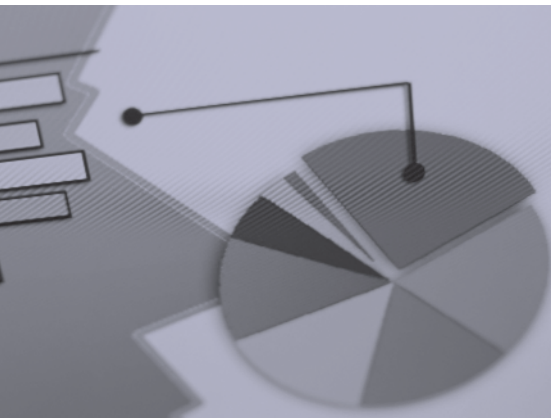


Best Practices

Choosing the Right Methods for Evaluation



When evaluating leadership development it is best to use multiple methods, but which methods should be chosen? Several factors can guide you.

As discussed in the lead article in this issue of *LiA*, the groundwork for evaluating a leadership development initiative is laid by identifying key stakeholders and doing a needs assessment. This enables you to design an evaluation that can effectively gauge the impact of the initiative. A crucial step in design is choosing the methods that will be used to implement the evaluation.

There is a wide variety of evaluation methods available. In what follows I will review several that have been found to be particularly useful in evaluating leadership development. Each has its values and limitations, and these have to be considered in light of the goals of evaluation and the context in which it is implemented.

DAILY EVALUATIONS

In some leadership development initiatives, participants complete evalua-

tion forms at the end of each day. The value of these forms is twofold. First, they give participants an opportunity to reflect on their daily experiences, which reinforces what they've learned. Second, they provide information that enables the people conducting the initiative to make necessary adjustments, thereby enhancing its effectiveness. This method is somewhat limited in that it does not offer participants much time to reflect on their experiences and does not provide a comprehensive picture of the experience.

END-OF-INITIATIVE EVALUATIONS

Participants can complete evaluation forms at the conclusion of each component of the leadership development initiative. You can design these forms to capture the extent to which a specific component met its target objec-

by **Kelly Hannum**

tives, how participants intend to apply what they've learned in the workplace, and how well facilitators, facilities, and logistics met a specified standard. Use these forms to gather evidence regarding how participants intend to use lessons learned, to collect impressions of how relevant and valuable the initiative is to potential participants, and to capture suggestions for changing the initiative. Capturing this information while it is fresh in participants' minds is helpful, but this method doesn't measure the actual implementation of the intended changes—only the intent to apply what has been learned.

INTERVIEWS

Interview questions are typically open-ended and provide qualitative data. They can be conducted face-to-face, via videoconferencing, or by telephone. Interviews can be used to gather information about knowledge, skills or behaviors, and attitudes gained from an individual's experience with a development initiative. They assess perceptions of the initiative from a stakeholder's perspective. If resources don't allow time for one-on-one interviews, you can adapt the interview format to an open-ended questionnaire that you can mail or e-mail to participants. Interviews are most effective when you want qualitative information to fulfill multiple purposes:

- Assisting in identifying training and learning needs, an initiative's design, or expectations for applying lessons from the development initiative (interviews would take place before the initiative).
- Determining participants' reactions, experiences, and satisfaction with the initiative and how they applied what was learned (interviews would take place during a long-term initiative).
- Determining participants' reactions, learning, and intentions to

apply their learning (interviews would take place during the initiative or right after it had ended).

- Developing surveys, focus-group interview questions, or the focus of an observation (interviews could take place at any point relative to the initiative, depending on the outcomes desired).
- Further interpreting survey results (interviews are typically most relevant for this purpose after an initiative has been completed).

Interviews have several advantages and disadvantages. One advantage is that they allow the evaluator to probe for clarification and deeper complexity, which results in richer data. They provide time for participants and stakeholders to reflect, which can be developmental in itself. If the evaluator (as opposed to the organization's HR department, for example) conducts the interviews, acceptance and endorsement of the evaluation may increase because participants and stakeholders often view the evaluator as an objective party.

