

Position Paper

The EvaluLead Framework

Examining Success and Meaning:
A Framework for Evaluating
Leadership Development Interventions
in Global Health

March 2002

By,

John Grove & PLP Team

&

Leadership Evaluation Advisory Group (LEAG) Members¹

*Sponsored by the Population Leadership Program (PLP)
A project of the Public Health Institute*

*PLP is funded through cooperative agreement
HRN-A-00-99-00012-00 with the United States
Agency for International Development.*

¹ See “Acknowledgements,” page 9

The EvaluLead Framework

Examining Success and Meaning: A Framework for Evaluating Leadership Development Interventions in Global Health

A Call for Action

Leadership development interventions present unique evaluation challenges. Desired results and evaluation data are often derived from highly personalized developmental processes, while outcomes are expected at much broader levels of the organization or society. Often, outcomes are difficult to capture using purely experimental procedures as they assume a static control over what is in actuality, an “open system.” In most cases, a control group is difficult to identify. The literature around evaluation of leadership and professional development in corporations reveals that the vast majority of training programs gather satisfaction-level data only (Olian: 1998). Additionally, the term leadership development is subject to many different interpretations based on contextual realities at the individual, organizational, and societal levels. This interpretation is made even more complex in working across multiple cultural realms and belief systems.

Practitioners in all sectors concerned with human and systems development employ a wide variety of development approaches in reaction to these different contextual realities. A mix of group training, one-on-one interventions and long and short-term job assignments are among the approaches used. How can leadership development interventions be evaluated and subsequently improved to lead to success at all levels: individual, organizational and societal? What do the effects of quality leadership development content in professional or organizational development interventions look like? Evaluation, as a professional field of practice, must examine these questions and create appropriate and constructive approaches to respond to the dynamic nature of these interventions and their associated effects. There is an opportunity, now, to discuss these issues, and emerge with a flexible framework for evaluating leadership development interventions.

Leadership Interventions

Leadership development interventions aim to enhance an individual’s ability to face challenges and achieve valuable outcomes. The content of such interventions may include exploring personal awareness and capability and providing strategies that focus one’s actions in a system. Even when the ultimate aim is to change an organization or a society, the individual is the focus of leadership development because people are the agents of change. Interventions assist leaders to apply a unique balance of objectively and subjectively driven actions. Effects of these interventions become more and more difficult to attribute as one moves away from the point of intervention. Evaluation approaches must be flexible, yet organized and sound, to capture the outcomes. Both an objective distance and a subjective presence with the point of intervention are critical to assessing for these outcomes (Kibel: 2001). We limit inquiry, and thereby discovery, if evaluation in this area is tied to one approach.

The Opportunity

Aside from accountability, why evaluate leadership development interventions? Ellen Van Velsor (1998: 263) of the Center for Creative Leadership proposes the following goals for assessing the impact of development experiences:

- Understanding a person’s development as a result of the experience;

- Evaluating and fine-tuning an intervention so that it better meets its goals;
- Documenting whether participation in development experiences is affecting the bottom line.

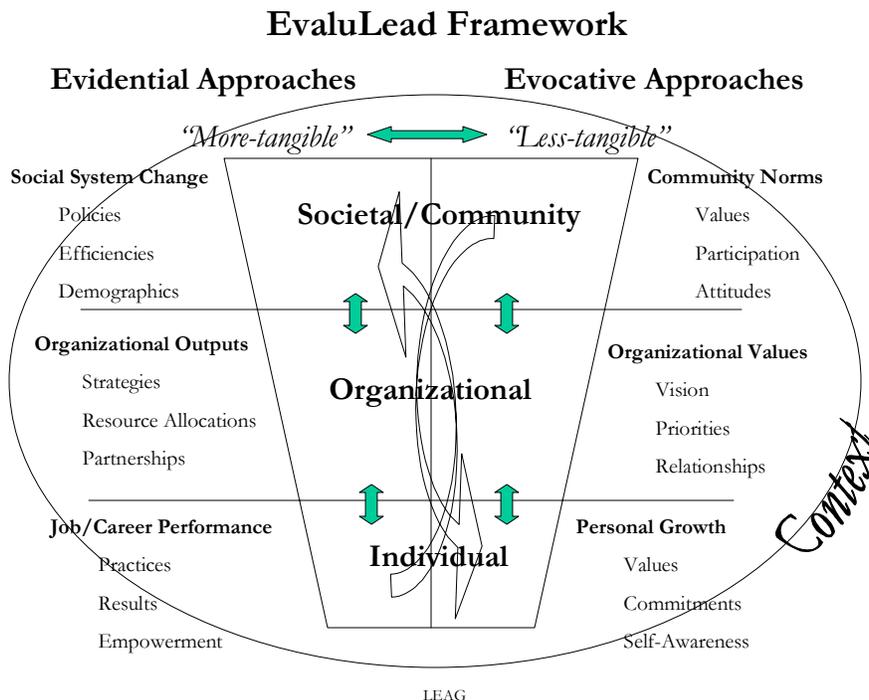
PLP adds the following goals:

