The following discussion provides general guidelines for measuring training program effectiveness using the Kirkpatrick model for training evaluation.
Contents

Contents ......................................................................................................................................... 2
Evaluation Levels ............................................................................................................................ 3
  Level 1 – Reactions ........................................................................................................................ 3
  Level 2 – Learning .......................................................................................................................... 3
  Level 3 – Behavior ......................................................................................................................... 3
  Level 4 – Results ............................................................................................................................ 3
Continuous Performance Improvement ............................................................................................ 3
Training Assessment Measures ......................................................................................................... 5
Performance Evaluation Measures .................................................................................................... 6
Guidelines ...................................................................................................................................... 7
  Determine the Purpose .................................................................................................................. 7
  Develop a Measurement Plan .......................................................................................................... 7
  Decide on reason(s) for Evaluating Performance ........................................................................... 9
  Determine Criteria That Supports Target Metrics .......................................................................... 9
  Create method of evaluating performance criteria ......................................................................... 9
    1. Instrument ............................................................................................................................... 9
    2. Scales ..................................................................................................................................... 10
    3. Raters .................................................................................................................................... 11
    4. Rating Periods ......................................................................................................................... 11
Maintaining the Evaluation Program ............................................................................................... 11
Kirkpatrick’s four levels of evaluation are often used to assess training a training program’s effectiveness. According to this model, evaluation should always begin with level one, and then, as time and budget allows, should move sequentially through levels two, three and four. Information from each prior level serves as a base for the next level’s evaluation. Thus, each successive level represents a more precise measure of the effectiveness of the training program, but at the same time requires a more rigorous and time-consuming analysis.

Level 1 – Reactions

Also known as “smile sheets” this type of evaluation mainly collects the learners’ reaction to their training experience, from the comfort level of the training environment, to their estimation of their level of knowledge prior to training and after training.

Level 2 – Learning

Assessing at this level moves the evaluation beyond learner satisfaction and attempts to assess the extent students have advanced in skills, knowledge, or attitude. This is often done by giving students a Pre-test before the class and a Post-test after the class. The results of each are compared to assess the degree of advancement the student made between the two.

Level 3 – Behavior

This level measures the transfer that has occurred in learners’ behavior due to the training program. Evaluating at this level attempts to answer the question: Are the newly acquired skills, knowledge, or attitude being used in the everyday environment of the learner? This assessment is often conducted at a set interval after training has occurred, e.g., two weeks to six months after the training event. Information for this level of evaluation is often collected via surveys and focus groups.

Level 4 – Results

Level 4 evaluations are also known as Return on Investment (ROI). This type of evaluation measures the success of the training program in terms that managers and executives can understand: increased production, improved quality, decreased costs, reduced frequency of accidents, increased sales, and even higher profits.

Read more at: http://www.kirkpatrickpartners.com

Continuous Performance Improvement

The levels also relate to the different stages of measuring the effectiveness of training of a particular capability. When an organization is confronted with the challenge to measure the value of a training program, there are two broad levels of measurement, Training Assessment and Performance Evaluation.

Training Assessment: used to determine the quality (i.e., retention) and accuracy (i.e., efficacy) of the training program. This information is gathered in three different stages: Pre-Training, Post-Training and Skill Retention. Data for all three stages can be easily determined by conducting surveys or assessments.
of the training attendees. The execution of the Training Assessment Program is completed in the “Release” stage of implementation.

**Performance Evaluation:** attempts to measure the degree in which training is transferred to a particular job(s) as well as measuring the increase in organizational performance. Determining the degree to which training has an influence on organizational performance is the most desired level when measuring training effectiveness. However, it is usually the most difficult to accomplish. It is the most difficult to measure as there are other variables in addition to training that affect job behavior, and ultimately organizational performance. To obtain measures of job behavior and organizational performance, a valid performance evaluation is required. Without an accurate performance evaluation system, there is no accurate way to measure the effect of training in the organization. The execution of the Performance Evaluation Program is completed in the “Maintenance” stage of implementation.

