Evaluating Leadership Development Programs

Leadership development programs (LDPs) vary in length and the type of activities included. OPM, for example, offers courses for aspiring leaders, supervisors, managers, and executives. These programs have the general purpose of helping participants identify their strengths and areas for improvement. Many agencies have implemented LDPs as a way to identify future leaders and provide them the necessary training to help them advance to the next level in their career. Many programs are targeted at specific levels and focus on competencies and skills associated with those levels. For example, a program may be targeted toward GS-9-11, GS-12-13 and GS-14-15 programs.

Most programs provide personal assessment inventories and personality and temperament profiles to help participants identify their personal leadership style and understand how to adapt that style to different situations and audiences. The courses offered depend on the target audience, so they range from basic (e.g., teamwork and collaboration, negotiation and conflict resolution, decision making) to higher-level courses that focus on areas such as creating a vision, leading change, or leading in a crisis environment. The programs also provide a variety of activities, including:

- Assessment (360, personality)
- Competency development
- Seminars
- Coaching
- Mentoring
- Learning portfolios
- Developmental assignments

While a LDP may include all or a subset of these activities, it is important to take the activities into account, as applicable, when evaluating the program. Some activities can be best assessed via survey while others, such as coaching, might more effectively be assessed through interviews with participants' bosses or direct reports. The program should also be evaluated as a whole and not just on the activities. There are many ways to do this, such as the percent of participants that advance to the next grade in a certain period, or the percent of graduates from the program, for example.

Planning for Evaluation

Effective program evaluation requires preparation and careful planning. Creating an evaluation plan will help you align the evaluation objectives with the program objectives, elements (e.g., seminars, learning portfolios), and expectations regarding the impact and outcomes of the training. An evaluation plan will allow you to identify necessary resources and any potential barriers to the evaluation process. It will also give you an opportunity to get involvement from key stakeholders, helping to focus stakeholders' attention on support for achieving the training objectives.

Your evaluation plan should consider the following questions:

- What are the scope, aims, and objectives of the evaluation?

- What is the evaluation time frame?

- Some results cannot be assessed until some months after the program ends. Some assessments require a pre- and post-program administration (e.g., 360-degree feedback). Who will be involved in developing and managing the evaluation process, and how can they be engaged in the process?

- What resources and inputs will be needed?

- What areas of expertise will be needed?

- What will be evaluated, which data will need to be collected, who will provide it, and how and when will it be collected?

- What data analyses will be needed, how will the results be reported, and to whom will they be reported?

- What criteria will be used to judge the success of the evaluation process?

Metrics

It is important to determine your metrics prior to evaluating your program. You can develop metrics to evaluate the individual activities, or the program as a whole. We strongly suggest that you do both. This type of evaluation will help you to understand which training activities are successful and which may need to be replaced or refined. Evaluating the program as a whole will allow you to understand if your LDP is accomplishing your goal, which should be to provide a mechanism for succession planning.

Some examples of metrics were provided above, but metrics should be outcome-oriented. Each metric should outline a desired result with a specific target.

Evaluation Framework

Most leadership development programs can be assessed using a four- or five-step framework. The following is based on Kirkpatrick's four-level evaluation framework. The four levels are:

- Level 1: Learner Reaction and Satisfaction

- Level 2: Learning

- Level 3: Application and Implication

- Level 4: Results or Business Impact

Level 1:
Learner Reaction, Satisfaction

Most courses include learner reaction or satisfaction with the information provided in the course, facilitation, materials, and the learning environment. These are sometimes referred to as "smile sheets." Although there is value in this kind of feedback from the learners, the research so far has indicated that there is little relationship between Level 1 (Reaction) and transfer of learning to the job. When participants consider the information or learning from the class to be useful, the correlation with transfer is slightly higher. However, reaction measures cannot be viewed as surrogates for other valued outcomes, such as learning and results.

