

*human resource*  
**selection**  
ninth edition

GATEWOOD ▪ FEILD ▪ BARRICK

**INSTRUCTOR'S MANUAL**



© 2019 Wessex Press, Inc. All rights reserved.  
[www.wessexlearning.com](http://www.wessexlearning.com)

# TABLE OF CONTENTS

## Plan of This Book 3

### PART 1 *Foundation for a Selection Program*

---

- Chapter 1** An Introduction to Selection 4
- Chapter 2** Job Performance Concepts and Measures 8
- Chapter 3** Job Analysis in Human Resource Selection 13
- Chapter 4** Legal Issues of Selection 21
- Chapter 5** Recruitment of Applicants 27

### PART 2 *Measurement in Selection*

---

- Chapter 6** Human Resource Measurement in Selection 33
- Chapter 7** Reliability of Selection Measures 38
- Chapter 8** Validity of Selection Procedures 42

### PART 3 *Selection Measures*

---

- Chapter 9** Application Forms–Biodata Assessment, Training & Experience Evaluations, and Reference & Social Media Checks 53
- Chapter 10** The Selection Interview 63
- Chapter 11** Ability Tests for Selection 72
- Chapter 12** Personality Assessment for Selection 80
- Chapter 13** Simulation Tests 86
- Chapter 14** Testing for Counterproductive Work Behaviors 92

### PART 4 *Using Selection Data*

---

- Chapter 15** Strategies for Selection Decision Making 99

## PLAN OF THIS BOOK

This book is divided into four parts:

- Part 1: Foundation for a Selection Program
- Part 2: Measurement in Selection
- Part 3: Selection Measures
- Part 4: Using Selection Data

### Part 1: Foundation for a Selection Program

- In today's competitive business environment, managers want to increase the performance of their employees.
- Selection is the basis for employee performance.
- Selection programs are useful if:
  1. They develop instruments that collect job-related information from applications.
  2. This information is used appropriately.
- Part 1 of this text will cover:
  1. The steps taken to develop a selection program (Chapter 1)
  2. The various forms of job performance (Chapter 2)
  3. The steps necessary to identify worker characteristics that lead to job success (Chapter 3)
  4. The specific legal demands of selection — laws, executive orders, court decisions, etc. (Chapter 4)
  5. The composition of recruitment programs that will attract appropriate applicants (Chapter 5)

### Part 2: Measurement in Selection

- Information is the basis for all decisions concerning the selection of job applicants
- Sometimes, however, HR selection decisions turn out to be wrong – faulty data
- Objectives of Part 2:
  1. Explore the role of HR *measurement* in selection decision making
  2. Examine the concepts of *reliability* and *validity* of selection data and their role in choosing useful selection measures and making effective selection decisions

### Part 3: Selection Measures

- Chapters in this part of the book are devoted to the main selection instruments that are used in choosing among applicants
- Each chapter discusses the necessary features of an instrument, common deficiencies to avoid, how to interpret scores, and legal issues in its use

### Part 4: Using Selection Data

- Once valid selection instruments have been developed, administered to applicants, and scored, the scores must be combined in some way and decision rules must be developed to evaluate individual applicants in order to offer employment.

## CHAPTER 1

# *An Introduction to Selection*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand what a selection program requires in terms of collecting and evaluating data on applicants and its contribution to improving the performance of the organization.
2. Know how selection interacts with the other human resource programs of recruitment, training, performance measurement, and compensation.
3. Understand the five main steps in developing a selection program and what type of data is necessary for each step.
4. Know the inherent constraints in developing and using a selection program and the limitations that these have on a selection program.

### DEFINITION OF SELECTION

[slides 4–10]

- The process of collecting and evaluating information about an individual in order to extend an offer of employment.
  - Such employment could be either a first position for a new employee or a different position for a current employee.
  - The selection process is performed under legal and market constraints and addresses the future interests of the organization and the individual.

### Collecting and Evaluating Information

- The selection specialist must systematically collect information from applicants about how much of the necessary characteristics each possesses:
  - WRCs (work-related characteristics)—human attributes related to job performance, including personality
  - KSAs (knowledge, skills, and abilities)—a traditional term used in HR, excluding factors such as personality
- Distinguish between selection and hiring.
- Selection occurs when job-related information is collected from applicants and offers of employment are given to those who possess the WRCs to do well on the job.
- Often, hiring occurs when a job is offered with no evaluation of the applicant's job-related qualifications.

## Selection for Initial Job and Promotion

### *Characteristics of Selection for Initial Job*

1. Applicants are external to the organization—commonly students, people who have recently completed an education, those currently not employed or employed at other organizations
2. Applicants are recruited through formal mechanisms—media advertisement, Internet contact, employment agencies, referrals by other employees
3. These mechanisms frequently produce a large number of applicants
4. With a large number of applicants, costs are a factor—a brief selection instrument (application form that collects limited information) reduces this number
5. Remaining applicants undergo a series of steps—interviews, ability tests, job simulations
6. Decisions formalized by statistical analysis or decision rules

### *Characteristics of Selection for Promotion*

1. Candidates are internal to the organization—existing members of an organization
2. A limited number of recruitment techniques used, or no formal recruitment techniques
3. Because applicants are members of the organization, there is already a great deal of information about them—performance reviews, training records, work history, records of attendance, reprimands, awards, etc.
4. Often the evaluation is not formalized—subjective decision making

## External Factors and Future Interests

- Ideally, an organization makes selection decision with a great deal of control over the number of applicants, information gathered from them, and the decision rules used in evaluating the information.
- However several factors limit the control of organizations over selection
  - great fluctuations in labor markets for applicants
  - economic and educational conditions in areas of recruiting
  - federal and state discrimination laws
  - future interests of both applicants and the organization must be considered

## Is There Evidence that Selection is Important?

- The resource-based theory of organizations holds that they can gain advantage over competitors by having and holding a valuable resource in short supply
  - Studies concluded:
    - Selection and internal training directly influence a company's profit because they influence labor productivity
    - Selection and training were related to customer service performance, employee retention

## SELECTION AND OTHER HUMAN RESOURCE SYSTEMS

[slides 11–12]

- Many systems are important for employee performance – recruitment, training, compensation, job performance measurement (Figure 1.1)
- To get the maximum benefit from these systems, firms must understand that HR activities are interconnected and greatly enhance employee work performance
- Selection is more closely related to recruitment than to other Human Resource Management (HRM) programs

## DEVELOPING A SELECTION PROGRAM

[slides 13–14]

- Steps in the development of a selection program determines the adequacy of the selection process (Figure 1.2)
  - Job analysis
  - Identifying relevant job performance measures
  - Identification of WRCs
  - Development of selection measures
  - Validation of selection instruments

## CONSTRAINTS IN DEVELOPING A SELECTION PROGRAM

[slides 15–17]

- The essence of selection is *prediction* (forecasting). A number of factors greatly affect the accuracy of the selection process:
  - Limited information on applicants
  - Applicant and organization at cross-purposes
  - Measurement of jobs, individuals, and work performance
  - Other factors affecting work performance – training programs, appraisal and feedback methods, goal-setting procedures, etc.
  - Selection research versus selection practice

## What is selection doing that is good for the world?

- Three of the many important areas that will be pertinent in the future:
  - Big Data—collecting, analyzing, and interpreting data about business-based problems
  - The magnitude of demographic group differences and their effects—in measuring WRCs that applicants possess, there are consistent differences among demographic, ethnic, and racial groups
  - The use of various Internet-based selection measures
    1. Proctored assessment
    2. Unproctored assessment

*The Use of Various Internet-Based Selection Measures:*

- Serious problems common to both proctored and unproctored testing:
  - Computer literacy
  - Graphic transmission
  - Technical failures
  - Equivalence of electronic testing results to those from traditional testing
- Problems characteristic of only unproctored online testing:
  - Cheating
  - Security of test items
  - Standardization of test environment

## CHAPTER 2

# *Job Performance Concepts and Measures*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand four key dimensions of performance.
2. Know different ways to measure performance.
3. Know when to use single or multiple criteria for validation.
4. Recognize the importance of retaining productive employees through selection.
5. Understand types of withdrawal behaviors and the ways to assess them.
6. Know different types of judgmental ratings used to assess performance.

### JOB PERFORMANCE CONCEPTS AND MEASURES

[slides 4–7]

- Applicants who score high on selection tests are predicted to do well in their future job performance
- What is meant by “job performance”?
  - Traditionally, it has meant task performance, since most jobs involved physical activities
  - An employee’s production was easy to measure
  - Selection tools measured knowledge of topics and tasks
- More service and knowledge-sector jobs
- Concept of job performance and nature of selection tests changed
  - Teams of workers
  - Collaboration required in complex fields
  - Workers think, plan, make observations, draw conclusions, interpret data—not easily measurable

#### *How Job Performance Is Viewed*

- Task performance still the primary component of job performance
- Work characteristics measured include numerous scales of social processes, culture, and work context
- Three facets of job performance in addition to task performance
  - Organizational citizenship behaviors (OCBs)
  - Adaptive performance (AP)
  - Counterproductive work behaviors (CWBs)

*How Job Performance Is Measured*

- Count the number of produced items or services rendered
- Supervisors make judgments on a worker's job behaviors

*The Type of Selection Measures That Are Used to Predict Job Performance*

- Measuring WRCs with job-knowledge tests and having applicants perform parts of the job continue to be used
- O\*NET database identifies four general categories of characteristics—abilities, occupational interests, work values, work styles
- Traditional applications expanded to include job- and task-related information

**TASK PERFORMANCE**

[slides 8–13]

**Production Data**

- The results of work comprising things that can be counted, seen, and compared directly from one worker to another
- Measures based on the specific nature of the job tasks
- The variety of measures so great makes it impossible to summarize—see Table 2.1 for examples showing both quantity and quality of production

**Judgmental Data**

- An individual familiar with the work of another required to judge this work—usually obtained using a rating scale with numerical values
- In most cases evaluations done by the immediate supervisor but can be done by others—subordinates, peers, customers

*Types of Judgmental Instruments*

- Trait rating scales (a bad method, do not use!)
  - supervisor evaluates a subordinate on personal characteristics
- Simple behavioral scale (better method, could be used)
  - supervisor rates a subordinate on major or critical tasks of the job
- BARS or BES (even better method!)
  - BARS (Behaviorally Anchored Rating Scales)
  - BES (Behavioral Expectation Scales)
- 360-degree feedback (a useful technique for evaluating managers)
  - gathers judgmental information from superiors, peers, subordinate

*Issues with Judgmental Scales*

- Intentional and inadvertent bias by the individual making the judgment
  - Intentional bias—when the rater deliberately distorts the ratings to either be favorable or unfavorable
  - Inadvertent bias (commonly called *rater error*)—halo, leniency or severity, central tendency
- May be based on production data

## ORGANIZATIONAL CITIZENSHIP BEHAVIORS

[slides 14–17]

- Not formally part of the job task behaviors but done by the individual to assist other workers or the organization itself
  - Mentoring new workers
  - Assisting other workers
  - Putting extra time and effort into work
- Dimensions of OCBs:
  1. Helping behavior
  2. Sportsmanship
  3. Organizational loyalty
  4. Organizational compliance
  5. Individual initiative
  6. Civic virtue
  7. Self-development

### What Prompts OCBs?

- Linked to organizational commitment, perceptions of fairness, and leader supportiveness

### Relationship of OCBs with Other Performance Measures

- Managers influenced by worker's OCBs, especially judgmental performance evaluations

### Measurement of OCBs

- Self-report judgmental scales used; bias?

## ADAPTIVE PERFORMANCE

[slides 18–20]

- A deliberate change in the thinking or behavior of an individual because of an anticipated or existing change in the work activities or environment
  - Differences in WRCs can be used to predict differences in AP
  - Both OCBs and AP are facets of job performance and can be included in operational selection programs
- Research work using AP in selection has concentrated on determining WRCs that would predict AP
  - Cognitive complexity—consider and integrate conflicting information
  - Frame changing—alternate between multiple ways of attending to and interpreting problems and solution strategies
  - Resiliency—persist and recover quickly
  - Problem solving—persist and work through details of a problem
  - Learning agility—apply lessons learned from previous experience

## COUNTERPRODUCTIVE WORK BEHAVIOR

[slides 21–23]

- Undesirable performance actions that harm the organization itself and often its employees and customers
  - Any intentional behavior by an employee that is viewed by the organization as contrary to its legitimate interests
  - Integrity tests developed to identify applicants with higher than normal probability of committing CWBs
  - CWBs cost billions of dollars
- CWBs can be divided into two groups:
  - Actions of deviance toward individuals (ID)
  - Actions toward the organization (OD)
- OCBs and CWBs are moderately negatively correlated:
  - Two distinct constructs rather than a single continuum
- Studies regarding the relationship of CWBs to individual and organizational characteristics

## EMPLOYEE WITHDRAWAL BEHAVIORS

[slides 24–25]

- Initiated at the discretion of the work and depict the individual's detachment from work:
  - Cognitive distancing
  - Physically removing oneself by being disengaged at work
  - Showing up late or not at all—absenteeism
  - Voluntarily severing employment ties
- Dynamic and influenced by many antecedents, with validated selection procedures:
  - Biographical information
  - Personality
  - Person-organization fit

## APPROPRIATE CHARACTERISTICS OF JOB PERFORMANCE MEASURES

[slide 26]

- Individualization—must collect data about performance that the individual controls
- Relevance—must measure production of critical or important parts of job
- Measurability—must be possible to generate a number that represents the amount or quality of the work performed
- Variance—scores that are generated must have differences among them

## USE OF CRITERIA FOR VALIDATION

[slides 27–28]

### Single versus Multiple Criteria

- Use of a single composite measure sees global performance; interpretation relatively simple
- Job analysis studies identify multiple tasks within jobs, indicative of the multiple aspects of job performance; global measure of performance many not reflect all activities

### When to Use Each

- For selection, use composite criterion
- For research, use multiple scores

### Forming the Single Measure

- Dollar criterion—what is the value of worker performance to the organizations?
- Factor analysis—a majority of the separate measures combined into one factor; provides weights that could be applied to each specific measure
- Expert judgment—must identify the weight of specific performance aspects

## CHAPTER 3

# *Job Analysis in Human Resource Selection*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand what job analysis is and its importance in human resource (HR) selection.
2. Describe several common job analysis methods.
3. Explain how job analysis results are implemented in choosing and developing HR selection procedures.

