
I. WHY PRINCIPAL PROGRAM EVALUATION IS IMPORTANT

Truly improving the quality of school leaders requires improving the way principals are prepared and selected (Wallace, 2012). Program evaluation can serve many purposes: supporting program improvement, communicating the value of the program to internal and external stakeholders, and holding the program accountable for results. The ability to measure and report impact on outputs (such as placement rates) and outcomes (such as impact on school practices, teacher effectiveness and student achievement) is particularly relevant for programs that want to hold themselves accountable to reaching higher outcomes, while also supporting those that have strong external expectations (from associations, funders, constituents, or the government) for results and improvement.

Many programs, however, have not adequately invested in program evaluation—and when they have the evaluation plans are often weak. Sanzo (2009), in a review of federally funded School Leadership Program grant evaluation plans, found that most evaluations of leadership programs lacked specificity in evaluation tools, data sources, analyses, and many did not have a robust theory of action to guide the evaluation. She concludes with a call for well-planned program evaluations that can provide evaluative feedback to the program as well as inform the field at large.

This paper is part of a [series](#) that describes lessons learned by New Leaders in conducting its own program evaluation. This paper describes the New Leaders approach—which is based on principles of high quality program evaluation—and how evaluation results have informed continuous learning and improvement.

PRINCIPLES OF HIGH QUALITY PROGRAM EVALUATION APPLIED TO PREPARATION PROGRAMS

High quality program evaluation—no matter the subject matter—examines inputs, processes, outputs, and outcomes or impacts (Rossi, Lipsey & Freeman, 2004; Wholey, Hatry & Newcomer, 2010). Both the Education Development Center (EDC) and the University Council of Educational Administration (UCEA) has created a conceptual frameworks for leadership preparation programs that describes the elements that should be measured and possible sources of evidence for measuring them.¹

1. For more information on EDC's framework and accompanying tools, see <http://www.wallacefoundation.org/knowledge-center/school-leadership/principal-training/Pages/Quality-Measures-Principal-Preparation-Program-Assessment.aspx>. For more information on UCEA's framework and accompanying tools, see: <http://www.edleaderprep.org>. All UCEA programs have a robust program evaluation in place, as required by membership in UCEA.

The following diagram is an adaptation of the UCEA Evaluation Pathway framework that New Leaders uses to guide its program evaluation. The accompanying papers in this series are organized by this framework, providing in-depth lessons and recommendations regarding: participant characteristics, program implementation, participant learning and program completion, participant job placement and career advancement, and student outcomes.

Figure 1. Evaluation Pathway for Leadership Preparation Programs (adapted from [UCEA, 2013](#)):



This paper series does not address leadership practices or school and teacher outcomes. While the New Leaders external evaluation by the RAND Corporation does address these components by leveraging our Memorandums of Understanding with districts to access district school climate survey data and to conduct third-party interviews of district partners, they are not addressed by our internal program evaluation. New Leaders has attempted to gather data on leadership practices in the past. However, we found that alumni and teacher survey response rates were too low to generate useful information and that staff visits to schools were too expensive to sustain. We have not been able to gain access to district principal and teacher effectiveness data and have found that publically available school survey data is often unavailable or inconsistent from year to year.

HOW PROGRAM EVALUATION CAN SUPPORT PROGRAM IMPROVEMENT

By tracking all of the elements in the evaluation pathway (inputs, processes, outputs, and outcomes), a well-designed program evaluation can provide formative feedback that informs single loop learning as well as summative findings that inform double loop learning.

Single loop learning includes improved decision-making about program execution based on fidelity of implementation of the program (Argyris & Schön, 1978). For example, New Leaders faculty members immediately use session survey results to inform their planning for sessions the following month (e.g., whether to reteach a particular topic). Recruiting staff use admissions data from previous years to identify gaps (e.g., not enough secondary applicants) and target recruiting efforts accordingly.