By asking the questions below, however, an evaluation can provide information that would be useful in revising the program. The questions could all be answered on a 5-point scale but it is important to provide space for participant comments to get specific information about what worked and what did not.

- How well did the participants like the learning process?
- Is the program relevant to participants' jobs and missions?
- Is the program important to participants' job/mission success?
- Does the program provide new information?
- Do participants intend to use what they learned?
- Will participants recommend the program to others?
- Can the program structure and contents, facilitation, materials, and the learning environment be improved?

Level 2:
Learning

Learning is often assessed by an end-of-course test or a capstone activity. The type of assessment used should mirror the content of the course. If the course or program was primarily factual, then a test that asks participants to apply the learning to new situations would be appropriate. For an LDP, a capstone experience or activity might be more appropriate. The capstone is a culminating activity that provides a way for participants to demonstrate the knowledge and skills they acquired during the program. It might be a group activity that requires participants to apply the knowledge from the class in a simulation of a real-world problem. A capstone exercise for an LDP might have the participant apply a technique or implement a policy learned in the workplace, write a paper describing the rationale behind the action and how the technique was implemented, and then analyze the results.

The learning activity should assess these types of evaluation questions:

- What did the participants learn?
- Are participants gaining knowledge and skills necessary to perform as desired by program developers?
- Do participants know what they are supposed to do with what they learned?
- Do participants know how to apply what they learned?
- Do participants feel confident about applying what they learned?

Level 3:
Application and Implementation

Application and implementation involves evaluating the extent to which participants have applied their new knowledge and skills to their work, the effect this has had on their work performance, and, for LDPs, the effect on the organization. More broadly speaking, it can help (along with Level 4 evaluations) in establishing the business value that the program has added to an organization.

The key questions that evaluators seek to answer at this level include:

- How effectively are participants applying what they learned?
- Were there noticeable and measurable changes in the activity and performance of the leaders and aspiring leaders when back in their workplace?
- Was the change in performance and new level of knowledge or skills sustained?
- Were there any particular barriers to or promoters of the application of learning to the workplace?
- What influence have factors such as the workplace environment, the learning culture, the support of leadership, and the availability of on-the-job support had on the application of learning?

Level 4:	This level of evaluation seeks to measure changes in business performance that have come about through participants applying their new learning to the workplace. Organization leaders may determine the bottom-line value of an LDP. Gathering evaluation data at this level, however, is a complex task that involves measuring the impact of the LDP on business performance measures as reported by learners' managers and other key stakeholders. The specific performance measures used will depend on the individual agency and, importantly, on the agreed objectives and expectations of outcomes from the program. In many organizations some or all of these measures will be in place alongside other performance measures within existing management and reporting systems.
Results or Business Impact	

- Which performance indicators (PIs) and/or other business impact measures are relevant to the LDP?
- Is data on these PIs/business impact measures currently collected for the group, team, department or organization? If not, then data should be collected before the program begins. If this is not possible, then key stakeholders should establish an estimate of pre-learning performance, where possible. In LDPs, 360-degree feedback is often used for pre- and post-assessments of learning.
- When, and over what timeframe, will changes to the PIs/business impact measures be measured? In most cases it takes some time (usually at least three months) before the impact of the LDP on workplace performance becomes evident.
- If you want to be able to calculate the Return on Investment (ROI) from the program, respondents will need to assign financial values to changes in performance, where possible.
- What are the tangible results of the learning process (reduced cost, increased efficiency, improved quality, increased production)?

Isolating the impact of an LDP on business performance is complex because there are likely to be a number of other factors which will have an influence on how the organization performs in any one area at any given time. For example, the organization may have introduced new working practices or performance incentives, or there may be new competitors, legislative, or environmental factors that can influence performance. These are important factors to consider especially when determining causal links between learning and business impacts. And even though there is a timeframe for the evaluation, it is also important to note that learning and its application are continuous and can be assessed as often as possible even months after the program to help ensure that outcomes of the learning program are linked to changes in the business performance.