- Promoting use of learning-centered reflection;
- Identifying useful leadership competencies in a particular setting;
- Increasing knowledge about what works.

We evaluate leadership development interventions *to improve activities that sustain the achievement of positive outcomes for organizations, communities, and countries by individuals*. Leadership development does not occur at a static point in time, but through a multitude of experiences. The relationship between an intervention and observable result is not direct. Evaluation approaches must explore learning in addition to job and career performance expectations. Descriptive data serves critical purposes, while stories and interpretive techniques serve essential purposes as well. A framework for evaluating leadership development requires that values and norms, *and* performance factors, are brought together in assessment to uncover meaningful development.

Framing an Approach

On March 1, 2002, PLP convened a one-day meeting of leadership program stakeholders and seasoned resource people to react to a proposed conceptual framework for examining evaluation of leadership development interventions. The basic premise and assumptions of the framework held firm, with key changes being proposed to advance the framework toward operation.



Concept

The framework conceptually outlines two types of evaluation approaches based on the particular kind of impact sought, making the critical point that both types are needed at all times. Further, three levels of effect related to elements of leadership development intervention outcomes are outlined. Six domains of *outcome elements* emerge. These domains are outcomes that leadership development interventions aim to influence. Outcomes are changes in the behavior, relationships, activities, or actions of the people(s), groups, and organizations with whom a program works directly (Earl: 2001: 1). Some changes may be observable, however, some are not immediately discernable such as personal assumptions, attitudes, and values. The elements within these domains are based on the goals and objectives of the particular program. (NOTE: The outcomes listed are example outcomes reflecting some of PLP's current leadership emphases). Given the personal nature of leadership development, and the expectation for those leaders to deliver results, we must employ evaluation approaches based on relevant levels of effect and both types of impact.

Evocative Approaches

The nature of evaluative inquiry for leadership development is presented in two types, "Evidential" and "Evocative." The "Evocative" column focuses on impacts that emerge through an interpretation or inference. These approaches attempt to uncover meaning by examining

relationships and personal learning in-depth. These impacts may be considered less tangible, however, critical to understanding a leader's development within a system. Evaluation approaches include in-depth interviews, stories, cases, textual analysis, and ethnographic techniques. Outcomes occur at three levels: individual, organizational, and societal/community. Outcome elements include shifts in personal and/or community values, re-orientation pertaining to ethical norms, change or adaptation in judgment, refinement or expansion of personal and/or community beliefs, and unexpected results or understandings acquired beyond one's plans.

Evidential Approaches

The "Evidential" column focuses on impacts of leadership development interventions that emerge through objective observation gleaned at a "distance" from the leader. These outcome elements could be thought of as more tangible, and important for determining successes and areas for improvement by focusing on gathering concrete evidence. Evaluation approaches include surveys, 360-degree instruments, pre and post intervention assessment, and retrospective reviews at a given point (static) in time and often in aggregate form. Again, outcomes occur at three levels: individual, organizational, and societal/community. Outcome elements include distinct accomplishments as reported by others, changes in behaviors as observed by colleagues, verifiable completion of tasks against a written plan, successful acquisition of monetary and human resources, and, expansion of progressive procedures that are designed to organize systems.

Approach	Purpose	Illustrative Methods	Sample Outcomes
Evocative	Uncover meaning by examining relationships and learning in-depth.	In-Depth Interview/Conversation Text Analysis Ethnographies Narrative/Stories/Vignettes	Changes in: Values Vision Self-Awareness
Evidential	Determine successes and areas for improvement based on numeric and/or physical evidence.	360 Feedback Survey Pre/Post Intervention Static Retrospective Reviews Experimental Designs	Changes in: Skills Strategies Policies

Context

An individual's leadership understandings, actions, and behaviors are contextual. Whether it is the intervention, the outcomes of the leader, or both for which we evaluate, context is a critical parameter. It is not enough to simply recognize context, rather, it is a factor to be understood as an integral element in the design, operation, and interpretation of evaluation findings at the individual, organizational, and societal levels. As a precursor, a program's interventions should be based on context-specific notions of leadership and related needs.