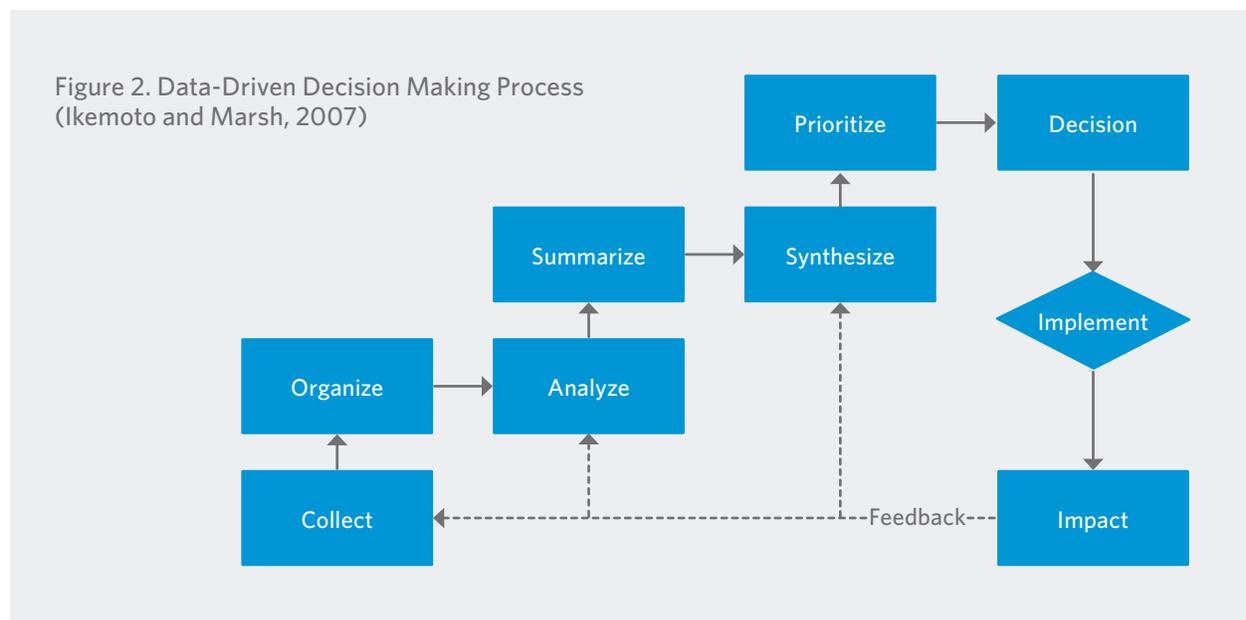
Double loop learning involves improving program design based on using evaluation results to examine the values and assumptions underlying the program design (Argyris & Schön, 1978). For example, New Leaders used results from an external evaluation of impact on student outcomes to question and change its underlying theory of action. While results of the external RAND evaluation have shown statistically significant positive program effects (Gates et al, 2014), New Leaders consistently seeks to improve the

consistency and magnitude of our effect. New Leaders therefore launched a series of efforts to gather more information to help understand and explain the student achievement results. It examined the results of a series of RAND case studies of early tenure principals (Burkhauser et al, 2012), which included an analysis of the relationship between the experiences of those principals and the subsequent school-level effects on student achievement. The organization also ran several analyses of existing data to see if there were any correlations between student outcomes and the characteristics of participants, program elements, or the schools led by New Leader principals. The organization also systematically gathered input and insights from coaches working directly with the principals.

New Leaders learned that the principals who were leading the most successful results were principals who could effectively lead adults. While New Leaders principals demonstrated consistent strengths in instructional leadership, the investigation also suggested that the national admissions process was not effectively screening for these adult leadership skills and coaches were not able to adequately remediate weak adult leadership skills once principals were in placed in the principalship. The organization concluded that it needed to redesign its program to better select and train for adult leadership. It has largely replaced its national admissions process with a teacher leadership program—called the Emerging Leaders Program (ELP)—that gives program participants additional authentic practice leading adults and provides them with feedback on those skills. This design not only develops program participant skills, it also generates scores on the authentic practice-based assignments that enable New Leaders to better assess whether someone is ready to be admitted to the Aspiring Principals Program (APP).

USING EVALUATION DATA TO INFORM IMPROVEMENT

Evaluations provide important information, but information and data alone are insufficient for organizational learning. Programs transform data and information into knowledge when they synthesize, interpret, and triangulate it with other information. This knowledge produces improvement when it informs decisions about design and implementation. Continuous improvement occurs when the cycle repeats itself.



New Leaders ensures that program evaluation informs continuous improvement by dedicating staff resources to analyze and summarize data. New Leaders staff produce analytical reports that are then reviewed by various stakeholders (including program designers and implementers) during meetings (which we call “step back” meetings) at least two to three times per year for each program. During these meetings, program implementers help to identify root causes of patterns observed in the data and the group strategizes on how to make improvements accordingly. For example, during a step back meeting, staff discussed low assignment completion rates. Coaches reported that the assignments were taking participants much longer to complete than the designers had planned. The program therefore reduced the number of assignments for the following year.

RECOMMENDATIONS BASED ON LESSONS LEARNED

Recommendations based on New Leaders’ experiences regarding using evaluation to inform improvement are:

- **Clarify your program’s theory of action.** The theory of action underlying the program needs to be clear in order to effectively identify the indicators that are important to track.
- Ensure your program evaluation **address the entire evaluation pathway**, including inputs, processes, outputs, and outcomes.
- **Generate hypotheses.** In addition to tracking current beliefs in the theory of action, identify factors that *might* be influencers and track them so that you can empirically test those hypotheses in future years.
- **Address single loop and double loop learning** needs by examining fidelity of implementation and questioning program design assumptions.
- **Create processes to translate evaluation data** into organizational learning and program improvement, such as step back meetings.

Other papers in this series provide more detail on New Leaders lessons learned evaluating inputs, processes, and outcomes. For more information, click [here](#).