A significant disadvantage is that interviews can be more costly than other methods. If they are conducted in person, for example, travel costs add to the total price of the evaluation. Interviews can also take more of the evaluator's time, in terms of both conducting them and analyzing the qualitative data obtained. Evaluators can reduce the pressure on resources by conducting phone interviews or by using Internet-based collaboration tools (such as electronic whiteboards or chat rooms), which reduce travel costs and also allow evaluators to conduct more interviews in a shorter amount of time.

LEARNING SURVEYS

A learning survey can assess the extent to which participants have absorbed new content during the initiative. This method is valuable when participants are expected to retain

factual information (such as their organization's leadership or competency models or its business policies or practices) or learn specific steps for implementing leadership responsibilities (such as giving feedback and coaching others).

One way to assess the attainment of factual information is to administer the learning survey twice: once before the initiative and once immediately afterward. If possible, conduct a pretest before the initiative as a means of assessing the needs of participants and stakeholders; this can guide you in focusing the measures you will use in the evaluation.

The learning survey's questions should be related to the content of the initiative. To analyze the data, you can compare the responses of the two surveys. Participants may have some knowledge of the content before the initiative, and they should know most if not all of the content after the initiative ends.

CHANGE SURVEYS

Change surveys are useful in assessing whether change has occurred as a result of a development initiative. They are typically used to measure changes in attitudes or behaviors specific to the initiative in question. A well-developed change survey should be based on what is already known about the impact of the initiative and the objectives of the initiative.

ABOUT THE AUTHOR



Kelly Hannum is a research associate at CCL in Greensboro. She holds a Ph.D. degree from the University of North Carolina at Greensboro.

Change surveys are most effective when the focus of the evaluation is on behavioral changes as measured by quantitative data. Administering a survey at different points in time will fulfill different purposes. For example, using a change survey before an initiative allows the evaluator to establish baseline information about the behaviors that the initiative is expected to change. When administered several months after the initiative has ended, a change survey can provide data to be compared with the results of the first survey in order to measure the actual behavioral change resulting from the initiative.

Valid, reliable change surveys can be time consuming to develop but relatively inexpensive to administer. They allow responses from a large sample of people and can collect easily analyzed quantitative data. They allow anonymity or confidentiality of responses, are useful when respondents are geographically dispersed, allow people to respond on their own time, and require that all respondents answer the same set of questions. On the minus side they don't allow changes to or clarification of questions, and it may be difficult to get the desired response rate.

Intuition can be misleading when developing a survey. Even seemingly simple choices, such as what response options to use, can have an unintended impact on results. If you plan to develop your own survey and do not have training in survey development, it is wise to seek advice from measurement or psychometric publications or from measurement professionals.

BEHAVIORAL OBSERVATION

Behavioral observation involves observing a set of activities, the people who are participating in those activities, and the environment in which the activities take place. Observations can produce qualitative data in the form of field notes or

quantitative data if observers note their information as ratings, rankings, or frequencies.

There are several effective uses for behavioral observation. Before an initiative begins, observation can determine participants' baseline knowledge, skills or behaviors, and attitudes. During the initiative it can determine levels of interaction, engagement, skill development, and satisfaction with the initiative. Observations conducted between one week and three to six months after completion of the initiative can determine changes in knowledge, skills or behaviors, and attitudes. Behavioral observation is especially effective when it is used as one of several data collection methods.

Compared with other evaluation methods, behavioral observation has several advantages. For example, data are collected where the activity is taking place, thus enhancing data validity; target activities are viewed in a context that may help you interpret data collected from other methods; a trained observer might see things that others close to the initiative may miss; and the observation process can illuminate issues that interviewees are unwilling to talk about.

This evaluation method also has some limitations. It requires a well-trained observer. Multiple observers may focus on different things, thus making analysis and synthesis more difficult. Also, participants may alter their behavior if they know they are being observed. Finally, behavioral observation can be disruptive to the work environment and can be expensive if a large sample is required.

