

Criterion Development and Performance Appraisal

Overview

- Why do we do performance appraisals?
- What should we measure?
- How should we measure it?
- How does the evaluation process work?
- How do employees react to performance appraisal?
- What have we learned, and how can we improve the appraisal process?

Purposes for Performance Appraisal

- Administrative Decisions
 - e.g., promotion, termination, compensation, training
- Employee Development
 - e.g., performance improvement, career development
- Validation of Selection Instruments
- Criterion Measures in Research

The Criterion Problem

Criteria = Standards for Success

The Problem

What standards to apply?

How to measure against the standards?

How to use these measures?

Some Important Distinctions

Attempts to address the criterion problem have led to some important distinctions

- Ultimate vs. Actual Criteria
- Multiple vs. Composite Criteria
- Task vs. Contextual Performance
- Static vs. Dynamic Criteria
- Objective vs. Subjective Criteria

Ultimate vs. Actual Criteria

The Problem: Deciding what to measure

The Solution (Thorndike, 1949)

- **Ultimate (Conceptual) Criterion**
 - Theoretical construct encompassing all performance standards constituting success on a job.
- **Actual (Operational) Criterion**
 - Measure intended to reflect the ultimate criteria as accurately as possible

Use ultimate criterion to guide measure development

Ultimate vs. Actual Criteria

Examples

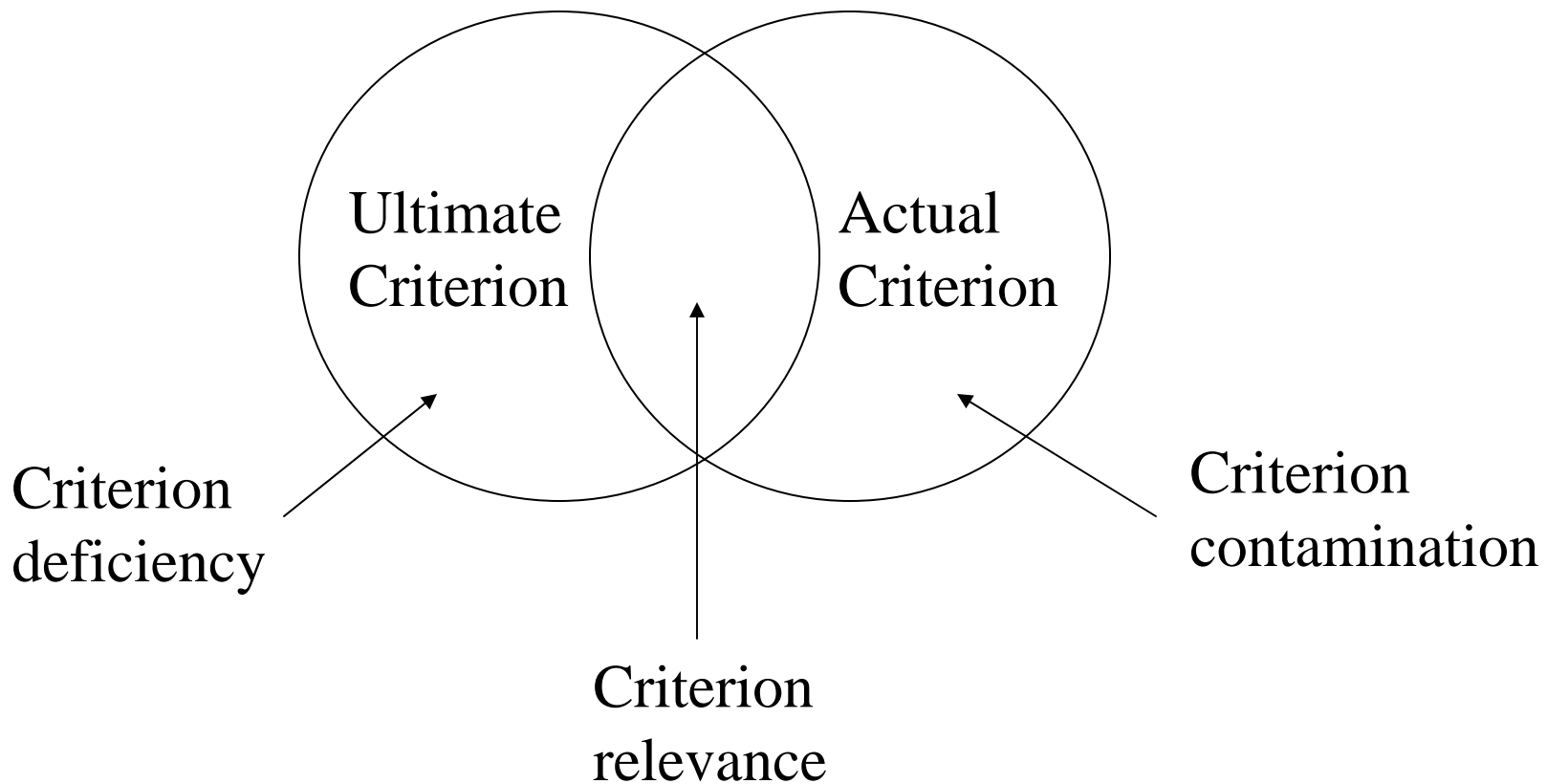
1. Student Performance

- What criteria *should* we use to evaluate student performance?
- What types of measures *do* we use?