### JOB PERFORMANCE: A DEFINITION AND ROLE IN HR SELECTION

[slides 4–14]

- A purposeful, systematic process for collecting information on the important work-related aspects of a job, including:
  1. Work activities – what a worker does; how, why, when these activities are conducted
  2. Tools and equipment used in performing work activities
  3. Context of the work environment – work schedule, physical working conditions
  4. Requirements of personnel performing the job—knowledge, skills, abilities, personality characteristics, or other specifications (WRCs)
- Job analysis data help to:
  1. Identify employee specifications or WRCs necessary for success on a job
  2. Select or develop selection procedures that assess important applicant WRCs to forecast those candidates likely to succeed on the job
  3. Develop criteria or standards of job performance that represent employee job success

### Growth in Job Analysis

- Within the past four decades or so, employers have given considerable attention to job analysis:
  - Jobs are not static entities—the nature of jobs change (technological advancements, seasonal variations, initiatives of incumbents)
  - Federal guidelines (*Uniform Guidelines on Employee Selection Procedures*) have had a significant effect
  - Professional standards (*Principles for the Validation and Use of Personnel Selection Procedures*) have emphasized the important role of job analysis in HR selection programs

## Legal Issues in Job Analysis

- Title VII makes it illegal to refuse an individual or to discriminate against a person with respect to compensation, terms, conditions, privileges of employment due to race, sex, color, religion, or national origin.
- Many Title VII cases have concerned the role of discrimination in selection, thus job analysis and its associated methodologies have become intertwined with the law.
- Two developments amplified the importance of job analysis in selection research:
  - a. Adoption of the *Uniform Guidelines on Employment Selection Procedures*
  - b. Litigation involving discrimination in selection
- Key court cases involving job analysis:
  - a. *Griggs v. Duke Power*—selection standards used without meaningful study of their relationship to job performance ability
  - b. *Albermarle Paper Co. v. Moody*—court criticized the lack of a job analysis in a validation study
- A review of 26 selected federal court cases (by Duane Thompson and Toni Thompson) found:
  1. Job analysis is mandatory and must be for the job for which selection procedures are used
  2. Analysis of the job should be in writing
  3. Job analysts should describe in detail the job analysis procedures used
  4. Knowledgeable job analysts should collect job data from a variety of current sources
  5. Sample size of individuals serving as subject matter experts should be large and representative of the jobs for which the selection procedures are used
  6. Tasks, duties, and activities should be included in the analysis
  7. The most important tasks should be represented in the selection procedures
  8. Competence levels of work performance for entry-level jobs should be specified
  9. WRCs including knowledge, skills, and abilities should be specified, particularly if a content validation strategy is used
- Federal Guidelines on Employee Selection
  - *Uniform Guidelines* agreed to by EEOC, Department of Justice, Department of Labor, Civil Service Commission
  - Many recent cases alleging discriminatory impact because of inferences made during human judgment of job analysis data

## Collecting Job Information

- The role of job analysis in HR selection is to identify the requisite WRCs and translate them into tests, interviews, etc.
- The process requires judgments or inferences at several points (Figure 3.2)
  1. Data collected from a job analysis to infer relevant employee specifications
  2. Content of selection procedures that reflect these identified specifications

## A SURVEY OF JOB ANALYSIS METHODS

[slide 15]

- Methods that seem most popular in current HR selection practice:
  - Job analysis interviews
  - Job analysis questionnaires (including task analysis inventories)
  - Critical Incidents Technique
  - SME or job expert workshops
- Many other methods available

## JOB ANALYSIS INTERVIEWS

[slides 16–24]

### Description

- The interview is one of the most frequently used methods of job analysis, meeting many purposes
- Consists of a trained analyst asking questions about the duties and responsibilities, WRCs required, conditions of employment
- Typically involves group or individual interviews with incumbents and supervisors – SMEs (subject matter experts) because of their familiarity with job content
- A job analysis interview typically is performed for one or more of the following reasons:
  1. To collect job information that will serve as a basis for developing other job analysis measures, such as a job analysis questionnaire
  2. To clarify or verify information collected previously through other job analysis methods
  3. To provide a method, preferably as one of several used, for collecting relevant job data for developing a selection system

### Considerations on Applicability

- The interview is applicable to a variety of jobs, from those composed of physical activities to those composed on functions and activities that are primarily mental
- An effective interview requires a detailed plan:
  - State the objectives of the interview—identification and rating of job tasks
  - Individuals to be interviewed—incumbents representative of work shift, gender, racial groups; incumbents with six months or more of job experience
  - Questions and means for recording answers—schedule listing questions and forms for responses
  - Who will conduct the interviews—consultants

### An Example

- Identification of critical job tasks:
  1. What the worker does, by using a specific action verb that introduces the task statement
  2. To whom or what he or she does it, by stating the object of the verb
  3. What is produced, by expressing the expected output of the action
  4. What materials, tools, procedures, or equipment are used
- Use task statement criteria to produce task statements representing important task activities (Table 3.1)

## Limitations of the Job Analysis Interview

- Lack of standardization
- Limited possibilities for covering large numbers of respondents
- Legal requirements likely will go unmet
- Requires a great deal of time and labor—not cost-efficient
- Interviewer may have to track through an entire job in specific detail—expensive and requires skill
- Information may be distorted

## JOB ANALYSIS QUESTIONNAIRES

[slides 25–33]

### Description

- Questionnaire distributed to respondents through various means—in person, by a job analyst, by mail, via e-mail including link to online questionnaire
- Questionnaire lists job information—activities or tasks, tools and equipment used to perform the job, working conditions, WRCs incumbents must possess
- Participants make some form of judgment about job information presented on the questionnaire
- Questionnaires may be *tailored* (developed for a specific purpose or job) or *prefabricated* (existing questionnaires)

### The Task Analysis Inventory

- Questionnaire or survey that includes a listing of job tasks on which respondents make some form of judgment (ratings)
- Ratings given use a task rating scale such as frequency of task performance
- Usually concerns only on job or a class of similar jobs
- Job incumbents complete a task analysis inventory—supervisors can complete it if they have current knowledge about the job
- Historically, this method widely used in military settings—in particular by the U.S. Air Force
- *The nature of task inventories*—often contains three major types of information
  - Background information on respondents—name, gender, ethnicity, tenure on the job rate, tenure with the employing organization, job location, work shift, title of the job rated
  - Listing of job tasks with associated rating scales
  - Other or miscellaneous information
- Development of task inventories
  - Process time-consuming and often expensive
  - Access to previous inventories or analyses important determinants of the cost and success of the method
  - Table 3.2 shows some of the major steps and guidelines for developing task inventories

- Application of task analysis in selection
  - Several statistical techniques useful for rating data
  - Techniques involve calculating simple descriptive statistics—means, standard deviations, percentages—and applying predetermined decision rules for defining critical job tasks
  - Figure 3.7 shows two example tasks and some associated descriptive statistics

#### *Advantages and Disadvantages of Task Analysis*

- Advantages of task analysis
  - An efficient means of collecting data from large numbers of incumbents in geographically dispersed location
  - Lead to quantifying job analysis data—most valuable in analyzing jobs and determining core job requirements
- Disadvantages of task analysis
  - Time-consuming, somewhat expensive—motivation problems when inventories are long or complex
  - Respondents not representative of the desired employee population—work shift, gender, race, age—results may not be generalizable

## **CRITICAL INCIDENT TECHNIQUE**

[slides 34–40]

### **Description**

- Involves a series of behavioral statements developed by supervisors and other SMEs
- Based on direct observation or memory, describing incidents of good and poor work behaviors
  - important to differentiate successful from unsuccessful work performance
- Important components of the job serve as a basis for developing descriptive information about a job

### **Application of Critical Incidents**

- Technique serves a variety of selection purposes. Implementing the methods involves:
  1. Selecting the methods for critical incidents collection
  2. Selecting a panel of job experts
  3. Gathering critical incidents
  4. Rating and classifying critical incidents into job dimensions

### **Advantages and Disadvantages of Critical Incidents**

- Advantages
  - Information is behavioral in nature, not trait based
  - Described behaviors “critical” incidents, so information represents important aspects of the job
- Disadvantages
  - Incidents don't represent the *full* scope of the job
  - Process is labor intensive, results often situation specific
  - Doubtful that the information is transferrable from one setting to another

## Integrating a Task Analysis Inventory with Critical Incidents: A Suggestion

- Integrate the task analysis inventory with the critical incidents technique—basic steps:
  1. Identify important job tasks
  2. Identify important WRCs
  3. Show critical job task and WRC information to SMEs
  4. Rewrite the critical incident into a selection interview question
  5. Develop a key for scoring responses to the interview question

## SME Workshops

- Consist of groups or panels of 10 to 20 job incumbents per workshop session who work with a group leader to produce a job analysis. Steps to collect relevant job data:
  1. Select and prepare SMEs
  2. Identify and rate job tasks
  3. identify and rate WRCs
  4. Judge selection measure—job content relevance

## INCORPORATING JOB ANALYSIS RESULTS IN SELECTION PROCEDURES: A COOKBOOK

[slide 41]

- How do we use our collected information to develop or choose selection procedures? Refer back to Figure 3.1
  - Job analysis results determine the relevant WRCs needed for effective performance
  - These WRCs serve as the basis for *constructing* (developing questions for an employment interview) or *choosing* (purchasing a commercially available clerical ability test) needed selection procedures

## IDENTIFYING EMPLOYEE WRCs

[slides 42–61]

- Employers use both *direct* and *indirect* methods:
  - **Indirect** method uses specific steps to break down large inferential leaps involved in deriving critical WRCs from job tasks
  - **Direct** method requires larger inferential leaps because SMEs simply rate the importance of WRCs listed on a survey for an entire job, not individual tasks
- WRCs are useful only if they are accurate and complete

## Determining Employee WRCs

- The following sequential steps are taken in specifying WRCs:
  1. Identifying job tasks and work behaviors
  2. Rating importance of job tasks and work behaviors
  3. Specifying WRCs necessary for successful job performance

4. Rating importance of identified WRCs
5. Linking important WRCs to important job tasks and work behaviors
6. Developing content areas of selection procedures—a *selection plan*

### Determination of Selection Procedure Content

- The preparation of a selection plan occurs in two phases:

1. Determining the relative importance of WRCs
2. Choosing the selection procedures to assess these WRCs

#### 1. Determining the relative importance of WRCs

- SMEs might complete a survey
- Questionnaire might consist of a listing of essential WRCs previously identified—respondents assign a relative importance weight from 0 to 100
- Multiple WRCs' *importance ratings* by the *task importance* ratings for those tasks requiring the WRC

#### 2. Choosing selection procedures to assess employee WRCs

- A personnel decision maker choosing a selection measure should ask the following types of questions:
  1. Have job applicants demonstrated past behaviors or had experiences before taking the job that are associated with successful performance of the tasks of the job?
  2. Can job applicants be observed performing the job or part of it? Is there a means for simulating the job in a test situation that is likely to require important behaviors as defined by the job?
  3. Would a written test be best for examining worker requirements in terms of eliciting desired reactions and providing practical scoring?
  4. Would giving job applicants an opportunity to express themselves orally through an interview cover job requirements that might go unassessed using other means?
  5. Can the assessment method produce reliable and valid data for evaluating job applicants' possession of a WRC?
  6. Is it practical and within our resources to use a particular method for measuring a WRC?

### An Example Selection Plan for the Job of HR Selection Analyst

- When developing employee specifications, research seems to suggest that:

1. A structured, systematic approach should be used to reduce the size of inferential leaps made by job incumbents when rating their jobs
2. Incumbents can reliably and validly rate the specifications for their position when the specifications deal with specific, more observable job descriptors
3. Ratings on more abstract traits are improved when raters are properly trained using methods such as frame-of-reference training

## EMPLOYEE SPECIFICATIONS FOR JOBS THAT ARE ABOUT TO CHANGE OR HAVE YET TO BE CREATED

[slide 62]

- The method consists of the following steps:
  1. An analysis of the job is made to identify current tasks and WRCs
  2. SMEs—job incumbents, supervisors, managers—are assembled in a workshop to discuss future issues (e.g., technological change) likely to affect the job
  3. Information on expected future tasks and WRCs is collected from individuals knowledgeable about these expected job changes
  4. Differences between present and future judgments about the job are identified to isolate those tasks and WRCs for which the great change is anticipated—the basis for selecting incumbents in a job that does not current exist

## CONCLUSION

[slide 63]

Robert Guion, in commenting on the amount of detail and comprehensiveness of job analysis information required in HR selection research, said:

*The level of detail desired may also be determined by the likelihood of litigation. Prudent personnel researchers attend not only to the best wisdom of their profession but to the realities of the courtroom. In a specific situation, a detailed job analysis may not be necessary technically, but failure to have evidence of an “adequate” job analysis, in the view of the trial judge, may result in an adverse decision in court.*

## CHAPTER 4

# *Legal Issues in Selection*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Explain the laws and executive orders that compose Equal Employment Opportunity laws and their implications for a company's selection program.
2. Understand the evidence that a person who is claiming discrimination must demonstrate in order to make a successful charge against a company and the evidence that the company needs to demonstrate in order to defend its selection practices against such a charge.
3. Interpret the statistical data that are used by each side in a claim of discrimination concerning a company's selection program.
4. Understand the purpose and components of affirmative action programs and when they should be implemented by a company.
5. Explain the principles of the Uniform Guidelines on Employee Selection Procedures (1978) and how those principles have been interpreted in court cases focusing on discrimination in selection procedures.