FOCUS GROUPS

You can use a focus group to gather information from six to twelve people at one time. The primary purpose of this method is to obtain qualitative information from a group of individuals (or a team) that has had a similar

experience (participation in a training program, for example). Evaluations usually make use of multiple focus-groups and employ a well-designed guide to focus the discussion. Focus groups should be carried out in a way that allows participants to feel safe disclosing information about their attitudes and perceptions about the initiative being evaluated.

There are several ways you can use focus groups in your evaluation. Before an initiative begins, you can use them to identify training or learning needs, to determine an initiative's design, or to assess expectations of how participants will apply what they learn. During an initiative (assuming it takes place over several days, weeks, or months), you can use a focus group to determine participants' reactions to and experiences and satisfaction with it. After the initiative ends, you can use this evaluation method to determine participants' reactions, what they've learned, how they intend to apply their new skills and knowledge, and the relative success or challenge they've had in doing so. Focus groups are also effective when it's necessary to assess reactions, learning, and intentions of a team rather than individual participants. Some evaluators use focus groups to develop survey questions or to further interpret survey results. Another interesting product of focus groups is that they allow participants to process their experiences together, which can help them build support networks to further aid in their development.

The focus-group data-collection method has several advantages. Because it captures the collective experience of individuals, interaction among participants tends to increase the number and quality of responses. It provides a forum in which teams can create additional impact and meaning related to their developmental experience. It lets the interviewer probe for clarification, explanations, and examples. Participants generally

enjoy being part of a focus group, resulting in high response rates. This method is relatively inexpensive and makes good use of time because it allows evaluators to collect data quickly from a large group of people.

The focus-group method also has some limitations. A skilled interviewer is required to ensure that the data collected are of high quality. The interviewer has less control in a group interview than in an individual interview and so needs to have the skills and ability to keep the group on track. Data collected from a focus group may be difficult to capture and organize. Groups vary widely—some may develop a collective energy and provide extensive data whereas others may lack energy and provide only superficial data.

DIALOGUE

Group dialogue is a special kind of conversation in which people listen intently for underlying meanings and assumptions. Dialogue requires that participants suspend their assumptions in a way that enables them to hear others' perspectives objectively. This technique allows an open, creative conversation to take place, often freeing the participants in the conversation to become aware of and better understand different perspectives. Unlike a focus group, a dialogue allows the evaluator to remain relatively invisible (at his or her choosing) and to let the conversation take place among group members.

The value in this method is that it allows team members and groups of participants to interact with each other, exploring their perspectives and insights more deeply than they might in a focus group. You can use

Editor's note: To learn more about dialogue, see "Conversation Piece: Using Dialogue as a Tool for Better Leadership," Leadership in Action, March–April 2003, 23(1), 8–11.

dialogue in your evaluation to understand the different perspectives of stakeholder groups and to clarify what the collected data mean.

MEASURING RETURN ON INVESTMENT (ROI)

Return on investment is a technical phrase that refers to financial considerations. However, it is also often used in a nontechnical sense to describe the broader impact of individual and team development initiatives on an organization. Several methods are available to evaluate such impact. Determining which is most appropriate for a particular evaluation depends on the type of impact expected. To identify the appropriate measure at the organizational level, some useful questions to ask include the following:

- What changes at the organizational level are expected as a result of the initiative?
- Why are these changes expected? (In other words, what is the connection between the objectives of the leadership development experience and these changes?)
- When can stakeholders in the organization expect to see changes?
- Who will be able to note and report on these changes?
- How can data about these changes be obtained?

True measures of ROI include data such as the costs of facilities, trainers, materials, and the time participants spend in training and away from their jobs. ROI formulas also include the financial benefits of training, such as cost savings, new revenue, and calculations of the value of perceived job improvement. A critical step in creating accurate ROI formulas is isolating and measuring the effects of development.

Although ROI formulas work well for skills-based training and many organizations apply them to leader-

ship development, the value of this method for evaluating leadership development is somewhat limited. It doesn't provide data related to the quality of improvement and to nonfinancial benefits such as enhanced communication and better delegation. For example, an ROI of 300 percent is impressive, but without data showing where the improvement has been made it's not possible to fully measure the impact of that improvement. (For more information on ROI, see Jack J. Phillips, *Return on Investment in Training and Performance Improvement Programs* [Gulf Publishing Company, 1997]).