Learning and Program Improvement

The relationship between the framework's outcome levels is dynamic, connected, and complex. Outcomes at the organizational and societal levels are often associated with those on the individual level, and vice versa. The arrows crossing all domains of the framework represent the fluidity and inter-connectedness of outcomes at all levels. Incorporating the nature of these connections in evaluation design and implementation will result in learning opportunities for

stakeholders. Program staff must incorporate learning-centered principles that emphasize the importance of personal and group learning and support program improvement.

Framework Application

The EvaluLead Framework is to be used as a lens through which evaluation for leadership development programs can be conceptualized. It is labeled a framework, as opposed to a “model” to connote flexible, yet organized, application. One single prescriptive model for making evaluation operational across multiple programs is impossible due to the multitude of contexts, goals, activities, outcomes, and accountabilities. Therefore, the framework serves as a view of programs from the perspective of a “higher altitude.” A program’s monitoring and evaluation plan and products are a separate, yet linked, piece. The EvaluLead framework provides parameters for designing and implementing a plan that would incorporate the basic conceptual elements outlined. Every leadership program has somewhat different emphases, both in terms of recipients of activities, the type of activities employed, and the actions and outcomes that are desired from the efforts. Additionally, where those outcomes are expected to be seen vary greatly across programs. Some programs will only seek outcomes at one level, while others will seek outcomes at multiple levels and to varying degrees. The following questions guide the development process to concrete plans and data collection approaches:

- How are notions of leadership different across the contexts we work in? What is leadership and associated outcomes as defined by our target audience?
- Do we expect, and/or, are we accountable, for outcomes at the (individual, organizational, societal) level?
- What outcomes do we expect from our interventions? (i.e. We expect to see changes in personal awareness, skills, and policies.)
- What is a realistic timeframe for seeing outcomes?
- How can we assess for those outcomes numerically (evidential) and through narrative (evocative)?
- Over time, what indications of change, and at what level do they emerge? What indications are we seeking?
- How will we use our findings? What can we learn from those findings to improve our program’s effectiveness?

Background Analysis: The Challenge of Measuring People

The breadth and depth of literature related to evaluation of leadership development are endless. PLP has sorted through a vast selection of this work to identify sources that seem to be influencing the current “state of the leadership-evaluation art.” This analysis provides an expansion of key points presented previously and outlines the theoretical underpinnings and operational imperatives for the framework.

Kirkpatrick’s straightforward four-levels of training evaluation is clearly a foundation for a large volume of work in training evaluation. Research confirms that the struggle for practitioners who use this model is actually conducting and interpreting the latter three levels of evaluation (learning, behavior, and results). Kirkpatrick (1998: 24) says it most succinctly, stating that “trainers should proceed to the [latter] three levels as staff, time and money are available.” The main reason to conduct evaluation of training programs has as much to do with justifying the existence of the

training department as it does determining if the program is useful to its participants or needs to be improved” (1998: 18). With this as an underlying assumption it is not hard to understand why level one is the one most used (the traditional “smile test”, in Kirkpatrick’s language, Reaction) and why there is currently a move to quantify the “intangibles” (Abernathy 1999: 4). While this model has withstood the test of time, it is limited for leadership development by its focus on classic time-bound group training interventions. Leadership development often involves one-on-one interventions such as mentoring, personal coaching, on-the-job experiences, and change management.

Developmental programs are often faced with “moving targets” in terms of delivering and evaluating interventions. Peter Senge’s (1990: 71-73) idea of “dynamic complexity” illustrates this reality, stating:

Situations where cause and effect are subtle, and where the effects over time of interventions are not obvious...or when the same action has dramatically different effects in the short run and the long, there is dynamic complexity. When an action has one set of consequences locally and a very different set of consequences in another part of the system, there is dynamic complexity. When obvious interventions produce non-obvious consequences, there is dynamic complexity...[There are] complex dynamics...when it takes...months to hire and train new people...nurture management talent....