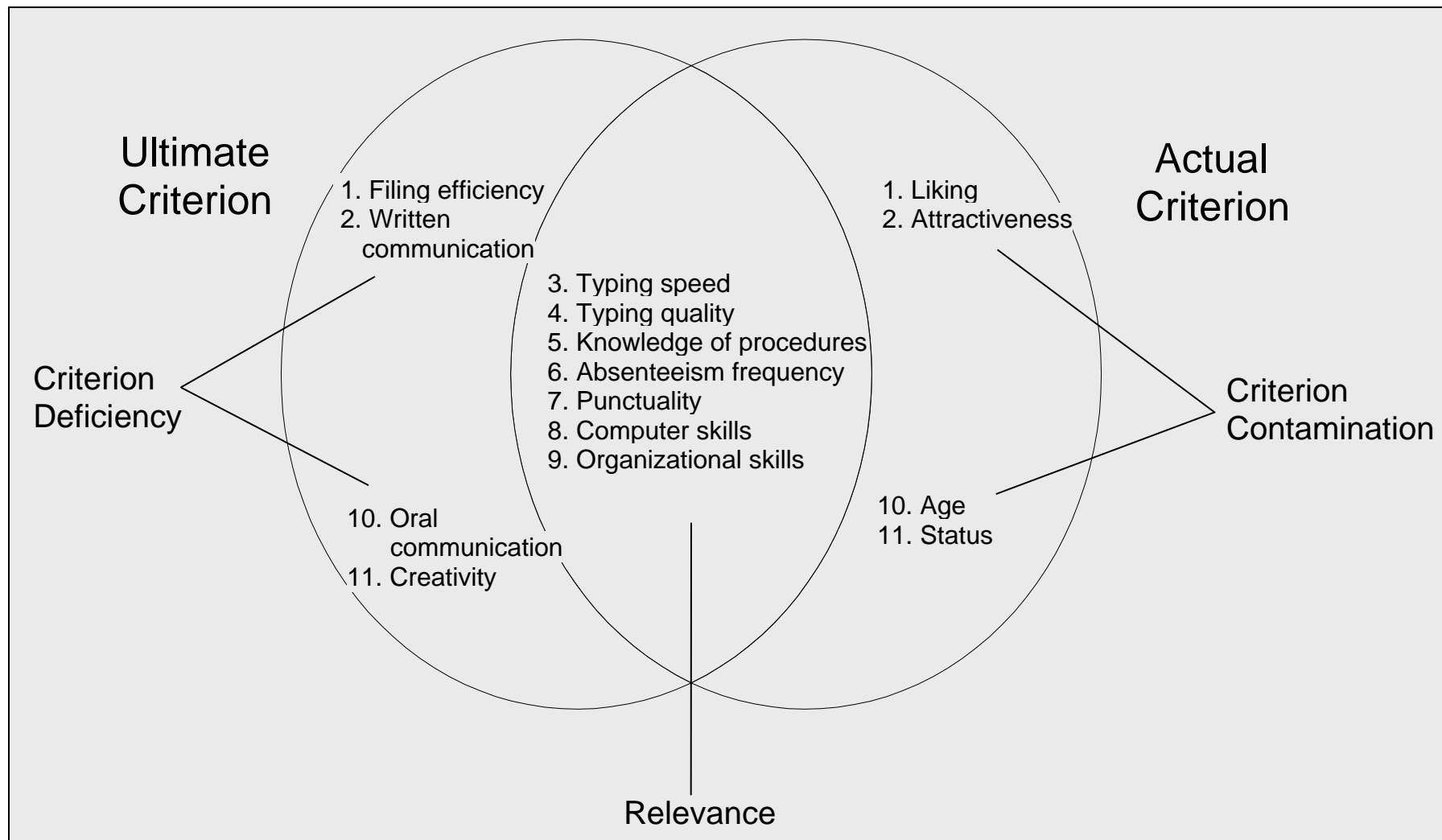
2. Employee Performance

- What criteria *should* we use to evaluate employee performance?
- What types of measures *do* we use?

Evaluating Actual Criteria



Work Example



Multiple vs. Composite Criteria

The Problem: Performance is multi-faceted *but* personnel decision (e.g., promotions) often require an overall evaluation

The Solution:

1. Acknowledge and measure multiple facets of performance (e.g., for theory, research, employee development)
2. When necessary, use weighted or un-weighted combination to form a composite criterion

Task vs. Contextual Performance

The Problem: Typical models of job performance focus only on *task performance* (aka, in-role performance)

The Solution: (e.g., Borman & Motowidlo, 1997)
Broaden the conceptualization of performance to include *contextual performance* (aka, extra-role performance, OCB)

Task vs. Contextual Performance

Task Performance: work-related activities performed by employees that contribute to the technical operations of the organization

Contextual Performance: activities performed by employees that help to maintain the broader organizational, social, & psychological environment

Static vs. Dynamic Criteria

The Problem: Performance can change over time, making it difficult to find a single set of predictors

Note. This is really a *prediction* problem

The Solution:

1. Identify the way(s) in which performance typically changes over time
2. Rule out explanations that can be addressed through more effective management practices
3. Identify indicators that will be effective in predicting performance in the short and long term

Objective vs. Subjective Criteria

The Problem: How to measure performance (i.e., selection of actual criteria)

The Options:

1. Objective criteria: use of quantitative measures that do not require subjective judgment (e.g., #widgets; \$ value of sales; days absent)
2. Subjective criteria: rankings or ratings made by individuals who are knowledgeable about performance (e.g., supervisors, peers, customers)

Potential Problems with Objective Measures

1. May not include all relevant aspects of performance (criterion deficiency)
2. Objective outcomes may be influenced by irrelevant factors (criterion contamination)
3. Objective outcomes are not available for some jobs

Subjective Measures of Performance

Potential Rating Sources

- Supervisor
- Peers
- Self
- Subordinates
- Clients/customers

Subjective Measures of Performance

Criteria for Selecting Evaluators

- Familiarity with performance (i.e., opportunity to observe)
- Knowledge of the job (purpose, constraints, etc.)
- Freedom from bias

Note. All sources have strengths and weaknesses.
Therefore it is often beneficial to use multiple raters (e.g., 360° feedback)

