Advanced Industrial Psychology
Industrial Training

Training Evaluation

What We Will Cover in This Section
• What is evaluation?
• The concept of the CRITERION.
• Evaluation procedures.
• Evaluation techniques.

Evaluation
Systematic collection of descriptive and judgmental information to make an assessment of instructional activities.
General Focus of Evaluation

1. Has change occurred on the job?
2. Was this change caused by training?
3. Will the same change happen to new participants?
4. Will the same change happen in other organizations?

Barriers to Evaluation

1. Management does not emphasize it.
2. People don’t have the skills to do it.
3. People don’t know what to evaluate.
4. People may fear the results.

Benefits of Evaluation

1. Determine where a program needs to be changed.
2. Evaluate acquisition, retention, and transfer.
3. Identify potential legal issues.
4. Evaluate the trainer.
5. Determine the overall benefit to the organization (UTILITY).
Criterion

A standard for assessing effectiveness, success, or failure of an activity.

Criterion As a Test

- RELIABILITY
  - Is the assessment done consistently.
    - Rater reliability.
    - Internal Consistency.

- VALIDITY
  - Are we measuring what we say we are measuring.
    - Face Validity.
    - Content validity.

Unidimensional vs. Multidimensional Criteria

UNIDIMENSIONAL CRITERION
Single overall indicator of success.

MULTIDIMENSIONAL CRITERION
Multiple indicators, factors, or facets of success.
Criterion Model

Needs Assessment  Criterion Data

Criterion Deficiency  Criterion Contamination

Criterion Relevance

Criterion Contamination

- OPPORTUNITY BIAS
  - People in different groups are treated differently.
- GROUP CHARACTERISTIC BIAS
  - Differential opportunities to use KSACs on the job.
- RATER REACTIVITY
  - Raters are biased on the basis of their knowledge of a person's training performance.

Levels of Criteria
Kirkpatrick, Level I

REACTION EVALUATION
Personal evaluation of the quality of the course and the amount of material learned.
Kirkpatrick, Level II

LEARNING (Acquisition)
An evaluation of the amount of material learned or the skills acquired in the classroom.

Kirkpatrick, Level III

BEHAVIOR (Transfer)
An evaluation of the extent to which the person uses the new KSACs on the job.

Kirkpatrick, Level IV

ORGANIZATIONAL RESULTS
The extent to which the organization benefited from the training.
**Types of Measurement**

**Criterion Referenced Vs. Norm Referenced**

- **Criterion referenced.**
  There is some absolute standard of performance against which a person’s behavior is judged.

- **Norm referenced.**
  A person’s behavior is measured and judged in relationship to some standard (norm) group.

**Objective vs. Subjective Measures**

- **Objective Criteria**
  There is a clear, unambiguous standard for assessing a person’s behavior.

- **Subjective Criteria**
  Measurement of the criteria is based on the expert judgment of the rater.
Obtrusive vs. Unobtrusive Measures

- **Obtrusive**
  The person is aware of the fact that his/her behavior is being assessed.

- **Unobtrusive**
  The person is not aware that his/her behavior is being assessed.

Uses of the Evaluation

- **FORMATIVE**
  Evaluate a training program with the goal of improving it before it is implemented.

- **SUMMATIVE**
  Evaluation of the overall impact of the final training program.

Evaluation Techniques
Preliminary Decisions

1. Decide on the level at which the evaluation will take place.
2. Decide on the type of measure you will use.
3. Decide on who will do the assessment.
4. Decide on the measurement technique.

1. Simple Survey

- Used for
  - Reaction
  - Transfer

- Benefits
  - Quick
  - Simple

- Issues.
  - Self-report.
  - Sample.

2. Experimental Designs
Basic Requirements

- Two or more groups.
- Participants randomly assigned to treatment conditions.
- One or more treatment conditions.

Experimental Designs

- Experimental Group
  The group that gets the treatment (training).
- Control Group
  Reference group that does not get treatment (training).
- Random Assignment of Subjects

True Experimental Designs (1)

- Randomized, Posttest Only, Control Group
  - Experimental group gets training.
  - Control group does not get training.
  - Two groups are evaluated on the criteria after the training.
Basic Design

<table>
<thead>
<tr>
<th>Treatment Groups</th>
<th>Independent Variable</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Training</td>
<td>Measurement(s) made after the treatments are applied.</td>
</tr>
<tr>
<td>Group 2</td>
<td>Something Else</td>
<td></td>
</tr>
</tbody>
</table>

True Experimental Designs (2)

- Randomized Pretest-Posttest, Control Group.
  - Both groups get a pre-test.
  - Experimental group gets training.
  - Control group does not get training.
  - Both groups are evaluated on the criteria after the training.

Randomized Pre-test Post-test Control Group

<table>
<thead>
<tr>
<th>SS, SS</th>
<th>Pre-test</th>
<th>Independent Variable</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>SS, SS</td>
<td>Measure A</td>
<td>Training</td>
<td>Measure A</td>
</tr>
<tr>
<td>SS, SS</td>
<td>Measure A</td>
<td>Something Else</td>
<td>Measure A</td>
</tr>
</tbody>
</table>
Internal Validity

Can I unambiguously conclude that the independent variable caused a change in the dependent variable.