Overview of Methods to Collect Information

The following table provides an overview of the major methods used for collecting data during evaluations.

Method	Overall Purpose	Advantages	Challenges
Questionnaires, Surveys, Checklists	When you need to get information quickly and/or easily from people in a non-threatening way	• Can be completed anonymously • Inexpensive to administer • Can administer to many people • Easy to compare and analyze • Can get lots of data • Many sample questionnaires already exist	• Might not get careful feedback • Wording can bias responses • They are impersonal • May need a sampling expert • May not get the full story
Interviews	When you want to fully understand someone's impressions or experiences, or learn more about their answers to questionnaires	• More likely to get the full range and depth of information • Develops a relationship with the client • More flexible for the client	• Can take a lot of time • Can be hard to analyze and compare • Can be costly • Interviewer bias
Documentation Review	When you want impressions of how a program operates without interrupting the program (From a review of applications, finances, memos, minutes, etc.)	• Get comprehensive and historical information • Doesn't interrupt the program or client's routine; information already exists • Less likely to produce biased information	• Often takes a lot of time • Information may be incomplete • Need to be quite clear about what you are looking for • Not a flexible means to get data; data is restricted to what already exists
Observation	To gather accurate post-program information about how a participant actually carries out duties	• View operations of a program as they are actually occurring • Can adapt to events as they occur	• Can be difficult to interpret observed behaviors • Can be complex to categorize observations • Can influence behaviors of program participants • Can be expensive

Focus Groups	Explore a topic in depth through group discussion (e.g., about reactions to an experience or suggestion) to better understand common problems or concerns	• Quickly and reliably get common impressions • Can be an efficient way to get a lot of information in a short time • Can convey key information about programs	• Can be hard to analyze responses • Need a facilitator with experience in focus groups and familiar with group dynamics • Difficult to schedule 6-8 people together
Case Studies	To fully understand or depict participants' experiences in a program and conduct a comprehensive examination through cross-comparison of cases	• Fully depicts participants' experience in program input, process, and results • A powerful means to portray a program to outsiders	• Usually quite time-consuming to collect, organize, and describe • Provides depth of information, rather than breadth

Summary

Planning is the key to successful evaluation of your agency's LDP. Creating an evaluation plan will help you align the evaluation objectives with the program's objectives, activities, and expectations regarding outcomes. Your evaluation plan should take into account the desired outcomes, identify what data can and should be collected to measure them, and the proper analyses to conduct to find the results.

Kirkpatrick's evaluation framework has four levels: Level 1: Learner Reaction; Level 2: Learning; Level 3: Application and Implementation; and Level 4: Results or Business Impact. Levels 1 and 2 represent success from the participants' perspective. In Level 1, the data show participants' perceptions of a program with regard to content, delivery, and learning environment. Level 2 data show the extent to which knowledge and skills were acquired. Level 1 and 2 data are usually collected during or at the end of a program. Results reported at these levels can sometimes be used to predict the extent to which knowledge will be used by the participants once they return to the job.

Although you want participants to find the program a positive experience and to learn new thoughts and perspectives, the organization that invests in an LDP should also expect to find positive changes (transfer of learning) from the program to the workplace. That is where Level 3 and 4 evaluations can provide answers. Level 3 data are typically collected several months after program implementation to allow time for participants to begin using their new skills on the job and for the application to have an effect. The results are an indication of the competence an individual displays in specific behaviors and/or skills. Level 3 data serve to measure how well the environment supports the transfer of knowledge and skills to the job. Level 4 data are also collected after the program, and represent the consequence of applying the newly acquired knowledge and skills. These measures focus on output (e.g., productivity), quality, cost, and time. They may also focus on more intangible measures such as customer satisfaction, job satisfaction, and work habits. Level 5 (ROI) calculations are based on Level 4 measures converted to monetary value and account for the fully-loaded cost of the program. Data that may not be converted to monetary value (e.g., improved teamwork and networking) are still important to collect as they represent intangible value.