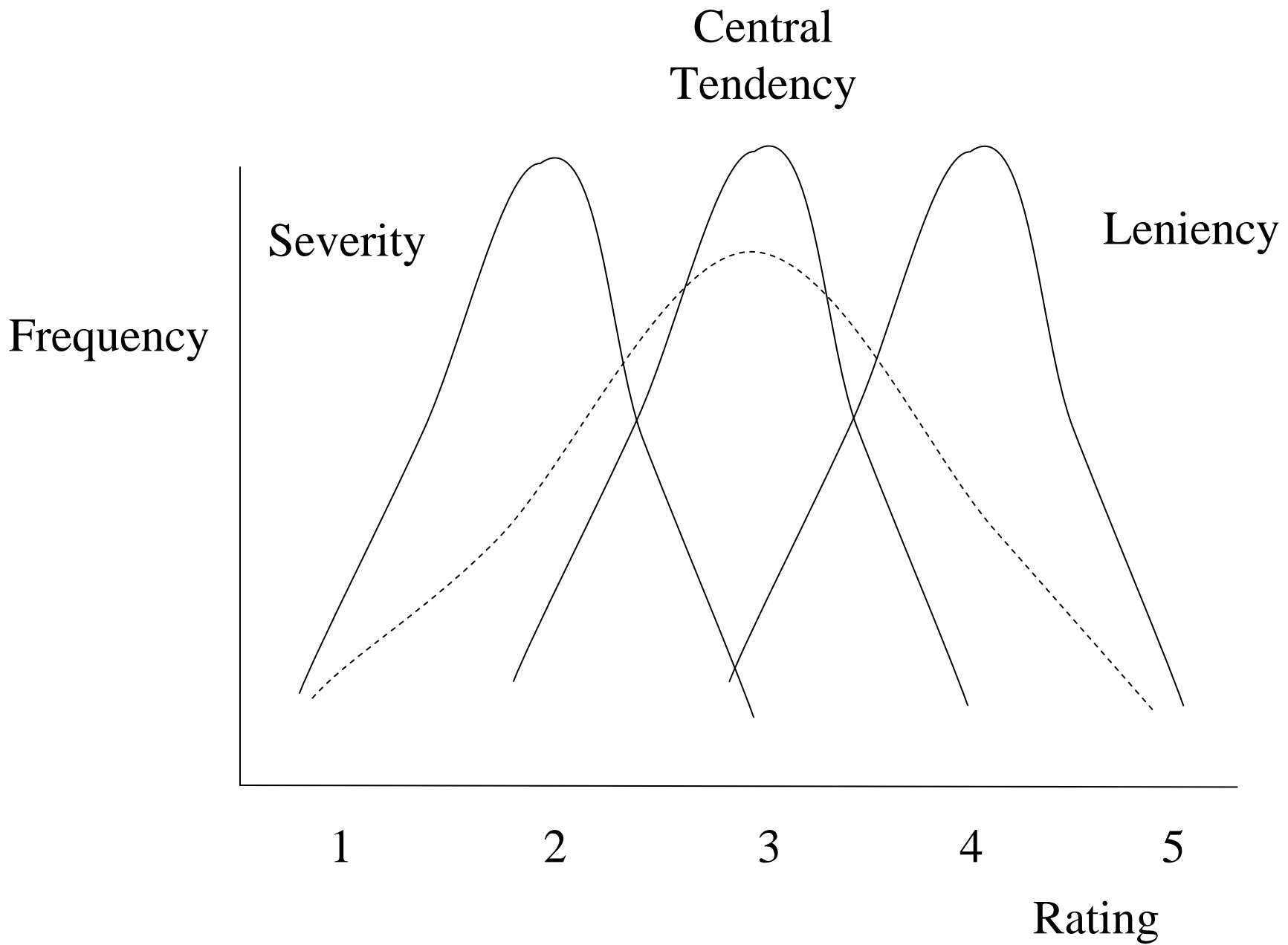
Film: Performance Appraisal

Things to look for:

- Reasons for performance appraisal
- Reactions to performance appraisal (by appraisers and appraisees)
- Problems with subjective evaluations
 - Unintentional errors
 - Intentional distortions
- Solutions to the problems

Common Errors in Subjective Evaluations

- Distributional errors
 - Leniency
 - Severity
 - Central Tendency



Common Errors in Subjective Evaluations

- **Distributional Errors**

- Leniency

- Severity

- Central Tendency

- Halo error (vs. true halo)
- Attribution errors (e.g., fundamental attributional error)
- Temporal errors (primacy; recency)
- Contrast errors
- Personal Biases (e.g., age, race, gender, similar-to-me)
- Intentional Distortion (motivation, political, economic)