While this view resonates with the challenge of delivering and evaluating leadership development interventions, we hope and strive to see tangible outcomes from these efforts. Both publicly and privately funded programs are accountable to tax payers or shareholders. Government-funded programs must show results for every tax dollar spent. Foundations must show results for every tax dollar diverted from government-funded programs (Porter: 1999). Therefore, most programs must be equipped to demonstrate outcomes within results frameworks and donor-established evaluation parameters, and to justify the value of activities to wider stakeholders.

USAID charges projects to “convey the development hypothesis implicit in the strategy and the cause-and-effect linkages between the intermediate results and the objective” (Haecker 2000: 1). However, at the same time a caution is made that “it is hardly ever possible to prove the cause-and-effect relationships between results that are identified in a results framework (nor is it necessary)” (2000: 2). This tension is particularly relevant with regard to leadership development as there are multiple actors receiving various kinds of interventions in an “open” system. Simon Zadek, an advocate in the field of social accounting and accountability systems (1994:3) points to an example of the interpretation process in which organizations often engage. He writes,

As one senior official in the World Bank argued recently, ‘we are an institution that makes very heavy use of tabulated data in justifying our decisions, but we base our decisions on qualitative arguments that can only be understood by analysing [sic] text and oral exchange.

While a Newtonian kind of cause and effect may be difficult (and perhaps not worth the trouble) to establish, clarity of programmatic purpose and theory coupled with an array of sound methodologies is critical to determine results and areas for improvement.

Dynamic complexity does not detract from the need to show outcomes and impact, or extinguish the value of organizational and programmatic learning. Rather, this complexity challenges us to study proximate relationships between interventions and outcomes in order to provide data to improve programs and satisfy accountabilities. Evaluation, in this kind of system, focuses on the individual and their development. An underlying assumption for evaluation within the Human

Performance Technology school of thought is not only to justify the training department, but also to implement interventions “with concern for those human beings who will be the objects of your attention” (Stolovitch:1999: xvi).

Given this complexity, Allen and Cherrey state the importance of “meaning-making” to sustain leadership and organizations oriented toward “organic” change in a community or organization. Allen and Cherrey (2000: 55) propose that sustainable change does not occur through the non-reciprocal delivery of messages from the top, but that change occurs through relationships, stating, “meaning in the organic change process is found through the ability to see connections and relationships between many variables. The system is seen as a living entity rather than an inert mass.” Leadership in this case rests on the individual’s placement with and among others involved in actions oriented toward meaningful change. How can we capture that process, *and*, the associated results? Herda (1999: 133) deals with subjectivity by placing the evaluator and the evaluated in a learning-based communicative relationship:

Rather than merely a researcher and data to be observed, counted, and controlled based on subjective-objective distinction, there exists in real life science communities and human judgments. Our transformation into this community encounters a shift in our model of rationality from one that searches for determinate rules to one that emphasizes interpretation.

Similarly, Kibel puts forth the recognition of the “open system”, or non-static environment of “transformational” programs that focus on individual learning and community change. Kibel introduces narrative as a viable approach to documenting outcomes through a system of journal entries (Kibel: 1999). Since this work, Kibel has re-worked this approach to focusing on a ‘journey’. Participants in programs submit journal entries, focusing on past, present, and future actions in relation to certain experiences and contributors to their journey. Kibel posits “objective distance” and “subjective presence” on an equal plane for evaluation of personal transformation (Kibel: 2001). This approach is a major shift from traditional experimental approaches that assume control over “variables” rather than recognizing the open nature of the system and the role of the individual on a particular journey. Here, a participant is the *subject* of assessment and sustainability rather than the *object* of purely descriptive data that will be interpreted free of that individual’s role in the plot of the story.

The Inter-American Foundation (IAF) developed The Cone Framework, which later became a driving factor in grassroots development program design. The framework represents a focus on both “tangible” and “intangible” types of impact based on three levels of results in community development. The team that developed the framework wanted to recognize the importance of both type of results, stating,

Tangible in this context refers to the material results that can be counted, measured, and substantiated by direct evidence. Intangible results are those gains or losses that can be observed and inferred but are harder to measure directly or in simple quantitative ways (Ritchey-Vance: 79).

The EvaluLead framework is greatly inspired by this work, as evaluating the results of leadership development interventions is similar to that of broadly mandated community development. However, all results are tangible if the most appropriate methodology is used based on

the context of the intervention. LEAG re-visited the terms “tangible” and “intangible” in order to establish appropriate evaluation approaches that are categorized as “evidential” and “evocative.”