WORKPLACE STATISTICS

Workplace statistics include information that organizations often gather on absenteeism, communication breakdowns, and grievances. Evaluators can obtain that information from the organization and analyze it in relation to the leadership development initiative and its objectives. Before requesting and using workplace statistics in an evaluation project, carefully determine which statistics are likely to change as a result of an individual's or team's participation in the initiative.

Workplace statistics are useful as an evaluation method when a development initiative is designed to meet specific organizational outcomes. For example, when participants are expected (as part of an action plan based on their development experience) to work differently with at-risk employees to prevent situations that

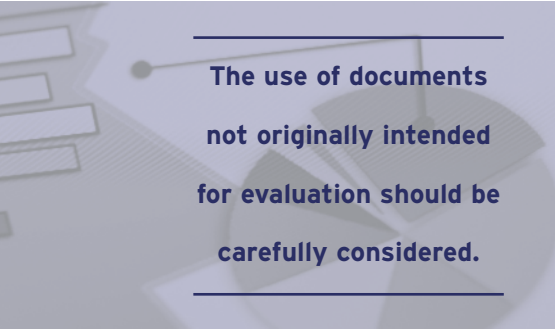
might lead to grievances, it's useful to examine the change in the frequency and severity of the grievances that employees file.

Workplace statistics aren't useful if there are no links between the statistics, the expected organizational outcomes, and the development initiative.

DOCUMENT ANALYSIS

Documents and records are written statements and other materials that attest to an event or provide an accounting of some activity.

Evaluators who use documents typically wish to make inferences about the values, sentiments, intentions, or beliefs of the sources or authors.



**The use of documents
not originally intended
for evaluation should be
carefully considered.**

Documents may include letters, journals, logs, position papers, notes, speeches, newspaper articles and editorials, annual reports, newsletters, case studies, evaluations or consultants' reports, and photographs. Records are typically used to keep track of events or transactions. They may include expenditure records; expense account vouchers; financial information; performance records; professional, business, or school directories; student or participant achievement or performance test records; state or federal regulatory records; attendance records; competency or attitude score records; and electronic mail records.

This method can be effective in understanding the history or background of a program or situation; in tracking the number of times that

something has occurred; in developing survey or interview questions or areas on which to focus an observation; in identifying patterns of participation in, interest in, or attitudes toward a program; and in better understanding an issue that people are unable or unwilling to talk about.

The use of documents not originally intended for evaluation should be carefully considered. Some documents may be seen as private and not appropriate for evaluation use.

Document analysis has several advantages as a method of data collection. Documents and records are often plentiful (stored usually in the organization's archives) and inexpensive to collect. Data from documents and records may provide useful chronological detail (possibly more accurately than a chronology built from interviews). The data gathered from documents and records are less subject to an evaluator's bias than data from interviews and observation. When used with other methods, documents and records provide contextual richness for interpreting other data collected during the evaluation.

There are also some limitations. For example, samples of documents or records may not be representative (notes may exist for some meetings but not for others). Personal documents may reflect a person's bias. Deliberate deception or manipulation of information is possible if the document's writer had a desire to express a certain viewpoint. Also, records may contain incomplete or out-of-date information.

ANALYSIS OF ORGANIZATIONAL SYSTEMS AND PROCESSES

At the organizational level, one type of outcome evaluators can look for is the extent to which organizational systems and processes—aspects of

the organization that directly affect employees—have been affected. Many leadership development initiatives will result in improvements to various organizational systems and processes. Initiatives that are most likely to have this effect include components that enable teams to work together more effectively, enable individuals to understand others' points of view more readily, or stimulate individuals and teams to more effectively and creatively problem solving and decision making. Some systems and processes for you to consider include operating procedures, educational processes, HR policies, formal and informal communication structures, financial accounting processes, and maintenance procedures.