The Continuous Performance Improvement pyramid represents a comprehensive platform by which to measure the effectiveness of a training program. Measuring the value of the training, at any level, often involves using the key performance indicators (e.g., faster and more reliable output, higher ratings on employees’ job satisfaction questionnaires). Please refer to the Training Assessment Measure and the Performance Evaluation Measures below which describe the Target Metrics for evaluating the effectiveness of a training program.
# Training Assessment Measures

<table>
<thead>
<tr>
<th>HP Categories</th>
<th>Possible HP Target Metrics</th>
<th>Data Collection Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Training Assessment</td>
<td>• The number of people who indicate a need for training during the needs assessment process</td>
<td>Survey Pre-test</td>
</tr>
<tr>
<td></td>
<td>• The number of people that sign up for the training session</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Attitudes toward training and change</td>
<td></td>
</tr>
<tr>
<td>Post-Training Assessment</td>
<td>• The number of people that attend the session</td>
<td>Survey Post-test</td>
</tr>
<tr>
<td></td>
<td>• Attendee satisfaction at end of training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Attendee satisfaction at end of training when attendees know the actual costs of the training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• A measurable change in knowledge or skill at end of training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ability to solve a “mock” problem at end of training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Willingness to apply or intent to use the skill/ knowledge at end of training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Popularity (attendance or ranking) of the program compared to others (for voluntary training programs)</td>
<td></td>
</tr>
<tr>
<td>Skill Retention</td>
<td>• Attendee satisfaction at X weeks after the end of training</td>
<td>Survey Post Test</td>
</tr>
<tr>
<td></td>
<td>• Attendee satisfaction at X weeks after the training when Attendees know the actual cost of the training</td>
<td>Observation</td>
</tr>
<tr>
<td></td>
<td>• Retention of knowledge at X weeks after the end of training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ability to solve a “mock” problem at X weeks after end of training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Willingness to try (or intent to use) the skill/knowledge at X weeks after the end of training</td>
<td></td>
</tr>
</tbody>
</table>
## Performance Evaluation Measures

<table>
<thead>
<tr>
<th>HP Categories</th>
<th>Possible HP Target Metrics</th>
<th>Data Collection Method</th>
</tr>
</thead>
</table>
| Job Behavior Change | • Trained individuals that self-report that they changed their behavior/ used the skill or knowledge on the job after the training (within X months)  
• Trained individuals who’s managers report that they changed their behavior/ used the skill or knowledge on the job after the training (within X months)  
• Trained individuals that actually are observed to change their behavior/ use the skill or knowledge on the job after the training (within X months)  
• Rank of training seminar in forced ranking by managers of what factors (among miscellaneous staff functions) contributed most to productivity/ profitability improvement | Survey  
Behavior-based Performance Evaluation |
| Organizational Performance | • Trained individuals that self-report that their actual job performance changed as a result of their changed behavior / skill (within X months).  
• Trained individuals who's manager's report that their actual job performance changed as a result of their changed behavior / skill (within X months).  
• Trained individuals who's manager's report that their job performance changed (as a result of their changed behavior / skill) either through improved performance appraisal scores or specific notations about the training on the performance appraisal form (within X months).  
• Trained individuals that have observable / measurable (improved sales, quality, speed etc.) improvement in their actual job performance as a result of their changed behavior / skill (within X months).  
• The performance of employees that are managed by (or are part of the same team with) individuals that went through the training.  
• Departmental performance in departments with X % of employees that went through training ROI (Cost/Benefit ratio) of return on training dollar spent (compared to our competition, last year, other offered training, preset goals etc.).  
• CEO / Top management knowledge of / approval of / or satisfaction with the training program. | Survey  
Behavior-based Performance Evaluation |
Guidelines

**Determine the Purpose**

Careful consideration of the purpose for which a solution is being developed is the first step toward accurate evaluation. Both extremes can be found within organizations. It is important to have balance and provide expectations on collecting such data. Too little and too much evaluation are both wasteful.

1) Determine if the measurement variables are organizational or individual. If the variables are organizational goals, refer to Develop Performance Evaluation Program (WBS 3.4).

2) Identify the measurement variables that are individuals rather than organizational.
   a) If variables are to measure individual traits, you must be able to prove that the traits accurately predict individual performance and that the solution used to impart knowledge is valid.
   b) Questions to ask are:
      i) Does the solution provide the skills necessary for employees to successfully complete the job into which they will be promoted?
      ii) Can you predict that employees’ professional development will increase as a result of successfully completing the develop activities
      iii) Does the training predict improved performance from successful completion?
      iv) Does the training result in self-development on the basis of the individual’s own perceptions?