REFERENCES

- Argyris, C. (1991). Teaching smart people to learn. *Harvard Business Review*, 4(2), 4-15.
- Argyris, C., & Schon, D. (1978). *Organizational Learning: A theory of action approach*. Reading, MA: Addison Wesley.
- Burkhauser, S., Gates, S. M., Hamilton, L., & Ikemoto, G. S. (2012). First-year principals in urban school districts: How actions and working conditions relate to outcomes. Santa Monica, CA: RAND Corporation. Available on-line: http://www.rand.org/pubs/technical_reports/TR1191.html
- Buskey, F., & Karvonen, M. (2012). Evaluating innovative leadership preparation: How what you want drives what (and how) you evaluate. *Journal of Leadership Education*, 11(1).
- Darling-Hammond, L., LaPointe, M., Meyerson, D., Orr, M. T., & Cohen, C. (2007). *Preparing school leaders for a changing world: Lessons from exemplary leadership development programs*. Stanford, CA: Stanford University, Stanford Educational Leadership Institute.
- Elmore, R. F. (2000). *Building a new structure for school leadership*. Washington, DC: Albert Shanker Institute.
- Farkas, S., Johnson, J., Duffett, A., & Foleno, T. with Foley, P. (2001). *Trying to stay ahead of the game: Superintendents and principals talk about school leadership*. New York, NY: Public Agenda.
- Farkas, S., Johnson, J., & Duffett, A. with Syat, B. & Vine, J. (2003). *Rolling up their sleeves: Superintendents and principals talk about what's needed to fix public schools*. New York, NY: Public Agenda.
- Gates, S. M., Hamilton, L. S., Martorell, P., Burkhauser, S., Heaton, P., Pierson, A., . . . Gu, K. (2014). *Preparing principals to raise student achievement: Implementation and effects of the New Leaders program in ten districts*. Santa Monica, CA: RAND. Available on-line: http://www.rand.org/pubs/research_reports/RR507.html
- Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the “data-driven” mantra: Different conceptions of data-driven decision making. In P. A. Moss (Ed.), *Evidence and Decision Making* (Vol. 106, pp. 105-131). Malden, MA: Blackwell Publishing.
- Johnson, J. (2007). *A mission of the heart: what does it take to transform a school?* New York, NY: Public Agenda. Available on-line: <http://www.publicagenda.org/files/missionheart.pdf>
- King, C. (2013). *Quality measures: Principal preparation program self-assessment toolkit*. Waltham, MA: Education Development Center, Inc.
- Levine, A. (2005). *Educating school leaders*. Washington, DC: The Education Schools Project.
- McCarthy, M. (1999). The evolution of educational leadership preparation programs. In J. Murphy & K.S. Louis (Eds.), *Handbook of Research on Educational Administration: A Project of the American Educational Research Association*. San Francisco: Jossey-Bass.
- Murphy, J. & Vriesenga, M. (2004). *Research on preparation programs in educational administration: an analysis*. Columbia, MO: University of Missouri-Columbia, University Council for Educational Administration.

- Murphy, J., Moorman, H., & McCarthy, M. (2008). A framework for rebuilding initial certification and preparation programs in educational leadership: Lessons from whole-state reform initiatives. *Teachers College Record*, 110(10), 2172-2203.
- Orr, M. T., & Barber, M. E. (2005). *Collaborative leadership preparation: A comparative study of innovative programs and practices*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach*, 7th ed. Thousand Oaks, CA: Sage.
- Sanzo, K. L. (2009). An analysis of evaluation plans in a federally funded leadership preparation program: Implications for the improvement of practice. *Planning and Changing*, 43(3/4), 243-259.
- The Wallace Foundation. (2012). *The making of the principal: Five lessons in leadership training*. New York: Author.
- University Council for Educational Administration. (2013). *Developing evaluation evidence: A formative and summative evaluation planner for educational leadership preparation programs*. Charlottesville, VA: Author.
- Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of practical program evaluation* (3 ed.). San Francisco, CA: Jossey-Bass Publishers.

ADDITIONAL RESOURCES ON PROGRAM EVALUATION:

- Designing Evaluations* by the U.S. Government Accountability Office: <http://www.gao.gov/assets/590/588146.pdf>
- W.K. Kellogg Foundation Evaluation Handbook: <https://www.wkkf.org/resource-directory/resource/2010/w-k-kellogg-foundation-evaluation-handbook>
- University of Wisconsin-Extension, Program Development and Evaluation: <http://www.uwex.edu/ces/pdande/evaluation/index.html>
- User-friendly Handbook for Project Evaluation, National Science Foundation: <http://www.nsf.gov/pubs/2002/nsf02057/start.htm>
- United Way Outcome Measurement: <http://www.yourunitedway.org/outcome-measurements>
- Harvard Family Research Project: <http://www.hfrp.org/evaluation>
- McNamara, C. "Basic Guide to Program Evaluation (Including Outcomes Evaluation) Adapted from the Field Guide to Nonprofit Program Design, Marketing and Evaluation." <http://managementhelp.org/evaluation/program-evaluation-guide.htm#anchor1575679>