APPRAISAL OF CUSTOMER SATISFACTION

When measuring the value of a development initiative, evaluators often find it useful to get the perspective of the customer. Although it is easy to understand this connection when the development initiative is directly related to customer service, it is not always as clear when the initiative focuses on developing individual leaders. Just how might such an initiative affect an organization's customers? Some components of effective leadership (communication and decision making, for example) have a direct effect on the relationship organizations have with customers and clients. For example, an organization whose leaders are unable or unwilling to understand what its clients need may see those clients move to a more responsive competitor.

Whether developing your own customer survey or interview or working with a vendor, you should investigate several important aspects of customer service as a means of measuring the impact of a development initiative. For instance:

- What do customers expect from their relationships with the organization?
- How long does it take for a customer problem to be solved or complaint to be addressed (after the development initiative compared with before the initiative)?
- Can the customer describe past experiences (before the development initiative) with the organization?
- Can the customer describe current experiences (after the development initiative) with the organization?

CLIMATE SURVEY RETESTS

When leadership development is a component in an organization's efforts to make significant improvements or changes to its work climate, it's useful to examine the extent to which that climate has changed over the course of the development initiative. This can be accomplished by making at least two assessments of the organization: one before the beginning of the initiative and one at an appropriate time after the initiative has ended. Many organizations conduct annual climate surveys as a general practice. The appropriate time to administer subsequent assessments depends on the type of climate change expected, the size and complexity of the organization, and the number of individuals who have participated in the initiative. Change in larger, more complex organizations takes more time than in less complex situations.

Organizational climate is typically defined by employees' satisfaction with specific features, such as pay and benefits, leadership, and opportunities for development. As with other measures of organizational improvement, climate change should be examined only if the leadership development initiative was designed to effect this type of large-scale change. The initiative should include a sufficient number of employees (the

appropriate number depends on the type of initiative and the change desired), should be of long enough duration to create change at the organizational level, and should be designed to encourage organization-level changes.

CULTURE SURVEY RETESTS

Culture in organizations is typically defined as the behaviors and values that pervade the organization. It determines, among other things, whether a particular behavior, appropriate in one organization, is deemed inappropriate in another. The culture of an organization is built by its people and is extended through its selection, attrition, reward, and reprimand processes. Organizational culture is greater than any one individual's values or behaviors.

When leadership development is a component in an organization's efforts to make significant improvements or changes to its culture, it's useful to examine the extent to which that culture has changed over the course of the leadership development initiative. You can make this part of your evaluation by making at least two assessments of the organization's culture—one before the beginning of the initiative and one or more at appropriate intervals after the initiative has ended. When to administer subsequent assessments depends on the type of change expected, the size and complexity of the organization, and the number of people who have participated in the initiative.

Additional administrations of assessments, including administrations during the initiative, may provide stronger evidence of impact.

Culture is more embedded in the organization than is climate, and so is more difficult to change. Assessments of culture change are unlikely to reveal significant data more than once a year. As with organizational climate change, the success and speed of culture change depend

largely on the size and complexity of the leadership development initiative, as well as on the positions and percentages of employees involved in the initiative.

Like other measures of organizational improvement, culture change should be examined only if the leadership development initiative was designed to effect this type of large-scale change. The initiative should include an appropriate number or percentage of employees from each organizational level capable of effecting the desired culture change, should last long enough to create change at the organizational level, and should encourage organization-level changes.

MULTIPLE METHODS

When doing an evaluation of a leadership development initiative, it is best to use a combination of methods so that you can assess outcomes at different levels. In regard to the methods just described, individual outcomes are most appropriately assessed by daily evaluations, end-of-initiative evaluations, interviews, learning surveys, change surveys, and behavioral observation; group or team outcomes by focus groups and dialogue; and organizational outcomes by ROI, workplace statistics, document analysis, analysis of organizational systems and processes, appraisal of customer satisfaction, climate survey retests, and culture survey retests.

In choosing techniques you must pay attention to the leadership needs addressed by the initiative, the type of data needed to answer critical evaluation questions, the type of data preferred by stakeholders, and the levels of mastery targeted (as discussed in the lead article in this issue).

If these issues are carefully considered in the selection of evaluation methods and in their implementation, the likelihood of a successful evaluation will be greatly increased. ✍