**Develop a Measurement Plan**

A measurement plan details the specifications of any measurement instruments included in the training program. Tests, Quizzes, questionnaires, and surveys must all be developed according to the purpose of evaluation.

If developing questionnaires or surveys, they must attain a minimum of content validity to be useful. Surveys present some difficulties with respect to being sure enough are returned to consider the resulting sample adequate. It is very optimistic to expect to get a 50% return rate on surveys.

3) Determine what to measure using the goals and objectives of the project. Decide if these goals and objectives should measure skills, knowledge or attitudes.

4) Develop your measurement instrument based on exactly you want to measure. Instruments can be one or a combination of:
   a) Matching Questions
   b) Multiple-choice Questions
   c) True-False Questions
   d) Completion (Short-Answer) Questions
   e) Simulation, Role Play, or Performance Tests
   f) Essay Questions

5) Calculate the length of each instrument

6) Calculate the weight of each item

7) Overall, all training assessments will occur during Stage IV: Release. However, it is important to determine if measure should be interspersed throughout the program or use a comprehensive test at the end. The following table will assist in the placement of measures.
<table>
<thead>
<tr>
<th>Type</th>
<th>Location</th>
<th>Level</th>
<th>Purpose</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embedded (Study)</td>
<td>Within a Lesson</td>
<td>Basic knowledge and understanding</td>
<td>Check basic knowledge of concepts(s) taught in the lesson</td>
<td>Interspersed throughout a lesson</td>
</tr>
<tr>
<td>Quiz</td>
<td>At the end of each lesson or periodically after several lessons</td>
<td>Knowledge of information within a lesson</td>
<td>Test student’s grasp of concepts</td>
<td>End of lesson or group of lessons</td>
</tr>
<tr>
<td>Application</td>
<td>Note: Questions should test only the major concepts from the objective(s) that the lesson covers</td>
<td>Note: At least one question should be asked on each major topic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>Interim and end of course</td>
<td>Application</td>
<td>Tests students’ knowledge or performance by integrating information from throughout the course</td>
<td>As required</td>
</tr>
<tr>
<td>Evaluation</td>
<td></td>
<td></td>
<td>Note: Test pool should contain at least one question from each lesson objective</td>
<td></td>
</tr>
</tbody>
</table>
Decide on reason(s) for Evaluating Performance

All organizations are different and will have different reasons and/or philosophies for evaluating performance. Although, organizational terminology may vary, most employee evaluations programs will be focused for one or more outcomes:

<table>
<thead>
<tr>
<th>Reason</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compensation</td>
<td>Typically the difference between two employees within the same job is tenure and job performance. Performance evaluations provide a fair basis upon which to determine an employee’s salary.</td>
</tr>
<tr>
<td>Promotions</td>
<td>Care should be taken to ensure that the employees are evaluated on job dimensions that are similar to the new position.</td>
</tr>
<tr>
<td>Employee Feedback and Training</td>
<td>A typically overlooked use of performance evaluations is to develop employees. These evaluations provide feedback on strengths and areas of improvement for employees. Training effectiveness can also be evaluated training programs.</td>
</tr>
</tbody>
</table>

Determine Criteria That Supports Target Metrics

Typically, the criteria you will need to evaluate are the same criteria that were highlighted during the recruiting and selection process. If the client is uncertain of what to measure, asking the following questions will assist in providing direction:

1. What part of an employee’s job are you trying to change?
2. How are you measuring the success of the IMPACT implementation?
3. What type of performance is important to the organization?
4. What is the prediction for organizational growth over the next five years?

Create method of evaluating performance criteria

There is a plethora of different types of performance evaluation tools, techniques, and forms. Performance evaluation instruments typically focus on one or a combination of traits, behaviors, and results. Before an actual performance evaluation instrument is created, three important decisions must be made:

1. What is the focus of the instrument?
2. What type of rating scale do you want to use?
3. Who is going to rate the performance?
4. How many time periods should the employee be rated?