### LEGAL ISSUES IN SELECTION

[slides 4–5]

- Legal policies address such issues as:
  - The records that must be kept on all employment decisions
  - Recording the movement of the various demographic group members through the steps of recruitment and selection
  - The methods for identifying the job relatedness of selection devices
- These tasks may be required to be used in the defense of the organization's employment decisions if a claim of discrimination is made
- Every selection program should have two objectives:
  1. Maximizing the probability of making accurate selection decisions about applicants
  2. Ensuring that all demographic groups are treated equally and fairly

### FEDERAL REGULATION OF BUSINESS

[slides 6–9]

- Title VII of the Civil Rights Act of 1964 was the first federal legislation to specify regulation across industries

- It addresses employment practices that are common across industries
  - Recruitment
  - Selection
  - Compensation
  - Training
  - Promotion of employees
- All laws that address employment discrimination are referred to as Equal Employment Opportunity (EEO) laws

### EEO Laws and Executive Orders

- EEO laws are federal laws designed to eliminate discrimination in HR management decisions
- EEO executive orders are statements made by the executive branch of the government intended for the same purpose as the laws, but are applicable only to organizations that do business directly with the government

### This and That Stuff

- Obesity—not a characteristic listed in any EEOC law or regulation, but morbid obesity (weight that is 100% or more over what is considered normal) is an impairment, and therefore can be considered a disability
- Smoking—not a protected characteristic—primarily a state law issue; most states allow employers to ban smoking in the workplace; in some states, employers can refuse to hire smokers
- Ethnic restaurants—issue whether or not an ethnic restaurant can select only those applicants in a specific demographic group, to be authentic; no such cases but selection likely constitutes national origin discrimination

## EMPLOYMENT DISCRIMINATION

[slides 10–22]

- Federal and state laws state principles concerning discrimination
- Discrimination is operationalized by
  - court decisions which yield specific comments about selection practices and set precedent for following cases
  - writings and actions of regulatory agencies such as EEOC's *Uniform Guidelines on Employee Selection Procedure*

### Discrimination Defined

- Two terms frequently used to reference equal employment laws—*adverse impact* and *discrimination*
  - **Adverse impact**—differences between demographically different individuals or groups with regard to the outcome of some selection procedure or process (e.g., male applicants may score higher on a physical ability test than female applicants)
  - **Discrimination**—occurs when adverse impact between demographic groups cannot be explained in terms of job-related reasons

- Two forms of discrimination:
  - **Disparate treatment**—different selection standards are applied to various groups of individuals even though there may not be an explicit statement made by the company using the different standards concerning preferences for one group rather than another (e.g. hiring minority group members for cleaning jobs in a restaurant while similarly qualified white people are made cashiers or waiters)
  - **Disparate impact**—selection standards are applied uniformly to all groups of applicants, but the net result is to produce differences in the selection of various groups (e.g. requirement of high school diploma, height minimums)

### Evidence Required in Court Case of Discrimination

- Plaintiff – starts case by presenting evidence of a *prima facie* case indicating discrimination. If this is done:
- Defendant – rebuts charges by providing defensible evidence and explanation for employment practices under question. If this is done:
- Plaintiff – attacks defendant's evidence and presents evidence that defendant's explanation was incorrect

#### *Disparate Treatment Evidence:*

- The McDonnell Douglas rule states that the plaintiff must show the following conditions exist:
  1. He or she belongs to a protected class—demographic group included in an EEOC law
  2. He or she applied and was qualified for the job
  3. Despite these qualifications, he or she was rejected
  4. After rejection, the position remained open and the employer continued to seek applicants from people who had the complainant's qualifications
- Employer provides clear and specific job-based explanation
- Plaintiff proves reason given is pretext

#### *Disparate Impact Evidence:*

- Plaintiff provides statistical evidence that shows employment practice negatively affects various groups differently; this establishes a *prima facie* case
- Employer demonstrates evidence that selection process is job-related—business necessity, bona fide occupational qualification, validation data
- Plaintiff proves an alternative practice is available that has less adverse impact

### The Use of Statistics

- Two main types of statistics used—*stock* and *flow*:
  - **Stock Statistics**—used to compare proportions of various demographic groups with regard to results of selection decisions (most common stock comparison is between the percentage of a specific demographic group and that same group in an external comparison group)

$$\frac{\text{number of women managers in organization}}{\text{total number of managers in organization}} \text{ vs. } \frac{\text{number of appropriately skilled women managers in labor force}}{\text{total number of appropriately skilled managers in labor force}}$$

- **Flow Statistics**—compares proportions that occur at two points in time
  - Most commonly used test is the *four-fifths rule*—the ratio of any group must be at least 80% of the *ratio of the most favorably treated group*

$$\frac{\text{number of minority applicants selected}}{\text{number of minority applicants}} \text{ vs. } \frac{\text{number of nonminority applicants selected}}{\text{number of nonminority applicants}}$$

### Definition of Internet Applicant

- Internet application can be through e-mail, résumé databases, job banks, electronic scanning technology, etc.
- An individual who satisfies all four of the following criteria as outlined by the OFCCP (Office of Federal Contract Compliance Programs):
  - The person must submit an expression of interest in employment through the Internet or related electronic devices
  - The organization is considering employing the individual for a particular position
  - The individual's expression of interest indicates that the person possesses the basic qualifications for the position
  - The individual must not remove herself from consideration for the position

## THE UNIFORM GUIDELINES ON EMPLOYEE SELECTION PROCEDURES

[slides 23–28]

### Determination of Adverse Impact

- The *Uniform Guidelines* only addressed to selection systems that produce adverse impact; if none exists no regulations concerning selection system—except recordkeeping
- The decision rule in judging discrimination is whether the use of a selection procedure leads to adverse impact as demonstrated by statistical tests—exceptions, including the four-fifths rule

### Selection Methods

- The *Uniform Guidelines* state that *any* method of selection that results in an employment decision is covered—not only scored selection tests—as the following statement indicates:

*When an informal or unscored selection procedure that has an adverse impact is utilized, the user should eliminate the adverse impact or modify the procedure to one which is a formal, scored, or quantified measure.*

### Options If Adverse Impact Exists

- For those selection programs that do have adverse impact, options of the organization are specified:
  - The organization may cease to use the selection devices under question and adopt others with no adverse impact—the organization may defend its practices by showing they are valid
  - If validation evidence is used, it should address the use of the selection instrument with regard to all groups being tested—should include a representative number of women and minorities
  - The organization should demonstrate that no other alternative selection programs are both valid and have less adverse impact

## Selection Requirements

- Skills and abilities easily learned during a brief training program not acceptable as selection requirements
- Requirements drawn from high level jobs permissible only if a majority of individuals move to the higher-level job within a reasonable time—seldom more than five years
- Cutoff scores used in selection must be justified

## Recordkeeping

- All organizations required to keep information about the demographic characteristics of applicants and those selected and to produce such information if requested
- Organizations with fewer than 100 employees should record by sex, race, and national origin the number of people selected, promoted, and terminated for each job level
- Selection procedures should be described

## Comments about the *Uniform Guidelines*

- The *Uniform Guidelines* have been given “great weight”—judges and juries have used the information when making decisions
- Selection researchers and managers have advanced our knowledge of several issues incorporated into the *Guidelines*—validation study in each selection site to ensure that all factors are included in the validation effort, but research demonstrates that validity can be accurately determined from several previous validity studies
- Evidence that a single validity study not a very good indicator of validity and could err by either overestimating or underestimating the validity coefficient

## AFFIRMATIVE ACTION PROGRAMS

[slides 29–32]

- A company would adopt an AAP in three situations:
  1. The company is a federal contractor
  2. The company has lost a court discrimination case or has signed a consent decree with the EEOC
  3. The company is voluntarily attempting to implement EEP principles

### Federal Contractor

- Requirement of affirmative actions for those with contracts of at least \$10,000
- OFCCP usually the regulatory agency
- The following three activities must be carried out by a contractor in an AAP:
  1. A utilization analysis is conducted
  2. Determine the goals that a company should strive to achieve; timetable should be developed
  3. Actions that are required to reach the goals must be taken

### Court Order or Consent Decree

- Legally required to engage in HRM actions that will directly lead to a balance between its workforce and the relevant labor market
- The company must meet specific numeric goals that have been stated and agree to in the court order or consent decree

## Voluntary AAP

- Potential conflict with Title VII's prohibition on discrimination-based employment decisions; *reverse discrimination* claims
- U.S. Supreme Court has ruled that voluntary AAPs must have the following characteristics:
  - Be temporary
  - Have no permanent adverse impact on white workers
  - Be designed to correct a demonstrable imbalance between minority and nonminority workers

## SELECTION COURT CASES

[slide 33]

- Table 4.5 summarizes a cross section of cases that deal with selection issues

## EEO AND JOB PERFORMANCE MEASUREMENT

[slides 34–35]

### Performance Measurement and the *Uniform Guidelines*

- Three sections of the *Guidelines* contain the direct statements about work measurement
  - Section 2.B—cited most often—makes clear that any procedures contributing to any employment decision is covered
  - Section 14.B (3)—scores on selection tests not to be treated as input to measure job performance; performance measure represents important or critical job behaviors; performance measures can be used even without referring to job analysis information
  - Section 15.B (5)—data used to identify and develop job performance criterion measures must be made explicit

### Court Decisions Addressing Performance Measurement

- Rulings made on the basis of the features of the appraisal systems
- Primarily concerned with the fairness of the appraisal system process
- Research examining differences in the average job performance scores of black and white employees—adverse impact may occur but may not be an indication of discrimination

## EEO SUMMARY

[slide 36]

- **Basis of discrimination**—race, color, religion, sex, national origin, age (older than 40), physical or mental handicaps
- **Evidence of discrimination**—based on a pattern of selection decisions over a period of time rather than on an isolated instance
- **Options of the organization**—if large selection differences in demographics groups:
  - Discontinue the current procedures and develop alternatives
  - Conduct a validation study to support the organization's contention that the selection instruments are job related

## CHAPTER 5

# *Recruitment of Applicants*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand the three key stages of recruiting.
2. Know what sources and information influence applicant choices at each stage.
3. Recognize differences across candidate “type,” such as college graduates, employed (passive) job candidates, and nontraditional workers (women, minorities, disabled, veterans, older candidates).
4. Understand how technology and globalization impact recruiting.
5. Recognize the impact the applicant pool (size, quality) has on the quality of selection decision.

### RECRUITMENT OF APPLICANTS

[slide 4]

- Recruitment is a distinctive process that is separate from, but parallel to, the selection process
- The more successful the recruiting process, the more the firm benefits from rigorous selection—larger applicant pool
- Recruitment critical to the organization because success of a company closely related to the quality of its workforce
- This chapter covers the best recruiting methods, use of new technologies—Internet, social media, key prehire applicant reactions and attitudes
- This chapter primarily focuses on external recruiting—requires generating interest when bringing a job opening to the attention of potential applicants not currently working for the organization

### A MODEL OF THE RECRUITMENT PROCESS

[slides 5–27]

- The recruitment process proceeds through a few well-defined sequential stages:
  1. Attracting and capturing the interest of potential applicants
  2. Maintaining applicant interest
  3. The organization’s decision to extend a job offer and postoffer closure
- The model recognizes the shifting contextual factors and applicant perceptions of organizational activities across these stages.

## Stage 1: Attracting and Generating Interest

- Stage 1 involves a number of key organizational decisions:
  - Establish its recruitment objectives—number of acceptable applicants, meeting the organization's legal and social obligations, increasing the success rate of the process
  - Strategic decisions for reaching objectives—type of applicants to pursue, target applicants broadly or focus on person-to-person methods, message orientation
  - Create a level of attraction for applicants to consider the firm's job opening—organizational reputation, organizational size and age, profitability, life cycle of firm and industry, product awareness, location, organizational social relationships, “celebrity effects”
- Applicant attraction also affected by job seekers' desires, values, self-image
- Experienced workers react differently than undergraduate college students
- Persuasion
  - Does the applicant pay attention?
  - Does the applicant understand the message?
  - Does it lead to relevant behavior change so that the applicant applies for the job?

### Targeting Strategies

- Types of individuals willing to apply or accept a position if offered—certain demographic groups or those with certain psychological characteristics
- Recruit broadly or to uniquely qualified people
  - Experienced workers?
  - College students?
  - Passive candidates?