Types of Appraisal Form

1. Narrative Evaluation

- **Strengths**

- Easy to develop and administer
- Allows rater flexibility

- **Weaknesses**

- Evaluations are not standard across ratees
- Potential for contamination and deficiency

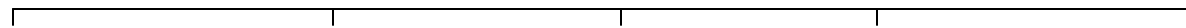
Types of Appraisal Form

2. Comparative Methods (e.g., ranking, pair comparisons, forced distribution)
 - **Strengths**
 - Reduces distributional errors
 - Useful in making administrative decisions
 - **Weaknesses**
 - Compares people to people and not to the standards
 - Less useful for employee development purposes

Types of Appraisal Form

3. Performance Ratings (e.g., graphic rating scales)

Judgment



1

2

3

4

5

Poor

Average

Excellent

Types of Appraisal Form

3. Performance Ratings (e.g., graphic rating scales)

- **Strengths**

- Quantification aids in administrative decisions
- Easy to construct and complete

- **Weaknesses**

- Subject to rating errors
- Does not permit elaboration or explanation

Improving the Rating Form

1. Different types of appraisal can be combined to take advantage of their relative strengths (e.g., graphic rating combined with narrative explanation)

2. Refinements have been made to minimize errors

Examples

- Behaviourally Anchored Rating Scales
- Behavioural Observation Scales

Behaviourally Anchored Rating Scales (BARS)

Purpose

Clarify constructs and anchors to improve rating accuracy and interpretation

Method of Development

1. Conduct a job analysis to identify relevant behaviour (typically Critical Incidents Technique)
2. Identify major job factors (categories)
3. Have SMEs classify behaviours into categories
4. Have SMEs rate behaviours for effectiveness
5. Select behaviours that SMEs agree on as scale anchors

Knowledge and Judgment

Extremely good	7		By knowing the price of items, this checker would be expected to look for mismarked and unmarked items
Good	6		
Moderately good	5		When operating the quick check, the lights are flashing, this checker can be expected to check out a customer with 15 items
Neither good nor poor	4		
Moderately poor	3		In order to take a break, this checker can be expected to block off the checkstand with people in line
Poor	2		
Extremely poor	1		

Behaviourally Anchored Rating Scales (BARS)

Strengths (relative to standard scales)

- Based on systematic job analysis
- Behaviours help to clarify meaning of construct and scale values
- Raters and ratees involved in the development (increases acceptance and understanding)

Weaknesses

- Costly and time consuming to develop
- Behaviours might not be those actually observed
- Only examples of behaviours are included

N.B. Evidence for superiority of BARS is limited

Behavioural Observation Scales (BOS)

Purpose

Improve ratings by focusing attention on concrete behaviour required for effective performance

Method of Development

1. Conduct a job analysis to identify relevant behaviour (typically Critical Incidents Technique)
2. Identify major job factors (categories)
3. Have SMEs classify behaviours into categories
4. Have SMEs rate behaviours for effectiveness & importance
5. Select appropriate behaviours for inclusion on the appraisal form along with a frequency rating scale

Overcoming Resistance to Change

1) **Describes the details of the change for subordinates.**

Almost never 1 2 3 4 5 Almost always

2) **Explains why the change is necessary.**

Almost never 1 2 3 4 5 Almost always

3) **Discusses how the change will affect the employees**

Almost never 1 2 3 4 5 Almost always

4) **Listens to employee's concerns**

Almost never 1 2 3 4 5 Almost always

5) **Asks the employees for help in making the change work.**

Almost never 1 2 3 4 5 Almost always

6) **If necessary, specifies the date for a follow-up meeting to respond to the employee's concerns.**

Almost never 1 2 3 4 5 Almost always

Total = _____

Below adequate

Adequate

Full

Excellent

Superior

6-10

11-15

16-20

21-25

26-30

Behavioural Observation Scales (BOS)