Confounding

Any variable other than the independent variable that could reasonably have caused changes in the dependent variable.
Confounding Variable: Example

Confounded by time of day.

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Given orange juice in the morning</td>
<td>Alertness two hours later.</td>
</tr>
<tr>
<td>Given distilled water in the evening</td>
<td></td>
</tr>
</tbody>
</table>

Pre-Experimental And Quasi-Experimental Designs

One-Shot Case Study

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Training Program</td>
</tr>
</tbody>
</table>
Benefits and Issues

Benefits.
1. OK for preliminary research.

Issues.
1. No comparison group.

One-group Pre-test Post-test

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test</th>
<th>Independent Variable</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure A</td>
<td>Training</td>
<td>Measure A</td>
<td></td>
</tr>
</tbody>
</table>

Benefits and Issues

Benefits.
1. OK for preliminary research.

Issues.
1. History.
2. Maturation.
3. Regression.
4. Testing.
5. Instrument decay.
1. History

*Any event that occurs between the first and second dependent measures that is not manipulated by the experimenter.*

<table>
<thead>
<tr>
<th>Pre-test</th>
<th>Training</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>Control</td>
<td>Post-test</td>
</tr>
</tbody>
</table>

2. Testing

*Participation in the pre-test may cause changes in the person.*

- Reactivity
- Memory

<table>
<thead>
<tr>
<th>Pre-test</th>
<th>Training</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>Control</td>
<td>Post-test</td>
</tr>
</tbody>
</table>

3. Maturation

*Changes in the individual over time that are not associated with the independent variable.*

<table>
<thead>
<tr>
<th>Training</th>
<th>Delay</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Delay</td>
<td>Post-test</td>
</tr>
</tbody>
</table>
4. Instrument Decay

Changes in the measuring instrument over time.
- Person gets bored.
- Test becomes obsolete.
- Machine wears out.

<table>
<thead>
<tr>
<th>Pre-test</th>
<th>Training</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Statistical Regression

Occurs when participants are placed into groups based on extreme scores. Extreme scores tend to move (regress) toward the mean.

<table>
<thead>
<tr>
<th>Pre-test</th>
<th>Training</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Field Research Example

- In 1984 Pittsburgh National Bank had a problem with their tuition reimbursement program.
- They were paying tuition and fees for employees seeking bachelors degrees.
- Approximately 45% of the people did not want to work in the field in which they majored.
- The bank was prepared to scrap the program.
### Evaluation Design

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental Group</td>
<td></td>
</tr>
<tr>
<td>342 people who attended workshop</td>
<td>Job posting Applications 70%</td>
</tr>
<tr>
<td>Control Group</td>
<td></td>
</tr>
<tr>
<td>450 people who did not attend the workshop</td>
<td>Job posting Applications 23%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Job posting Applications</th>
<th>Promotions</th>
<th>Salary/grade change</th>
</tr>
</thead>
<tbody>
<tr>
<td>70%</td>
<td>12%</td>
<td>91%</td>
</tr>
<tr>
<td>23%</td>
<td>3%</td>
<td>66%</td>
</tr>
</tbody>
</table>

### Non-equivalent Control Group

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>Training</td>
</tr>
<tr>
<td>Group X</td>
<td>Something Else</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
</tr>
</tbody>
</table>

### Benefits and Issues

**Benefits.**

1. May be the only alternative in field experimentation.

**Issues.**

1. Treatment difference is CONFOUNDED by group difference.
Interactive Problems

- Diffusion of information. One group talks to another.
- Intergroup competition. One group competes with another.
- Demoralization. One group resents the other group.

Time Series Designs

- Similar to Pretest Post Only design but with multiple preliminary measures.
- No control group.
- Usually have
  - Baseline measurement on dependent variable.
  - Some manipulation or event.
  - Second measurement on the dependent variable.

Interrupted Time Series Design

Baseline (Pre-test)  Event  Post Test
When Used

- Assess the impact of some event.
- Assess the impact of some broad treatment within an existing group.

Utility

Process of showing that the training produces some organizational benefit.
### Berkshire Hotels

<table>
<thead>
<tr>
<th>Should</th>
<th>86% occupancy</th>
<th>Strong Customer Relations Skills</th>
<th>Authority to make decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gap</td>
<td>5%</td>
<td>Training</td>
<td>Organization change</td>
</tr>
<tr>
<td>Is</td>
<td>81% occupancy</td>
<td>Mediocre Customer Relations Skills</td>
<td>No authority</td>
</tr>
</tbody>
</table>

### Did It Work?

1. Randomly selected 40 of the 80 hotels for training.
2. Compared the occupancy rates for the 'Trained' hotels against the Control group.
3. Found an increase in occupancy rate of 2 rooms per day.

### So What!???

Marketing research, design, consulting, materials, and travel expenses: $275,000.

2 x $135 = $270 additional income each day.

$270 x 200 workdays = $54,000 additional annual income per hotel.

$54,000 x 40 hotels = $2,160,000 for 40 hotels.
The End