### The Recruiting Message

- Recruit with a positive message or a realistic one which allows job seekers to opt out
- Decide how much job information to give
- For targeted recruiting:
  - For applicants with high cognitive ability, describe job as *challenging, stimulating*, requiring *intelligence*, etc.
  - For applicants high in conscientiousness, use *reliable, works hard, well-organized, self-disciplined*, etc.
  - To attract conscientious racial minorities, use *innovative* and *progressive*

### Recruiting Methods

- Effectiveness of recruiting methods revolves around three issues:
  - Media richness—multiple informational cues
  - Media credibility—expertise and trustworthiness
  - Degree of attention required during information processing—central processing requires more
- Importance of organizational reputation, PJ (person–job) fit and PO (person–organization) fit shift during the process
- Considerations when evaluating recruiting methods—employee referrals and walk-ins, use of more informal methods has potential to create disparate impact, magnitude of different methods tends to be relatively small

- The takeaway point regarding recruitment methods is:
  - Use informal, personal recruitment methods in cases in which voluntary turnover is a problem during employees' first two years of employment
  - However, a cautionary note: minorities use formal recruitment methods more frequently than informal ones
  - A related point: word-of-mouth information about jobs and employers also has been identified as an important source of recruits
- Ways to use the Web for recruiting:
  - Posting jobs on a company's home page
  - Posting jobs on websites devoted to job announcements—GlassDoor, Indeed, Monster, Vault
  - Posting links to a job advertisement on social networking sites—Twitter, Facebook, YouTube
  - Screening for applicants using available search options on business social networks—LinkedIn
- Ways to use the Web for recruiting – best results:
  - Use individuals—current and past employees—who know your organization and the job to refer applicants
  - Develop a recruiting website
  - Use employee testimonials
  - Find active candidates
  - Find passive candidates
  - Recruit passive job candidates
- Research indicates that 77 percent of employers recently reported using social networking sites to research job candidates
  - More than a third (35%) reported finding information that led them to not hire a candidate
    - professional image, company culture fit
  - May result in disparate impact

## Stage 2: Maintaining Applicant Interest

- The organization must manage three aspects of recruiting well:
  - Recruiter has to be an effective administrator of this process
  - Two-way communication eases applicant anxiety
  - Think strategically about the information conveyed—applicants begin to transition from focusing on organizational reputation and attraction to fit with the organization

### *Applicant Self-Selection*

- Applicant can decide to withdraw from further consideration—poor fit based on perceptions of job
  - Fit perceptions may be incorrect
  - Firm must present accurate fit information during recruiting process
  - Fairness of recruiting based on perceptions of:
    - a. Procedural justice—consistency of administration
    - b. Information justice—provision of (un)timely feedback
    - c. Interpersonal perceptions of justice—how recruiter treated the applicant

*Recruiter Characteristics*

- Recruiters usually influence early opinions and behaviors of recruits
- Recruiters who have race and gender in common with applicants may have some influence
- Job and organization characteristics have a stronger influence on recruits than do recruiters
- Training of recruiters important
- Recruiters often seen as less trustworthy and credible than members of intended work group
- Recruiters perceived as competent, informed, trustworthy, warm, personable, concerned are regarded favorably
- Recruiters can have a strong, negative influence on applicants
- Recruiters vary in the way they mix recruitment and selection components of applicant interviews

*The On-Site Visit Also Matters*

- Communicates more useful comprehensive information about the job, work group, organization
- Details of the site visit that matter to the candidate:
  - Likability of the employee hosting the visit
  - Opportunity to meet high-level executives and interact with potential supervisor and coworkers
  - Visit arrangements for a more effective site visit – hotel accommodations, well-organized schedule, being able to bring a spouse

*Internal Selection*

- Identifying pivotal jobs that are critical to the firm's strategic goals
- Multinational enterprises rely on expatriates to fill top job in overseas affiliates
- Hiring interns—internship seen as a “working interview”
- Internal recruitment nominated by a supervisor affected by:
  - Whether supervisor is reluctant to lose a good employee to another department
  - Whether supervisor sees value in establishing network across the firm

**Stage 3: Securing Offer Acceptance and Postoffer Closure**

- Factors that may influence the applicant's decision:
  - Research indicates that early in recruiting compensation (71%), organizational culture (63%), the work itself (52%), and benefits (52%) would be important factors in future job choice
  - After the offer was extended and accepted, these factors not as important as anticipated
    - Compensation before hire—71%
    - Compensation after hire—19%

*Administration of Recruitment*

- Proper administration of this complex process will significantly improve it
  - Prompt responses by the company to the applicant indicate that the company is efficient and a desirable place to work
  - Highly qualified applicants who have multiple job alternatives are most strongly affected by delays in communications

### *Technology's Role in Recruiting*

- Using the Internet for recruiting improves speed, efficiency, and effectiveness, but
  - Legal issues—discrimination and invasion of privacy
  - Practical issues—sifting through a larger pool of applicants
  - Scientific issues—increase the number of highly qualified applicants
- Organizations must present website content in an easily usable, interactive, yet engaging manner

### *Global Recruiting*

- Two critical issues that are likely to affect global recruiting success
  - Cross-cultural differences and variations in socioeconomic conditions across regions
  - Use of referral programs is much higher in some countries than in the U.S.
- Which recruiting method is more frequently used differs for blue-collar jobs versus white-collar jobs—again, these practices can differ country to country

### *Recruiting Metrics*

- The last part of recruiting is to determine the results of recruiting—the number of people who applied, their skill levels, their diversity, the time it took to hire recruits, how many actually became employed. Several metrics suggested:
  - New employee job performance
  - New employee turnover
  - New hire failure rate
  - Manager satisfaction with new employees
  - New employee training success
  - Cost per hire

## **SUMMARY RECOMMENDATIONS FOR ENHANCING RECRUITMENT**

[slides 28–29]

1. Carefully consider the sometimes-competing objectives of attracting versus selecting job seekers
2. Identify specific objectives of a recruitment program
3. Formulate a specific strategy for accomplishing recruitment program objectives
4. Ensure that recruitment screens reflect the desired job-related characteristics of candidates
5. Identify who represents the potential applicant population
6. Use targeted recruitment to identify underrepresented protected groups
7. Develop a recruitment website that is attractive, easy to navigate
8. For organizations using websites and who receive large numbers of résumés, consider using résumé screening and tracking software
9. Evaluate an organization's image being communicated by recruitment and related media
10. Consider using recruitment websites for organizations searching for employees
11. Recognized that all your employees are “recruiters” – encourage referrals of potential applicants
12. Select recruiters who can play an important role in attracting initial interest in organizations on characteristics associated with dealing with people

13. Use technology to aid in recruiting, but use more personal means of communications to contact applicants
14. Use RJPs (realistic job previews)—positive results more likely to occur
15. Use metrics and other measures to evaluate recruitment programs and identify “what works”

## CHAPTER 6

# *Human Resource Measurement in Selection*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand the role of measurement in human resource (HR) selection.
2. Explain four different scales of measurement and the conclusions that can be legitimately drawn from data produced by each scale.
3. Understand why standardized measures are important to use in HR selection.
4. Know how to research a published test being considered for use in HR selection.
5. Understand the fundamentals in developing and implementing a selection procedure.
6. Interpret the meaning of percentile norms.

### FUNDAMENTALS OF MEASUREMENT: AN OVERVIEW

[slide 4]

An assumption in selection decision making is that information is available for making these decisions.  
But

- What types of information can be used?
- Where does this information come from?
- What characteristics should this information have to be most useful?

This chapter specifically focuses on:

- the basics of psychological measurement as they apply to HR selection
- locating, developing, and interpreting measures commonly used in HR selection

### THE ROLE OF MEASUREMENT IN HR SELECTION

[slides 5–14]

#### The Nature of Measurement

- From the perspective of HR selection:
  - HR measurement involves the systematic application of rules for assigning numbers to objects—usually people—to represent the quantities of a person's attributes or traits
  - *Rules* suggest that the basis for assigning numbers is clearly specified and consistently applied—it is important that different users who use a test administer it under the same conditions and score it in the same manner as all other users

- Physical attributes—gender—can be assessed through direct observation
- Psychological attributes (constructs)—conscientiousness, intelligence—are not directly observable and must be inferred from a score – numbers or units of measurement

#### *Criteria and Predictors in Selection Research*

- Identification of two types of variables is critical:
  - *Criterion*—measure of what represents successful performance on a job (supervisory ratings of work performance)
  - *Predictor*—a measure (of employee WRCs) used to forecast or predict the likelihood of job candidates' success on a job

#### *Measurement and Individual Differences*

- Measuring individual differences helps identify those individuals who should be hired for a job
- Figure 6.1 shows a hypothetical distribution of quantity of output per worker for a large number of workers:
  - Individual employees differ in their levels of productivity
  - Relatively few produce a very large or very small number of baskets
  - We need a selection procedure that will predict productivity

### **Scales of Measurement**

- A means by which individuals are distinguishable from one another on a variable such as a predictor or criterion
- Predictor or criterion variables can differ dramatically in their precision—e.g. trainability
- Figure 6.2 shows hypothetical distributions of trainees' scores for two methods of measuring trainability
- Four types of scales or levels of measurement exist:
  - a. Nominal—composed of two or more mutually exclusive categories (e.g., male or female)
  - b. Ordinal—rank-orders objects (individuals) from “high” to “low” (e.g., test scores as percentiles) (Figure 6.3)
  - c. Interval—uses constant units of measurement—differences between numbers take on meaning (Figure 6.4)
  - d. Ratio—has an absolute zero point—differences between numbers also have meaning (e.g., most scales involving physical measurements)
- The degree of precision increases as we move from nominal to ratio scales (Figure 6.5)

### **STANDARDIZATION OF SELECTION MEASUREMENT**

[slides 15–33]

- The systematic application of pre-established rules or standards for assigning scores to the attributes or traits of an individual
- A selection measure may provide information to be used as either a predictor or a criterion
  - differences in scores must be attributable to ability, and not to other factors

- A predictor or criterion measure is standardized if it possesses:
  1. Content—all people assessed are measure by the same information or content
  2. Administration—information is collected the same way in all locations and across all administrations
  3. Scoring—rules for scoring are specified before administering the measure and applied the same way with each scoring

### Measures Used in HR Selection

- Criteria are employed as part of a research study designed to determine which selection procedures are related to job success and should be used in selection decision making—a validation study
- Criterion measures help serve as a standard for evaluating how well predictors do the job they were intended to do
  - Predictors have a *direct* impact on decisions—a manager reviews an applicant's scores
  - Criteria play an *indirect* role—which selection procedures should be used in making selection decisions

#### *Predictors or Selection Procedures*

- Numerous types but in general the major types fall into three broad categories:
  1. Background information – application forms, training and experience evaluations, reference checks, biographical data used to collect information
  2. Interviews—employment interviews used to collect additional information
  3. Tests—hundreds of tests have been used for selection purposes—aptitude, ability, achievement, personality

#### *Criteria or Measures of Job Success*

- Measurement methods used to collect data include the following:
  1. Objective production data—physical measures of work
  2. Personnel data—personnel records and files can serve as criterion measures
  3. Judgmental data—performance appraisals or ratings
  4. Job or work sample data—obtained from a measure of specific aspects of the work process or outcomes
  5. Training proficiency data—how quickly and how well employees learn during job training activities

### Standards for Evaluating Selection Measures

- Table 6.1 lists some of the factors to be considered when choosing or developing a selection measure
- If a measure doesn't meet these standards:
  - a. Determine whether you can adjust your data or the way you calculate the score itself so it will meet each measurement evaluation criterion
  - b. If this option is not viable, find or develop another, more suitable measure
- The factors listed in Table 6.1 are not of equal importance

## Finding and Constructing Selection Measures

- Once we know the criteria selection measures for successful work performance, the process of identifying and implementing the selection procedures may begin:
  - A consultant—industrial–organizational psychologist—is needed
  - Two choices:
    - a. We can locate and choose from existing selection measures
    - b. Construct our own
  - We might need to consider both options

## Locating Existing Selection Measures

- Advantages of using existing measures:
  1. Use of existing measures usually less expensive and less time-consuming than developing new ones
  2. If previous research was conducted on these measures, we will have some idea about the reliability, validity, and other characteristics of the measure
  3. Well-developed, existing measures often superior to what could be developed in-house

### *Information Sources for Existing Measures*

- Sources in print and on the Internet (Tables 6.2 and 6.3)
  - Text and reference sources—provide excellent reviews of predictors and other measures that have been used
  - Buros' Mental Measurements Yearbooks—the most important source for information on tests for personnel selection
  - Other reference sources—journals, test publishers (levels A, B, C), professional associations

### *Other Reference Sources*

- Suggestions for choosing an existing selection measure:
  1. Be sure you clearly and completely understand the attribute or construct you want to measure—decide on the best means for assessing the attribute
  2. Search for and read critical reviews and evaluations of the measure—Buros' Mental Measurements Yearbook an excellent source
  3. If a specimen set of the measuring is available, order and study it
  4. Once steps 1-3 are completed, ask “are there more compelling arguments for using this measure? Or, are there compelling arguments against using it?”