Strengths (relative to standard scales)

- Based on systematic job analysis
- Designed to reduce rating errors by focusing on concrete behaviour
- Raters and ratees involved in the development (increases acceptance and understanding)
- Can be used for job description and goal-setting

Weaknesses

- Costly and time consuming to develop
- Relies on rater's perception and recall
- Lengthy and time-consuming to complete

The Appraisal Process

Evaluation process

Observe Behaviour



Encode Observations
in Memory



Interpret request for
evaluation



Recall Behaviour



Rate Performance

Potential Problems

Selective exposure
Selective attention
Distorted perception

Information overload
Encoding problems

Ambiguous criteria
Hidden agendas

Failed or Selective recall
Mood influences

Motivated distortion
Judgment errors

Potential Solutions

- Job Analysis (clear criteria)
- Use of logbooks (to aid memory)
- Multiple observers/raters (e.g., 360° feedback)
- Unambiguous rating forms
- Incentives/support/evaluation of raters
- Training (e.g., Frame of Reference Training)

Frame of Reference Training

Step 1

- Identify and describe performance dimensions
- Explain performance standards

Step 2

- Present examples of performance on video
- Have trainees evaluate performance
- Give feedback (illustrate errors and show expert ratings)

Step 3

- Follow-up consultations

Employee Reactions to Performance Appraisal

Factors found to affect reactions:

- Favourability of the rating
- Job relevance of the rating
- Accuracy of the information used
- Discussion of plans for future performance
- Opportunity to respond or provide self-evaluation

N.B. These factors all relate to the fairness of the evaluation and the evaluation process

Three Forms of Fairness (Organizational Justice)

Distributive Justice

- Fairness of the outcomes received

Procedural Justice

- Fairness of the procedures used to distribute the outcomes

Interactional Justice

- Fairness in the way the procedures are enacted, and in the way people are treated

Criteria for Procedural Justice

Source: Leventhal (1980)

- Consistency across time and people
- Adherence to shared ethical standards
- Recognition of everyone's concerns
- Freedom from bias
- Use of accurate information
- Opportunity to correct unjust decisions

Film: Computer Monitoring

Things to Consider:

- What is the “case for” computer monitoring?
- What is the “case against” computer monitoring?
- Is computer monitoring “fair” (according to Leventhal’s criteria)?
- Can anything be done to improve computer monitoring to make it more fair and effective?

Computer Monitoring

The Case For

- Increased productivity (e.g., increased sales at reduced cost)
- Increased competitiveness
- Improved service and customer satisfaction
- More efficient management (e.g., more accurate evaluation of performance)

Computer Monitoring

The Case Against

- Hidden costs (e.g., insurance, turnover)
- Increased stress and health problems
- Decreased satisfaction and commitment
- Decreased customer service
- Invasion of privacy

Computer Monitoring

Ways to Improve

- Ensure that the criteria are appropriate
- Inform employees of monitoring and its purpose
- Use for developmental purposes (use feedback to improve performance)
- Involve employees in design

Improving the Performance Appraisal Process

Preparation

- Raters should know in advance that ratings are to be made and should know what the criteria are
- Raters should have the opportunity to observe the behaviours to be evaluated

Conditions

- Raters should avoid doing evaluations when in an unusual (positive or negative) frame of mind
- Sufficient time should be allotted for doing appraisals and they should be done under quiet conditions

Improving the Performance Appraisal Process

Instruments

- Ratings should focus on specific and observable outcomes and behaviours
- Rating scales should be carefully constructed with clearly defined rating categories

Training and Incentives

- Raters should be trained on how to observe performance objectively and systematically, how to avoid errors, and how to improve feedback
- The importance of accurate ratings should be emphasized and raters should be evaluated on their ability to evaluate

Improving the Performance Appraisal Process

Feedback

- Managers should deal with problems as they arise rather than saving them up. Overly negative reviews can evoke defensiveness.
- Managers should include planning and goal setting (both organizational and personal) as part of the feedback process.

Purpose

- When possible, do separate evaluations for administrative decisions and employee development