Without losing the critical data that evidential approaches to evaluation provide, the EvaluLead framework offers a way to broaden our inquiry with evocative approaches so as to gain greater insights into what, how and why leadership development interventions have the effects they do. These effects, discovered through the examination of concrete pieces of evidence, uncovering meaningful relationships and personal learning, reveal not only the success of the intervention but its meaning in the lives of those who participate and those communities and people who are its beneficiaries.

Acknowledgements

PLP wishes to thank all participants in the LEAG Meeting held on March 1, 2002 at the Waterfront Hotel in Oakland, California. We look forward to LEAG 2003.

Advisors: Barry Kibel, Pacific Institute for Research and Evaluation
 Jennifer Martineau, Center for Creative Leadership
 Jerome Peloquin, Performance Paradigm
 Hallie Preskill, University of New Mexico

Program Representatives:

Gbolahan Oni, The Bill and Melinda Gates Institute for Population and Reproductive Health
 Carol Woltring & Liz Schwarte, Center for Health Leadership and Practice
 Martina Frank, Institute of International Education
 Robert Minnis, International Health Programs
 Ayliffe Mumford, PLP Evaluation Consultant
 Carole Leland, Leadership Enterprises
 Joan Galer & Melanie Powers, Management Sciences for Health
 Sharon Rudy, Susan Lavezzoli, Laurel Halsey, Gail Rae, Alida Bouris, Angela McKusker, PLP/PHI
 Jim Wiley, Public Health Institute
 Babatunde Osotimehin, University of Ibadan, Nigeria
 Mary Anne Mercer, University of Washington.

Designer/Facilitator: Loni Davis, Davis & Associates

For more information on the EvaluLead Framework and the LEAG, Contact John Grove, PLP Learning & Evaluation Officer at jgrove@popldr.org.

References

- ABERNATHY, D.J.
1998 *Thinking Outside the Evaluation Box. In Training and Development. Vol. 53, no. 2: 18-24.*
- ALLEN, K.
CHERRY, C.
2000 Systemic Leadership: Enriching the Meaning of our Work. Lanham, Maryland: University Press of America.
- EARL, S.
CARDEN, F.
SMUTYLO, T.
2001 Outcome Mapping: Building Learning and Reflection into Development Programs. International Development Research Centre: Ottawa, Canada.
- HAECKER, J., editor
2000 *Building a Results Framework. In Tips (No Vol. #), no. 13: 1-6.*
- HERDA, E.
1999 Research Conversations and Narrative: A Critical Hermeneutic Orientation in Participatory Inquiry. Westport, CT: Praeger Press.
- KIBEL, B.M.
2000 *Outcome Engineering Toolbox.* Chapel Hill, NC: Pacific Institute for Research and Evaluation.
- 1997 Success Stories as Hard Data: An Introduction to Results Mapping. New York: Kluwer Academic/Plenum Publishers.
- KIRKPATRICK, D.L.
1998 Evaluating Training Programs, Second Edition: The Four Levels.

San Francisco: Berrett-Koehler Publishers, Inc.

OLIAN, J.D. et al.

2000 Designing Management Training and Development for Competitive Advantage: Lessons from the Best. in *Human Resource Planning*. Vol. 21, no. 1: 20-32.

PORTER, M.

KRAMER, M.

1999

Philanthropy's New Agenda: Creating Value, in *Harvard Business Review*, Nov.-Dec. 1999

RITCHEY-VANCE, M.

1998

Widening the Lens on Impact Assessment, in *Mediating Sustainability: Growing Policy from the Grassroots*, Jutta Blauert and Simon Zadek, editors. Bloomfield, CT: Kumarian Press.

SEN, A.

1999

Development as Freedom. New York: Anchor Books.

SENGE, P.M.

1990

The Fifth Discipline: The Art and Practice of the Learning Organization. New York: Currency/Doubleday.

STOLOVITCH, H.

KEEPS, E.

1999

Handbook of Human Performance Technology. San Francisco: Jossey-Bass.

VAN VELSOR, E.

1998

Assessing the Impact of Development Experiences, in, *Handbook of Leadership Development*. The Center for Creative Leadership. San Francisco: Jossey-Bass.

ZADEK, S.

1994

Beyond Fourth Generation Evaluation. Unpublished Material on website, www.zadek.net.