## Constructing New Selection Measures

- Developing of measures is a complex, resource-consuming process that usually requires expert advice
- Consultants likely will be needed
- Knowledge of the basic issues involved in selection measure development, validation, and application can help bridge any possible communications gap between the organization and the consultant

### *Steps in Developing Selection Measures*

1. Analyze the job for which a measure is being developed
2. Select the method of measurement to be used (Figure 6.6)
3. Plan and develop the measure (Figures 6.7 and 6.8)
4. Administer, analyze, and revise the preliminary measure
5. Determine the reliability and validity of the revised measure for the jobs studies
6. Implement and monitor the measure in the HR selection system

## INTERPRETING SCORES ON SELECTION MEASURES

[slides 34–38]

### Using Norms

- To interpret the results of measurement, we need two pieces of information:
  - a. How others scored on the selection procedure
  - b. Validity of the selection procedure
- Predictor scores of relevant others in groups are called norms
- Keep the following in mind when using norms to interpret scores:
  - a. The norm group selected should be *relevant*
  - b. Accumulate and use *local* norms when appropriate
  - c. Norms are *transitory*—specific to the point in time when they are collected

### Using Percentiles

- Percentile scores show the percentage of a persons in a norm group who fall below a given score on a measure—a percental score is not a percentage score
- The higher the percentile score, the better a person's performance relative to others (Figure 6.9)
- Percentile scores are useful in interpreting test scores, but are subject to misuse

### Using Standard Scores

- Standard scores represent adjustments to raw scores so it is possible to determine the proportion of individuals who fall at various standard score levels—these scales indicate how far above or below the mean score any raw score is
- Some of the more common are *z*, *T*, and *stanine* scores
  - *z* score the most common—can be obtained for all individuals for whom test score data are available
  - *T* scores – similar to *z* scores but adjusted so that all T scores are positive
  - *Stanine* scores—computed by rank-ordering scores from lowest to highest

## CHAPTER 7

# *Reliability of Selection Measures*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Explain the meaning of reliability and why it is important in human resources selection.
2. Contrast the concepts of true scores and errors of measurement for selection procedures.
3. Compare and contrast several methods for estimating reliability.
4. Explain what a reliability coefficient means.
5. Understand why certain factors affect a reliability estimate.
6. Explain why the standard error of measurement is important in comparing individuals' scores on a predictor.

### WHAT IS RELIABILITY?

[slides 4–21]

- Degree of dependability, consistency, or stability of scores on a predictor of criterion used in HR selection
- Figure 7.1 shows an example involving the dependability of information, in the context of selection, after a computer programming aptitude test was administered to 10 individuals applying for a job as a computer programmer
- The test was readministered after the first test results went missing; results show that each of the applicants had different scores for the two tests—test results are not consistent, therefore not reliable

### A Definition of Reliability

- The term *reliability* has a host of definitions, but in the context of HR selection it simply means the degree of dependability, consistency, or stability of scores on a measure used in selection research—predictors or criteria
- Reliability of a measure is determined by the degree of consistency between two sets of scores on the same measure
- If such scores are inconsistent, then “errors of measurement” are present

## Errors of Measurement

- Reliability deals with *errors of measurement* – free of errors—but none of our selection measures will be free of measurement errors
- Selection measures designed to assess important job-related characteristics—knowledge, skills, personality traits—may be prone to error due to the sample of items used, the test taker, the examiner, or the situation in which testing takes place
- We want to know the “true” scores of applicants for each characteristic being measured—unless our measure is perfectly reliable, we will encounter difficulties in knowing precisely these true scores
- The score obtained on a measure—obtained score/raw score—consists of two parts: a *true* component and an *error* component

### True Score

- A hypothetical score for a person assuming no errors of measurement were present at the time of measurement or scoring
  1. Individuals answered correctly the same percentage of problems on the test that they would have if *all possible* problems had been given and the test were a construct valid measure of the underlying phenomenon of interest
  2. Individuals answered correctly the problem they actually knew without being affected by external factors—lighting or temperature of the room in which testing took place, emotional state, physical health

### Error Score

- Represents errors of measurement—those factors that affect obtained scores but are not related to the characteristic, trait, or attribute being measured
- These factors distort respondents' scores either over or under what they would have been on another measurement occasion—fatigue, anxiety, noise during testing
- Figure 7.2 shows the relationship between reliability and errors of measurement for three levels of reliability of a selection measure
- Table 7.1 summarizes some of the more common sources of error that contribute to the unreliability of selection measures

## Methods of Estimating Reliability

- We cannot measure reliability; we can only estimate it
- Statistical procedures are commonly used to calculate what are called reliability coefficients—an index of relationship
  - Summarizes the relationship between two sets of measures for which a reliability estimate is being made
  - The calculated index varies from 0.00 to 1.00—the correlation coefficient obtained is regarded as a direct measure of the reliability estimate
  - The higher the coefficient, the less the measurement error and the higher the reliability estimate
  - With high reliability, more confidence that a particular measure is giving a dependable picture of true scores for attribute being measured

- Four principle methods most often employed in selection research studies:
  - a. Test-retest – the same measure used to collect data from the same respondents at two different points in time
  - b. Parallel or equivalent forms—two versions of a selection measure collected from the same respondents at two different times, scores on the two forms then correlated
  - c. Internal consistency reliability estimate—shows the extent to which all parts of a measure are similar in what they measure (split-half reliability, Kuder-Richardson—K-R 20—reliability, Cronbach's coefficient alpha ( $\alpha$ ) reliability)
  - d. Interrater reliability estimates—the determination of consistency or agreement among raters (interrater agreement, interclass correlation, intraclass correlation)

## INTERPRETING RELIABILITY COEFFICIENTS

[slides 22–32]

### What Does a Reliability Coefficient Mean?

- Specific to the reliability estimation method and group on which it was calculated
- A necessary but not a sufficient condition for validity
- Based on responses from a group of individuals
- Expressed by degree
- Determined ultimately by judgment

### How High Should a Reliability Coefficient Be?

- There is not generally agreed upon value above which reliability is acceptable and below which it is unacceptable
- The more critical the decision to be made, the greater the need for precision of the measure on which the decision will be based, and the higher the required reliability coefficient
- Imprecise predictors can have long-term consequences for an organization—dependable predictors are essential for accurately evaluating key personnel
- Criterion measures should be reliable, however their reliability need not be as high as predictors for them to be useful
- Test users must consider the specific circumstances surrounding their situations to determine how much measurement error they are willing to put up with – is the reliability coefficient adequate?

### Factors Influencing the Reliability of a Measure

- Method of estimating reliability (Figure 7.9)
- Individual differences among respondents
- Stability
- Sample
- Length of a measure (Figure 7.10)
- Test question difficulty (Figure 7.11)
- Homogeneity of a measure's content
- Response format
- Administration and scoring of a measure

## Standard Error of Measurement

- Reliability is a group-based statistic
- To obtain an estimate of the error for an individual, we can use the standard error of measurement—a number in the same measurement units as the measure for which it is being calculated
- Formula for calculating standard error:

$$\sigma_{meas} = \sigma_x \sqrt{1 - r_{xx}}$$

$\sigma_{meas}$  = the standard error of measurement for measure  $X$

$\sigma_x$  = the standard deviation of obtained scores on measure  $X$

$r_{xx}$  = the reliability of measure  $X$

- To interpret differences in individuals' scores:
  1. The difference between two individuals' scores should not be considered significant unless the difference is at least twice the standard error of measurement of the measure
  2. Before the difference between scores of the same individual on two different measures should be treated as significant, the difference should be greater than twice the standard error measurement of either measure

## Evaluating Reliability Coefficients

- The Buros Institute of Mental Measurements reviewed more than 1,000 commercially available tests published in *The Eighth Mental Measurements Yearbook*. For the tests listed, the Institute found:
  - Over 22% appeared without any reliability information
  - 7% showed neither reliability nor validity data
  - 9% showed no reliability data for certain subtests or forms
  - 28% did not report any normative data

## Reliability: A Concluding Comment

- Even though the assessment and interpretation of reliability can be complex, it is a fundamental element to the proper use of HR selection measures
- The validity of a measure depends on its reliability—reliability of predictor scores and criterion scores is necessary, but not sufficient, for a score's validity or interpretation
- Knowledge of reliability information and other associated statistics are critical for making accurate assessments and decisions about individual seeking employment

## CHAPTER 8

# *Validity of Selection Procedures*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Explain the difference between and importance of a selection procedure's reliability and validity.
2. Know the major steps a consulting firm should take and the deliverables it should provide if your company has contracted with the firm to undertake a selection procedure validation study.
3. Explain the differences among the major types of validation strategies.
4. Understand how it is possible to use statistical methods and validated selection procedures to predict the future, that is, in terms of job-related employee behaviors.
5. Communicate to managers and executives the meaning and importance of a statistically significant validity coefficient for a selection procedure.

### AN OVERVIEW OF VALIDITY

[slides 4–7]

- In this chapter, we focus on *validity*—its relation to reliability, and the principal analytic strategies available for determining the validity of selection procedure data
- Validity represents the most important characteristic of data produced from measures used in HR selection
- Like reliability, the importance of validity applies to both selection procedures as well as criteria
- Validity shows what is assessed by selection measures and determines the kinds of legitimate inferences or conclusions we can draw from data such measures produce

### Validity: A Definition

- Validity concerns the accuracy of judgments or inferences made from scores on selection measures, including predictors and criteria
- We want to know the accuracy of hypothesized predictions about employee work behaviors for job success
- The research process for discovering what and how well a selection procedure measures is called *validation*—involves the research processes we go through in testing the appropriateness of our inferences from our selection procedures

## The Relation between Reliability and Validity

- It is possible to have a measure that is reliable yet does not assess what we want for selection
- The quantitative relationship between validity and reliability is:

$$r_{xy} = \sqrt{r_{xx} r_{yy}}$$

$r_{xy}$  = maximum possible correlation between predictor  $X$  and criterion  $Y$ , the validity coefficient

$r_{xx}$  = reliability coefficient of predictor  $X$

$r_{yy}$  = reliability coefficient of criterion  $Y$

## Types of Validation Strategies

- A *validation study* provides the evidence for determining the legitimate inferences that can be made from scores on a selection measure
- Three classical approaches used for validating measures in HR selection:
  - a. Content validation
  - b. Criterion-related validation—includes both *concurrent* and *predictive* validation strategies
  - c. Construct validation

## CONTENT VALIDATION STRATEGY

[slides 8–19]

- A selection procedure (test) or a criterion (performance evaluation) has *content validity* when it is shown that its content (items, questions, behaviors, etc.) representatively samples the content or aspects of the job associated with successful performance
- “Content of the job” is a collection of job behaviors and the associated knowledge, skills, abilities, and other characteristics (competencies, personality, physical requirements, licenses, certifications, etc.) necessary for effective work performance
- Emphasizes the role of expert judgment in determining the validity of a measure rather than relying on statistical methods
- *Judgments* used to describe the degree to which content of a selection method reflects important aspects of work performance
- *Face validity* sometimes confused with the concept of content validity
  - Content validity deals with the representative sampling of the content domain of a job by a selection measure
  - Face validity concerns the *appearance* of whether a measure is measuring what is intended
- Perceived face validity of selection procedures the strongest correlate of participants’ beliefs regarding both the procedure’s effectiveness in identifying qualified people and the procedure’s fairness

## Major Aspects of Content Validation

1. Conducting a comprehensive job analysis
  - a. Describing the tasks performed on the job
  - b. Measuring criticality and/or importance of the tasks
  - c. Specifying WRCs required to perform these critical tasks

- d. Measuring the criticality and/or importance of WRCs
    - operational definition of each WRC
    - relationship between each WRC and job task
    - complexity/difficulty of obtaining each WRC
    - whether an employee is expected to possess each WRC
    - whether each WRC is necessary for successful job performance
  - e. Linking important job tasks to WRCs
2. Selecting experts participating in a content validation study
  3. Specifying selection procedure content
    - a. Selection procedure as a whole
    - b. Item-by-item analysis
    - c. Supplementary indications of content validity—predictor validity
  4. Assessing selection procedure and job content relevance

### Some Examples of Content Validation

- A reading skills test based on actual safety procedures and operating procedures employees need to read upon job entry and their *importance* to work performance
- Job-related employment interview questions to assess work performance dimensions
- For content validity, derive test content from what incumbents do on the job (Figure 8.1)

### Appropriateness of Content Validation?

- In Section C1 of the *Uniform Guidelines*, it is noted:

*A selection procedure based upon inferences about mental processes cannot be supported solely or primarily on the basis of content validity. Thus, a content strategy is not appropriate demonstrating the validity of selection procedures which purport to measure traits or constructs, such as intelligence, aptitude, personality, commonsense, judgment, leadership, and spatial ability.*

- Recently, however, some industrial psychologists agree that “content validity is appropriate scientifically and professionally for use with tests of specific cognitive skills used in work performance”

#### *Job Analysis and Content Validation*

- Figure 8.2 summarizes major inference points that take place when using job analysis to help establish the content validity of a selection procedure
  - Point 1 is from the job itself to the tasks identified as composing it
  - Point 2 is from the tasks of the job to identified WRCs
  - Point 3 is the most critical – final judgments regarding content validity of the selection measure are made
- To make the inferential leap supporting content validity, three important issues that contribute to physical and psychological fidelity must be addressed:
  1. Does successful performance on the selection procedure require the same WRCs needed for successful work performance?
  2. Is the mode used for assessing performance on WRCs the same as that required for job or task performance?
  3. Are WRCs *not* required for the job present in the predictor?