Below are some guidelines for answering each of the above items. It is important to note that there is no one perfect system and the client may need to be guided through this process using the information below.

1. Instrument
   
   **Traits**
   
   Almost always, traits are poor instruments for measurement and should not be used. Traits provide poor feedback to the employee and are difficult to translate into development and growth initiatives.

   **Behaviors**
   
   Traits focus on who an employee “is,” as where behaviors concentrate on what an employee actually “does.” This measure provides specific feedback that employees can translate into new practices, which should improve performance. Using the behavior measure also reduces an
employee’s defensiveness during the evaluation meeting as well as many legal problems associated with performance evaluations.

Results
Lastly, tools which focus on results tend to concentrate on what an employee accomplishes during a given time period. Results based performance evaluation systems work to directly relate an employee’s contributions to the organizations bottom line. Remember to be extremely cautious when using this type because many times, employees may do everything asked of them and factors outside their control directly affect the outcome. If the client wishes to use this type of scale, it would be best to conclude some type of accompanying behavioral based scale.

Results criteria are derived from some type of hard and/or objective data. Typical data consists of quality of work, quantity of work, attendance, and safety measures.

2. Scales

There are multiple rating scales that can be used when developing a performance evaluation. The following are a list of and definition of each.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graphic Rating Scale</td>
<td>Simple Likert type scale where number ranging between one to five or one to seven accompanied by words such as poor and excellent anchoring each end.</td>
</tr>
<tr>
<td>Behaviorally Anchored Rating Scales (BARS)</td>
<td>Similar to the scale above, BARS uses samples of behavior describing each rating number along the scale.</td>
</tr>
<tr>
<td>Forced-Choice Rating Scales</td>
<td>Similar to BARS, scales have specific behavior examples, except the scale points are hidden. This allows for the rater to choose the behavior that is most typically performed by ratee.</td>
</tr>
<tr>
<td>Mixed-Standard Scales</td>
<td>List of behavioral examples representing various (usually three) levels of performance on multiple dimensions. Raters then assign a (+), (0) or (-) for each example. Overall performance scores are obtained by summing the score from each scale.</td>
</tr>
<tr>
<td>Rank Order</td>
<td>Comparing employees to each other by placing employees in an order based on performance of relevant dimensions. Do not provide information on how an employee is actually doing.</td>
</tr>
<tr>
<td>Paired Comparison</td>
<td>Comparing each possible pair of employees and choosing which employees is better based on the pair match. This scale is not feasible with large organizations. Do not provide information on how an employee is actually doing.</td>
</tr>
<tr>
<td>Forced Distribution</td>
<td>Assigning a predetermined percentage of employees into one of five categories for each dimension. Do not provide information on how an employee is actually doing.</td>
</tr>
<tr>
<td>Behavioral Observation Scales (BOS)</td>
<td>Similar to BARS, specific behaviors are rated on the frequency with they occur for each employee. Provide high levels of feedback and are better for motivating behavioral changes.</td>
</tr>
</tbody>
</table>
3. Raters

Typically, performance evaluations are a top-down communication method for the ratee. In other words, feedback is given by the ratee’s supervisor or manager. However, using multiple sources other employees can provide a more accurate picture of an employee’s performance. Other raters can consist of peers, subordinates, and even self-appraisal.

4. Rating Periods

Most organizations will conduct performance evaluations one time a year. Feedback is a critical part of an employee’s development, especially when they are learning a new skill such as IMPACT. If at all possible, performance evaluations should at least be bi-annually. If feedback is less frequent, it may take longer for IMPACT to be utilized throughout the organization.

Maintaining the Evaluation Program

A Metrics Evaluation Program is a program by which Human Performance, Application and Technological information is gathered, and target metrics (identified during the planning stage) are applied. This process ensures that each change that is adopted by the capability is extrinsically tied to the overall target metrics, and therefore supports the overall vision of the clinical operations.

The main steps in maintaining the evaluation program are as follows:

1. Gather Human Performance Information
2. Gather Application Information
3. Gather Technological Information
4. Apply Target Metrics (continuous, systematic)
5. Identify Best Practices and/or Opportunities for improvement
6. Assess the Effect of Change
7. Assess the Technological Effect of Change
8. Incorporate the Change in the Change Management Process