & M. Weissinger (Eds.), *Program evaluation: A resource handbook for vocational rehabilitation* (pp. 45-60). New York: ICD Rehabilitation and Research Center.

Walls, R.T., Misra, S., & Majumder, R.K. (2002). Trends in vocational rehabilitation: 1978, 1988, 1998. *Journal of Rehabilitation*, *68*(3), 4-10.

Walls, R.T. (2001). Measurement of client outcomes in rehabilitation. In B. Bolton (Ed.), *Handbook of measurement and evaluation in rehabilitation* (3rd ed., pp. 311-332). Austin, TX: Pro-Ed.

Weiss, C.H. (1988). Evaluation for decisions: Is anybody there? Does anybody care? *Evaluation Practice*, *9*(1), 5-19.

Wholey, J.S., Hartry, H.P., & Newcomer, K.E. (2004). *Handbook of practical program evaluation* (2nd ed.). San Francisco: Jossey-Bass.

Wilkerson, L. (1999). Statistical methods in psychology journals: Guidelines and expectations. *American Psychologist, 54*, 594-604.

Wright, G. N. (1980). *Total rehabilitation*. Boston: Little, Brown and Co.

About the Authors

Michael J. Leahy, Ph.D., is a professor and director of the Office of Rehabilitation and Disability Studies at Michigan State University. He is also the principal investigator for Project Excellence at MSU.

Virginia A. Thielen is a senior research associate and co-principal investigator for Project Excellence at MSU.

Michael Millington, Ph.D., is the director of the National Clearinghouse of Rehabilitation Training Materials and assistant research professor at Utah State University.

Bryan Austin and **Allison Fleming** are second year doctoral students in the Rehabilitation Counselor Education Program at MSU, and research assistants for Project Excellence.