- *Uniform Guidelines* specify some situations in which content validation alone is not appropriate—in these situations, other validation methods must be used:
  1. When mental processes, psychological constructs, or personality traits—judgment, integrity, dependability, motivation—are not directly observable but inferred from the selection method
  2. When the selection procedure involves WRCs an employee is expected to learn on the job
  3. When the content of the selection device does not resemble a work behavior or the work setting

#### *How Content Validation Differs from Criterion-Related Validation*

- a. In content validity, the focus is on the selection procedure and its manifest relation with the job content domain, whereas in others the focus is on the relations of the selection procedure with an external criterion
- b. Criterion-related validity is narrowly based on a specific set of data, whereas content validity is based on a broader base of data and inference
- c. Criterion-related validity is couched in terms of quantitative indices, whereas content validity is characterized using broader, more judgmental descriptors

## CRITERION-RELATED VALIDATION STRATEGIES

[slides 20–32]

Two approaches typically undertaken when conducting an empirical, criterion-related study:

- *Concurrent validation*
  - Information obtained on both a predictor and a criterion for a *current* group of *employees*
  - Validity of the inference signified by a statistically significant ( $p \leq 0.05$ ) relationship
  - Figures 8.3 and 8.4 demonstrate an example of a concurrent validation study
- *Predictive validation*
  - Involves the collection of data over time
  - Job *applicants* rather than job incumbents serve as the data source
  - Figure 8.5 illustrates five variations in which a predictive study might be conducted

Table 8.1 outlines the basic steps taken in both concurrent and predictive validation studies

### Concurrent Validation

#### *Strengths and Weaknesses*

Several factors mitigate usefulness of a concurrent validation study:

- a. Availability of a large sample working in comparable settings who will participate in the study
- b. Differences in job tenure or length of employment
- c. Representativeness of present *employees* to job *applicants*
- d. Certain employees failing to participate
- e. Motivation of employees to participate or employee manipulation of answers

## Predictive Validation

### *Strengths and Weaknesses*

- a. Because of the inference tested by predictive validation, the method is appropriate for measures used in HR selection
- b. Predicts how well job applicants will *be able* to perform on the job
- c. One big weakness is the time interval required to determine the validity of the measure being examined
- d. Can be difficult to explain to managers the importance of filing selection measure information before using the data for HR selection purposes

## Concurrent versus Predictive Validation Strategies

- Generally assumed that a predictive validation design is superior to a concurrent one because it more closely resembles an actual employment situation—predictive designs have been thought to provide a better estimate of validity
- Minimal differences found in the validation results of two types of designs – another review revealed no significant differences
- For ability tests, studies suggest that a concurrent validation approach is just as viable as a predictive one
- Studies have reported different results for predictive versus concurrent validation designs for both personality and integrity measures

## Requirements for a Criterion-Related Validation Study

- At least four requirements are necessary before a criterion-related study should be considered:
  1. The job should be reasonably stable and not in a period of change or transition
  2. A relevant, reliable criterion that is free from contamination must be available or feasible to develop
  3. It must be possible to base the validation study on a sample of people and jobs representative to which the results will be generalized
  4. A large enough, and representative, sample of people on whom both predictor and criterion data have been collected must be available

## Stability of Criterion-Related Validation Over Time

- A review indicated that the predictive validity of some measures rapidly decayed over time
- Critics of the review noted that only one study reviewed incorporated an ability test to predict actual performance on the job
- Another study found that the predictive validity of mental ability tests actually increased over time, job experience validity decreased, and predictive validity remained about the same for dexterity tests

## The Courts and Criterion-Related Validation

- There is no guaranteed outcome of a legal case—only 5 of 12 defendants won their case. Among the findings:
  1. Some courts preferred to judge validity on the basis of format or content of the selection instrument
  2. Some courts were swayed by a test's legal history even though existing evidence was available of the validity of the test, others influenced by the type of test used
  3. Some judges had preferences for the use of a predictive validation strategy versus a concurrent one
  4. A statistically significant validity coefficient alone did not guarantee judgment for the defendant
  5. Judges differed on their willingness to accept statistical corrections to predictor scores or correction for unreliability of the criterion

## Content versus Criterion-Related Validation: Some Requirements

- Drawing from the *Uniform Guidelines, Principles for the Validation and Use of Employee Selection Procedures*, and other sources, the major feasibility requirements for conducting content and criterion-related (concurrent and predictive) validation methods are summarized in Table 8.2
- The requirements are not meant to be exhaustive, only illustrations of major considerations when HR selection is involved

## CONSTRUCT VALIDATION STRATEGY

[slides 33–36]

- Psychologists use the term *construct* to refer to a theoretical psychological concept, attribute, characteristics, or quality
- When a psychological test is used in selection research, it assesses a construct—intelligence, sociability, clerical ability are all theoretical abstracts called constructs
- Specific measures are operational measures hypothesized to represent a specific construct
- Construction validation helps us determine whether a measure does indeed reflect a specific construct
- Figure 8.6 shows the hypothesized links between the constructs and their measures
- The example highlights the major steps for implementing a construct validation study:
  1. The construct is carefully defined and theoretically developed—hypotheses are formed concerning the relationships between the construct and other variables
  2. A measure hypothesized to assess the construct is developed
  3. Studies testing the hypothesized relationship between the constructed measure and other, relevant variables are conducted
- Results of studies such as the following are particularly warranted in construct validation research:
  1. Intercorrelations among the measure's items, questions, etc. should show whether the items cluster into one or more groupings
  2. Items of the measure belonging to the same grouping should be internally consistent or reliable

3. Different measures assessing the same construct as our developed measure should be related with the developed measure
4. Content validity studies show how experts have judged the manner in which items, questions, etc. of the measure were developed and how these items sample the job content domain

## EMPIRICAL CONSIDERATIONS IN CRITERION-RELATED VALIDATION STRATEGIES

[slides 37–61]

Even when we have conducted validation studies on a selection procedure, we will want to answer two important questions:

1. Is there a relationship between applicants' or employers' responses to the selection procedures and their performance on the job?
2. If so, is the relationship strong enough to warrant the measure's use in employment decision making?

### Correlation

#### *Computing Validity Coefficients*

- A validity coefficient is an index that summarizes the degree of relationship between a predictor and criterion
- Table 8.3 shows example data from a hypothetical collection of sales ability score (predictor) and job performance rating (criterion) on 20 salespeople – an example scattergram of the data is shown in Figure 8.7
- If a validity coefficient is not statistically significant, then the selection measure is not a valid predictor of a criterion (Figure 8.8)

#### *Importance of Large Sample Sizes*

1. A validity coefficient computed on a small sample must be higher in value to be considered statistically significant than a validity coefficient based on a larger sample
2. A validity coefficient computed on a small sample is less reliable than one based on a large sample—greater variability
3. The chances of *detecting* that a predictor is truly valid is lower for small sample sizes than for large one

#### *Interpreting Validity Coefficients*

- By squaring the validity coefficient, we obtain an index—*coefficient of determination*—that indicates our test's ability to account for individual performance differences
- The coefficient of determination represents the percentage of variance in the criterion that can be explained by variance associated with the predictor
- In addition to the coefficient of determination, *expectancy tables* and *charts* are useful
- *Utility analysis* can be used—its computation is far more complex

## Prediction

- A statistically significant validity coefficient is helpful in showing that for a *group* of persons a test is related to job success
- For individual prediction purposes, linear regression and expectancy charts can be used to aid in selection decision making—these tools should be employed only for those predictors that have a statistically significant relationship with the criterion

### *Linear Regression*

- Involves the determination of how changes in criterion scores are related to changes in predictor scores
- A regression equation is developed that mathematically describes the functional relationship between the predictor and criterion
- Two common types of linear regression—*simple* and *multiple*
  - *Simple* regression – only one predictor and one criterion (Figure 8.9 shows the *regression line* which summarizes the relationship between inventory scores and work performance ratings)
  - *Multiple* regression—assumes *two or more* predictors used to predict a criterion

### *Cross-Validation*

- Involves the following steps:
  1. A large group of people on whom predictor and criterion data are available is randomly divided into two groups
  2. A regression equation is developed on one of the groups—the “weighting group”
  3. The equation is used to predict the criterion for the other group – the “holdout group”
  4. Predicted criterion scores are obtained for each person in the holdout group
  5. For people in the holdout group, *predicted* criterion scores are then correlated with their *actual* criterion scores

### *Expectancy Tables and Charts*

- An expectancy *table* is a table of numbers that shows the probability that a job applicant with a particular predictor score will achieve a defined level of success
- An expectancy *chart* presents essentially the same data except that it provides a visual summarization of the relationship between a predictor and criterion
- The construction of expectancy tables and charts is a five-step process:
  1. Individuals on whom criterion data are available are divided into two groups—superior performers and others
  2. For each predictor score, frequencies of the number of employees in each group are determined
  3. The predictor score distribution is divided into fifths
  4. The number and percentage of individuals in each group are determined for each “fifth” of the predictor score distribution
  5. An expectancy chart that depicts these percentages is then prepared
- Figure 8.10 shows the scattergram of the interview scores plotted against the performance ratings
  - Table 8.4 is the expectancy table developed from the plotted data

- Two types of expectancy charts can be prepared—*individual* and *institutional*
  - *Individual* expectancy chart shows the probability that a person will achieve a particular level of performance (Figure 8.11)
  - *Institutional* expectancy chart indicates what will happen within an organization if all applicants above a minimum interview score are hired (Figure 8.12)

## Factors Affecting the Magnitude of Validity Coefficients

- Reliability of criterion and predictor
- restriction of range
- Criterion contamination
- Violation of statistical assumptions—Figure 8.13

## Utility Analysis

### *A Definition of Utility Analysis*

- The goal is to translate the results of a validation study into terms that are important to and understandable by managers
- Using dollars-and-cents terms as well as other measures, such as percentage increases in output, utility analysis shows the degree to which use of a selection measure improves the quality of individuals selected versus what would have happened had the measure not been used

### *Some Preliminary Work on Utility Analysis*

1. *Validity coefficient*—the magnitude of the correlation of the selection procedure with a criterion
2. *Selection ratio*—the ratio of the number of people to be hired to the number of applicants available
3. *Base rate*—the percentage of employees currently successful on the job using the selection procedure

### *Applying Utility Analysis to Selection: Some Examples*

1. Costing the value of selection procedure
2. Enhancing recruitment
3. Using a method with low validity

## BROADER PERSPECTIVES OF VALIDITY

[slides 62–70]

## Validity Generalization

### *An Overview*

- The idiosyncrasies of jobs, organizations, and other unknown factors contributed to the differences in validity results that were obtained
- Wide variations in the magnitudes of validity coefficients across validation studies, even when the same test had been used
- Validity is generalizable across situations

### *Validity Generalization Methods*

1. Obtain a large number of published and unpublished validation studies for a selection procedure
2. Compute the average validity coefficient for these studies
3. Calculate the variance of differences reported
4. Subtract the variance due to the effects of small sample size
5. Correct the average validity coefficient and variance for errors due to other methodological deficiencies
6. Compare the corrected variance to the average validity coefficient to determine the variation in study results
7. If the differences are small, then differences are concluded to be due to methodological deficiencies and not to the nature of the situation

### *Conclusions from Validity Generalization Studies*

- Schmidt and Hunter concluded it is not necessary to conduct validity studies within each organization for every job
- Mental ability tests can be expected to predict work performance in most employment situations

### *Criticisms of Validity Generalization*

- The correction formula used to generalize results usually not based on sufficient data but on hypothetical values derived from other research work assumed appropriate
- Correction formulas may be inappropriate and may overestimate the amount of variance attributable to study deficiencies

### *Validity Generalization Requirements*

1. The user must be able to show that the proposed selection procedures assesses the same WRCs or that it is a representative example of the measure used in the study database
2. The user must be able to show that the job in the new employment setting is similar to the jobs or group of jobs included in the study database

## **Job Component/Synthetic Validity**

### *An Overview*

- There are a number of different approaches to job component validity—*synthetic validity*
- Involves demonstrating that a correlation exists between a selection procedure and at least one specific aspect or component of a job
- Once established, it is assumed that the selection procedure is valid for predicting performance on that job component if it exists on other jobs

### *Conducting a Job Component Validity Study*

1. Conduct an analysis of the job using the position analysis questionnaire (PAQ)
2. Identify the major components of work required on the job
3. Identify the attributes required to perform the major components of the job
4. Choose tests that measure the most important attributes identified from the PAQ

### *Accuracy of Job Component Validity Studies*

- Job component validity estimates generally lower and more conservative than validity coefficients obtained in actual validation studies
- The Department of Labor's O\*NET database consists of job analysis information collected on a large array of occupations—information on 42 generalized work activities that occupations may involve is available
  - Use of the O\*NET for job component validation is in its infancy
  - More developmental research is needed

### *Criticisms of Job Component Validity Strategy*

- Mossholder and Arvey have noted that the method has been less successful in predicting actual validity coefficients
- The strategy has been relatively less useful in predicting psychomotor test data
- The strategy generally has reported test results from the General Aptitude Test Battery—available only to public employers

## **VALIDATION OPTIONS FOR SMALL SAMPLE SIZES**

[slides 71–72]

- Content validity
- Validity generalization
- Job component validity or some other form of *synthetic validity*—a logical process of inferring test validity for components of jobs (Figure 8.16 illustrates three jobs and work performance components common to each)

## **THE FUTURE OF VALIDATION RESEARCH**

[slide 73]

- Changes occurring in the workplace:
  1. Increasing numbers of small organizations without the resources, technical requirements, and technical skills available to undertake traditional validation strategies
  2. Increasing use of teams of workers rather than individuals
  3. Changes in the definitions of job success to include such criteria as organization and job commitment, teamwork, and quality of service delivered to customers
  4. The changing nature of work—jobs and requirements for performing them are becoming more fluid, requiring job analytic methods that focus on broader work capacities rather than on molecular task requirements

## CHAPTER 9

# *Application Forms – Biodata Assessments, Training & Experience Evaluations, and Reference & Social Media Checks*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand the design and use of employment application forms and how biodata application material is useful in predicting job-relevant criteria.
2. Review employment application forms for content that might be illegal.
3. Develop rudimentary application form items relevant to a position opening.
4. Screen applicant résumés for content that might signal distortion of an applicant's reported background and accomplishments.
5. Know when and when not to use training and experience evaluations with job applicants.
6. Understand the role of applicant reference checks and why use of social media content for human resource selection purposes is currently not appropriate.

### APPLICATION FORMS – BIODATA ASSESSMENTS

[slides 4–8]

#### Nature and Role of Application Forms and Biodata in Selection

- An application form consists of a series of questions designed to provide information about the general suitability of applicants for jobs to which they are applying
- It serves as a means for
  - a. deciding whether applicants meet the minimum requirements of a position
  - b. assessing and comparing the relative strengths and weaknesses of individuals making application
- An application form, however, can provide information that unfairly discriminates against some individuals—federal and state laws (such as Title VII) were passed to prevent discrimination
- A key issue facing any application reviewer is in deciding *what* application data are most beneficial in choosing successful job applicants
- We refer to application information empirically developed and scored in a way to maximize prediction as *biodata*—self-report data provided by the applicant

- Biodata is based on the notion that a deeper study of applicants and their employment backgrounds and life experiences can be employed as an effective predictor
- At least two ways to categorize biodata:
  - a. *Response-type*—based on the type of response options offered to a respondent (Table 9.1—examples)
  - b. *Behavior-type*—based on the specific behavioral content of the item itself (Table 9.2—examples)

## EVALUATING APPLICATION FORMS AS PREDICTORS

[slides 9–32]

- Information obtained from application blanks can be useful, but to realize its potential, the form must either:
  - rely on an empirical analysis to identify items that predict the outcome of interest, or
  - begin with a job analysis to ensure that only job-related questions are posed
- Biodata can predict supervisory ratings of performance, productivity, promotion, tenure, turnover, and training success (Table 9.3) but only when the criterion is taken into account when identifying which items to rely on
- Self-reported application data are susceptible to distortion—applicants believe it is advantageous to “look good”
- Applicants should be told—preferably both verbally and in writing—that the information they give will affect their employability
- Extent of faking can be reduced through instructions that include statements about the consequences (e.g., “Deliberate attempts to falsify information ... may be grounds for not hiring you ... or for terminating you after you begin work.”)
- Applicants should be required to sign and date their application, and sign a statement certifying the accuracy of the information they provided
- In states covered by an employment-at-will doctrine, an employer should be sure that no contract of permanent employment is implied in the application or any resulting job offer letter

## Legal Implications of Application Forms

- “The more information (on application forms), the better” mentality can create major problems for an employer
  - Federal and state laws affect the kinds of information that can be requested
  - Generally assumed that *all* questions asked are used in making hiring decisions
  - The burden of proof may be on the employer to demonstrate that *all* application questions are fair and not discriminatory
- The law according to EEOC preemployment guidelines cautions against questions that
  - a. disproportionately screen out minority group members or members of one sex
  - b. do not predict successful performance on the job
  - c. cannot be justified as a business necessity
- Table 9.4 lists rating criteria useful for examining the appropriateness of application form questions

## Composition of Application Forms

- Most job applications consist of two major parts:
  - Instructional information for completing and submitting the application
  - Questions whose answers are used in deciding applicant suitability for a position in the organization
- Instructions for applicants:
  - The form should be clear and understandable
  - The form should state that applicants giving unsolicited information will be rejected
  - The form should state that disabled applicants can request reasonable accommodation in completing the application
  - The form should state that it has been reviewed to ensure it is job-related and fair
  - The form should address privacy concerns – who will see the information or how applicant responses will be used
  - The form should be reviewed for its attractiveness, fairness, ease of use
- Questions for applicants:
  - Requesting information other than that necessary for initially judging applicant qualifications to perform a job opens an organization to the possibility of a discrimination charge
  - Organizations should ask job-related questions only and avoid those relating to personal information
  - Including discriminatory questions on an application form influences applicants' perceptions of the employing organization—applicants viewed the company as less attractive, were less motivated to pursue employment with the organization, were less likely to accept an offer of employment, were less likely to recommend the organization

## Selecting Application Form Content

- Some essential data should be assessed in all forms:
  - a. Name
  - b. Current address
  - c. Telephone number
  - d. Work experience
  - e. Level of education and training received
  - f. Skills possessed
  - g. Social Security number
- The lower the level of a job or job class, the shorter, less detailed the content of the application
- Job analysis methods can identify items that could be useful in screening applicants for a job but each must be reviewed for its fairness and usefulness
  - Employers should first research the fair employment practice laws that exist for their state
  - An employer conducting business in more than one state should review each state's laws, regulations, and guidelines concerning use of preemployment inquiries—one excellent source for review is *The Commerce Clearing House Employment Practice Guide*, State FEP laws

## Developing and Revising Application Forms

- Employers should study carefully the development or revision of their biodata forms
- Several steps involved in developing biodata questions, including:
  1. Because there are often many type of jobs in a firm, more than one application form may be needed
  2. Job analysis data should serve as one basis for choosing employment application questions
  3. Develop a pool of biodata items
- These guidelines should be followed as biodata items are formatted into an application form:
  1. Should principally deal with past behavior and experiences
  2. Items dealing with family relationships or other personal matters are usually viewed as offensive
  3. Specificity and brevity of of items and response options are desirable
  4. Numbers should be used to define options or alternatives
  5. All possible response options and an “escape” option (“other”) should be given—where possible, response options that form a continuum should be used
  6. Item options should carry a neutral or pleasant connotation
  7. Items dealing with past and present behaviors and with opinions, attitudes, and values are generally acceptable
  8. Items should reflect historical events that are important in shaping a person’s behavior and identity
  9. To lessen the effect of individuals responding in ways considered to be socially desirable (faking), biodata items should reflect external events, be limited to firsthand recollections, be potentially verifiable, and measure unique, discrete events (such as age when first licensed to drive)
- Prescreening and pilot-testing biodata items:
  - Biodata items developed are reviewed by a panel of judges—items that objectionable or have potential bias are eliminated
  - Ideally, items will be tested on a large group representative of the applicant or employee population for which the biodata form will be used
  - Analyses are performed in order to choose those items most useful
  - Items passing prescreening and pilot-testing reviews are retained for inclusion in the final version of the application form
- Scoring the biodata form:
  - Various scoring options for biodata measures usually fall into two categories:
    - a. Calculation of a *single*, empirically keyed overall score predictive of employee success
    - b. Development of multiple scores for dimensions or groups of related items appearing on a biodata inventory
  - Several methods of empirical keying are available including vertical percentage, horizontal percentage, correlation, differential regression, deviant response, rare response

## Résumés

- The first impression many employers have of applicants often is from their résumés
- Résumés are subject to the same kinds of distortions that plague many application forms
- Résumé items used by recruiters to infer job applicant attributes are shown in Table 7.6

## Internet-Based Résumé Screening

- The rate of unsolicited résumés has exploded along with growth in Internet usage
- Companies frequently automate scanning and scoring résumés using keyword search to identify specific attributes that meet minimum qualifications or credentials—work experience, training or education
- Two legal issues particularly salient with online screening
  1. Disparate impact due to lack of Internet access
  2. Privacy concerns of applicants

## Using Application Forms in HR Selection

- Research shows that application information managers rely to make the initial screening decision often differs from they actually use—e.g., some relied on gender rather than applicant qualifications for the job
- Methods that have been demonstrated to provide valid information tend to be either based on an empirical scoring key, resulting in an overall score, or they rely on multiple job-relevant constructs with separate scores
- Application forms can be used liked a checklist—training and experience evaluations (T&E evaluations)

## TRAINING AND EXPERIENCE (T&E) EVALUATIONS

[slides 33–40]

### Nature and Role of T&E Evaluations in Selection

- T&E evaluations are a way to rationally assess previous experience, training, and education information given by job applicants
- Scores from the evaluations can be used in a number of ways:
  - a. as the sole basis for deciding whether an an individual is qualified
  - b. as means for rank-ordering individuals from high to low based on on a T&E score
  - c. as a basis for prescreening applicants prior to administering more expensive, time-consuming predictors
  - d. in combination with other predictors used for making an employment decision
- Examples of T&E evaluations in selection
  - Only a brief check needed of relevant portions of a job application for minimal qualifications—Figure 9.1
  - When a more thorough review of minimum qualifications is being made—Figure 9.2
  - Determination of what experience, education, and training and relevant for successful task performance—Figure 9.3
  - T&E calculations are based on ratings rather than using empirically keyed life history responses like those used in biodata

## Reliability and Validity of T&E Evaluations

- *Reliability*—T&E evaluations tend to reflect rather high interrater reliability estimates
- *Validity*—although most organizations use prior training and work experience as a first cut in selecting applications, surprisingly few studies examine the validity of these predictors
  - One study found that the validity of T&E ratings varied with the type of procedure used
  - Behavioral consistency method demonstrated the highest validity with a mean correct validity coefficient of 0.45

## Legal Implications

- Organizations should ensure that training and experience qualifications they rely on for initial selection decisions are fair and not discriminatory
- Organizations should avoid collecting information that disproportionately screens out members of one sex or minority group, particularly when that information
  - a. does not predict successful performance on the job
  - b. is not related to the requirements of the job
  - c. cannot be justified as a business necessity
- Predictors should be based on a competent job analysis, have some validity evidence available, and be uniformly applied to all applicants

## Recommendations for Using T&E Evaluations

1. Use T&E evaluations to set specific minimum qualifications job candidates should hold rather than using a selection standard such as a high school diploma
2. T&E evaluations are subject to the *Uniform Guidelines*
3. T&E evaluations should be used only as rough screening procedures for positions where previous experience and training are necessary for job performance
4. Forms and procedures for collecting and scoring T&E evaluations should be standardized as much as possible
5. Some form of data verification should be made from self-report data
6. Where distortion of self-evaluation information is likely to be a problem, final hiring decisions based on other selection measures (ability, job knowledge, tests) can minimize risks associated with T&E evaluations

## REFERENCE CHECKS

[slides 41–62]

### Nature and Role of Reference Checks in Selection

- Collection of information about prospective job applicants from people who have had contact with the applicants—information used for the following purposes:
  - a. To *verify* information given by job applicants
  - b. To serve as a basis for either predicting job success or screening out unqualified applicants
  - c. To *uncover* background information that may not have been provided by applicants

## Types of Reference Data Collected

- Generally speaking, four types of information are solicited through reference checks:
  - a. Employment and educational background data
  - b. Appraisal of an applicant's character and personality
  - c. Estimates of an applicant's job performance capabilities
  - d. Willingness of the reference to rehire an applicant
- Table 9.7 illustrates the usefulness of employment and personal background information collected through reference checks

## Usefulness of Reference Data

### *Reliability and Validity of Reference Data*

- little research evidence available regarding reliability and effectiveness in predicting job performance
- low validity of reference data
- multiple factors affect validity

### *Applicant Acceptability*

- applicant reactions to references likely to vary
- references viewed more favorably by applicants than cognitive ability evaluations, personality tests, biodata, integrity tests, but not as favorably as employment interviews, work samples, résumés

## Legal Issues Affecting the Use of Reference Checks

- Two broad categories particularly critical:
  - Discriminatory impact reference checks may have on a job application and the defamation of an applicant's character through libel or slander
  - Complaints filed against employers for “negligent hiring” of employees

### *Discriminatory Impact and Defamation of Character*

- For defamation to occur, several elements must be present:
  1. A written or oral defamatory statement must have been given
  2. There must be a false statement of fact
  3. Injury must have occurred to the referee
  4. The employer does *not* have absolute or qualified privilege

### *Negligent Hiring*

- For an employer to be held liable, five points must be covered:
  1. Injury to a third party caused by an individual employed by a firm
  2. Employee must be shown to be unfit for the job he or she holds
  3. Employer knew or should have known that the employee was unfit for the job
  4. Injury received by the third party must have been a foreseeable outcome resulting from the hiring of the unit employee
  5. Injury is a reasonable and probable outcome of what the employer did or did not do in hiring the individual

- Recommendations to former employers to supply information on a past employee
  - Submit a written request for information on specific questions relevant to making an employment decision, and use the same questions with every applicant
  - Include a release form signed by the applicant stating the applicant has read and approves the information requested and the applicant requests that the information be given
- If the previous employer refuses to provide such information, then
  - Call the person in charge of human resources and ask why the request was not honored
  - Ask how a request should be made so it will be honored
- If the previous employer still refuses, then
  - Inform the individual that failure to cooperate is being documented with date, time, and name of the person refusing
  - Verify the candidate's statements about the position held, number of years in that position, and the final salary
  - If the missing information relevant that the applicant will not be hired without it, indicate that previous employer's refusal the reason for lack of an offer

### Methods of Collecting Reference Data

- *Telephone checks* – relatively fast, high reference return rate, allows follow-up questions or clarity of type of information needed, inexpensive to conduct
- *Internet and e-mail reference checks*—fast, inexpensive, but not as interactive, less desirable due to lack of privacy and informality, may be necessary when searching public records, references submitted electronically considered legally binding
- *Mail checks*—a written questionnaire or letter, candidate should sign a release form giving former employers permission to release information without liability, low return rate
- *Letters of recommendation*—probably restricted to high-skill or professional jobs, may provide greater depth of information, negative comments seldom given, disadvantages include:
  1. Writers have the difficult task of organizing the letter and deciding what to include
  2. Letter quality will depend on the effort expended by the writers and their ability to express their thoughts
  3. Writers tend to be positive in their evaluations and often lack specificity and accuracy in letter writing
  4. The same job-relevant information will not be obtained on each applicant
  5. Information relevant to hiring organization may be omitted
  6. Scoring of the letter is subjective and based on reader's interpretation
- *In-person checks*—involve face-to-face personal contact with reference giver (allows high level of interaction which may lead to more useful information being exchanged), often part of background investigations and concern jobs in which an incumbent is a potential security of financial risk, expensive, time-consuming, often impractical, not frequently used in most selection programs

## Sources of Reference Data

- Reference givers must meet four conditions for their data to be useful:
  - a. They must have had a chance to observe the candidate in relevant situations
  - b. They must have knowledge of the candidate
  - c. They must be competent to make the evaluations requested and be able to express themselves so their comments are understood as intended
  - d. They must want to give frank and honest assessments
- *Former employers*—an important source for verifying previous employment records and evaluating an applicant's previous work habits and performance, available data likely released through personnel office, information from previous supervisors particularly valuable
- *Personal references*—most applicants choose individuals they believe will give a positive evaluation, can provide information about prior employment and candidate's qualities and behavior characteristics, important to ask how long and in what capacity the reference has known the applicant
- *Investigative agencies*—will conduct background checks that focus on résumé and application information, educational accomplishments, credit ratings, police and driving records, personal reputation, lifestyle; expensive and require more time; checks take the form of consumer reports (two basic types):
  1. Consumer reports—any written or oral communication collected concerning an individual's credit standing, character, general reputation, personal characteristics used to establish eligibility for employment
  2. Investigative consumer reports—based on personal interviews with friends, neighbors, acquaintances
- *Public records*—use caution—be sure information solicited does not discriminate and can be justified by the nature of the job for which applicant is being screened; records include:
  1. Criminal records
  2. Motor vehicle department records
  3. Workers' compensation records
  4. Federal court records
  5. Educational records

## Social Media

- Many employers are using social media—Facebook, LinkedIn, Twitter—to conduct their own informal “reference checks”
- Information learned just as often used to screen out candidates as to hire candidates
- Can serve as an important predictor of job performance and organizational fit – current empirical evidence does not support that use
- Information solicited should not discriminate against a protected group
- Until reliability and validity evidence for collecting social media information available, use of same should be avoided in HR selection

## Recommendations for Using Reference Checks

1. Reference data most properly used when involves job-related concerns
2. Because reference check tailored to a specific job, likely need more than one general form for all positions in an organization
3. Reference checks subject to the *Uniform Guidelines*
4. A more structured reference checking system rather than an unstructured system less likely to be open to charges of discrimination
5. Applicants should be asked to give written permission to contact their references
6. Reference takers collecting information should be trained in how to interview reference givers
7. All reference check information should be recorded in writing
8. If applicant provides references but reference information cannot be obtained, ask for additional references
9. Check all application form and résumé information
10. Negative information received frequently serves as a basis for rejecting an applicant—caution is advised in using *any* negative data as a basis for excluding applicants
11. If background investigation firm used, understand that firm is serving as the employer's agent – employer may be held liable for agent's actions
12. Do not ask for applicants' social media usernames and passwords

## SUMMARY

[slide 63]

- Application forms, resumes, reference checks, and ratings of training, education, and experience are relatively inexpensive and easy to collect
- Collecting basic background information enables one to ensure applicant has the minimal skills and qualifications to do the job
- Those making hiring decisions should carefully consider each piece of information requested in the early stages of selection – considerable utility to be gained by rigorously measuring minimum qualifications and applicant ability and motivation at this stage

## CHAPTER 10

# *The Selection Interview*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand the ways a structured interview differs from an unstructured one.
2. Know the steps required to design an effective interview.
3. Realize the interview involves competing objectives: for the interviewer, the objectives include generating interest and maintaining the applicant's interest in the company, hiring the best candidate, and increasing the likelihood the applicant will accept an employment offer; whereas for the applicant, the objectives include obtaining a job offer and finding out whether he or she "fits" the job, the potential coworkers, and the supervisor, as well as the organization itself.
4. Understand which WRCs the interview is best suited to measure, because it is a unique social exchange.
5. Identify different interview types: screening versus selection interviews and behavioral description versus situational structured interviews.

### THE SELECTION INTERVIEW

[slides 4–6]

#### ■ Two approaches:

- Option A—conduct a standardized interview following a few “key” rules—*structured* interview
  - Predictive validities range from 0.44 to 0.62
- Option B—conduct the interview as the hiring manager thinks best – *unstructured* interview
  - Predictive validities range from 0.20 to 0.33

#### ■ Structured interview—a few key rules:

1. Ask the same questions for all candidates
2. Use only job-related questions capable of differentiating high from low performers
3. Have an established “scoring key” complete with examples of a 3- to 5-point rating scale
4. Mechanically sum the ratings over all questions to establish an overall score
5. Rely on multiple well-trained interviewers rating each candidate independently from other interviewers

- Unstructured interview—preferred by many:
  - Flexible—liked by both applicants and interviewers
    - Candidates like unstructured interviews because they have more discretion in answering these questions – enables applicant to better convey their own skills and abilities, applicant has the chance to learn more about the job opportunity itself (what the job, work group and manager, organization are like)
    - Interviewers like this format because it allows them flexibility to get to “know the person”, provides an opportunity to win the applicant over and portray the best side of their job, products or services, the organization

## BRIEF HISTORY

[slides 7–23]

- Early studies show unstructured interviews suffer from both low reliability and low validity due to a combination of both the use of inappropriate questions and reliance on idiosyncratic, extraneous factors that affected an interviewer's evaluation of an applicant
- After 25 years of research, validity generalization studies have produced strong evidence of the predictive validity of structured interviews
- Schmidt and Hunter concluded that the corrected validity coefficient for the structured interview (0.51) is comparable to similar coefficients produced for cognitive ability tests (0.51) and higher than for assessment centers (0.37)
- Unstructured interview better at recruiting
- How to combine both approaches?

## Developing and Designing Effective Interviews

- Range of outcomes possible when using the interview fall into five main categories:
  - a. Provides an opportunity to select the best set of candidates from the applicant pool
  - b. Is an opportunity to recruit highly qualified candidates and sell them on the organization, product or service, job
  - c. Enables applicant to assess their fit with the organization/job as well as ask about procedural aspects linked to job decision
  - d. Is an efficient and practical method for measuring a number of applicant WRCs
  - e. Can be used to make either an early decision or a later one about an applicant's acceptability

*Recruiting the Applicant to the Organization: Providing Job Information*

- Oral discussions between two individuals being recalled differently by each person—give a written job description to applicants that supplements much of the information transmitted during the interview
- Design the employment interview using a mix of selection and recruiting—increases the applicant's willingness to accept an offer and allows the employer to assess the candidate's qualifications

*Effect on Attracting Applicants: Recruiting Outcomes*

- The interview is only part of the information that applicants use to form impressions about the organization and to make job search and acceptance decisions

- Recruiter characteristics and behavior are part of this information—those with the “right” qualities and trained to consistently and fairly provide correct information have a significant impact on candidate attitudes towards a job, demographic characteristics have no impact
- Nevertheless, the effect of recruiter behavior becomes a relatively modest effect

### Making a Selection Decision

- An interviewer's rating of candidate suitability can be affected by physical attractiveness and professional demeanor, use of impression management behavior, and verbal and nonverbal behavior of the candidate—need to ensure interviewer ratings are focused on job-relevant WRCs
- Some biases most relevant to the interview:
  - *Anchoring*—locking onto salient information too early
  - *Confirmation bias*—tendency to look for confirming evidence rather than for evidence to refute an early judgment
  - *Illusory correlations*—perceiving associations where none exist
  - *Overconfidence bias*—tendency to act on incomplete information or hunches

#### *Screening versus Selection Interview*

- Screening interview—used to assess applicants on general characteristics
  - Focuses on the checking of credentials and licensure requirements and evaluation of an applicant's minimum work requirements needed for the job, involves more recruiting information
- Selection interview—used to assess more job-related WRCs
  - Composed of questions concerning job-related knowledge, interpersonal skills, problem-solving skills, other work-related experiences and behaviors
- Recommendations for developing questions for a screening interview:
  1. Use job analysis information to identify general or fundamental WRCs that an applicant must possess and for the organization does not provide training for
  2. Use “job experts” to identify the most important of these characteristics
  3. Use a modified crucial-incidents technique to identify questions

#### *Measuring Applicant WRCs*

- More is not better—focus on only a few key WRCs and assess those thoroughly
- Appropriate WRCs for the interview—interview is designed to evaluate seven major dimensions (Table 10.1), but these predictors not highly correlated with the single-best predictor of performance (that is, measures of general mental ability)

#### *Type of Structured Interviews*

- Structured versus unstructured interviews
  - Unstructured results in subjective, global evaluations that are not very useful
  - Structured rely on a disciplined method for collecting job-relevant information (Table 10.2 shows components of the structure interview that interviewers do not use very often)

### *Developing Appropriate Interview Questions*

- Evaluation of candidates should be more accurate if more information is obtained about the candidate's job qualifications
- Two types of structured interviews for developing specific job-related questions:
  - The situational interview
  - The behavior description interview
- *The Situational Interview*
  1. Do a job analysis of the position using the critical-incidents technique
  2. Incidents are then sorted into groups of similar behaviors – *behavioral dimensions*
  3. Review incidents for each behavioral dimension, select a small number of the most appropriate incidents, and use these to write interview questions
  4. Score applicant responses using a five-point scale – see Table 10.3 for examples
- *The Behavior Description Interview*
  1. Do a job analysis of the position using the critical-incidents technique
  2. Incidents are then sorted into groups of similar behaviors – *behavioral dimensions*
  3. Review incidents for each behavioral dimension, select a small number of the most appropriate incidents, and use these to write interview questions
  4. Create follow-up questions for each question
  5. Score applicant responses using a five-point scale – see Table 10.4 for examples

### **Conclusions about Designing the Interview**

- Number one purpose should remain on accurately assessing a few construct valid WRCs, but should also try to accommodate other objectives. Suggestions:
  1. Train interviewers about rapport building and consider scoring based on job relevance
  2. Before transitioning into the structured interview, train the interviewer on a script that tells the applicant what to expect during the remainder of the interview
  3. If this is acceptable to the applicant, begin the structured interview
  4. When all questions have been asked and all answers have been rated, the structured interview ends, notes are put away—remainder of the time dedicated to answering questions applicant may have and providing information on the position and organization

### **THE ROLE OF TECHNOLOGY AND GLOBAL TRENDS**

[slide 24]

- With the spread of Skype and videoconferencing, interviews increasingly are occurring at a distance
  - Preliminary evidence suggests comparable predictive validities to more traditional face-to-face interviews
  - Future research will explore whether the impact of impression management behaviors changes the amount or quality of information gleaned from an interview
  - Cultural differences may influence the rate and interpretation of impression management behaviors in the interview
  - Reactions to structured interviews may differ and might even be illegal in different countries

## EVALUATING INTERVIEWS AS PREDICTORS

[slides 25–27]

### Predictive Validity

- Structured interviews are more valid than unstructured interviews
  - Behavior description interviews are past-oriented questions involving prior work experiences
  - Situational interviews are future-oriented questions that ask applicants to imagine a work situation
  - Even when designed to assess the same construct, the two types of interviews seem to reflect different measures
  - Behavioral interviews have somewhat higher predictive validities than do situational interviews
- Situational interviews—much of the validity evidence produced from studies involving entry-level, clerical, hourly jobs; jobs for which applicants have no prior work experience
- Behavioral interviews – questions related more to what the candidate has done in actual situations; higher predictive validity for more complex jobs
- Interviews should include both situational and behavioral interview questions—a combination of these questions will predict better than either alone

### Reliability

- Why structured interviews work better than unstructured interviews relates to differences in reliability
- By improving the structure or standardization of the interview, one increases reliability and higher reliability leads to higher validity
- To increase reliability is to increase the number of interviews given to each job candidate—increases the amount of information on which a hiring recommendation can be based, should correct for idiosyncratic biases of individual interviewers, averaging the ratings of three or four independent interviewers resulted in the same level of predictive validity

## DISCRIMINATION AND THE INTERVIEW

[slides 28–30]

- Discrimination could occur if
  - a. decision of the selection interview led, or assisted in leading, to disparate treatment or a pattern of disparate impact
  - b. the interview could not be defended by showing job relatedness
- Employers should avoid asking questions relating to race, color, religion, sex, national origin, marital status, sexual orientation, age, handicap or disability, status as a Vietnam-era or disabled veteran

### Court Cases

- *Watson v. Ft. Worth Bank & Trust* case has had a major impact on how the selection interview is treated by courts and companies in discrimination cases
  - Because unscored interviews considered subjective selection devices, cases most often heard as disparate treatment issues—easier to defend than a disparate impact charge

- The *Watson* decision stated that a case in which the selection interview is central to the charge of discrimination could be heard as a disparate case if the appropriate data were presented—many companies now use a scored interview
- Table 10.5 shows a summary of other cases that have focused on the specific practices used by companies to conduct the selection interview

## RESEARCH FINDINGS ON DISCRIMINATION

[slide 31]

- Legal restrictions against using race, ethnicity, sex, age, and disability in employment decisions should minimize or eliminate the influence of these characteristics when making hiring decisions
- Direct effects resulting from candidate race, sex, age, and other demographic characteristics on interviewer ratings generally are small and inconsistent, particularly in structured interviews
- Applicant disclosure of nonobvious disabilities increases the likelihood of a hiring recommendation
- Black and Hispanic candidates received evaluations that, on average, were about one-quarter of a standard deviation lower than those for white candidates

## A MODEL OF INTERVIEWER DECISION MAKING

[slides 32–39]

- Models that focus on factors that affect decision making during the interview provide a way to organize the conduct of research and the interpretation of results
- Such models assume the interviewer and candidate are gathering and processing information about each other as well as the organization and job
- Dipboye's model (shown in Figure 10.1) briefly summarizes the existing large body of research that identifies critical factors affecting the decision-making process over the different phases of the interview—before, during, and after the interview

## Expectations, Beliefs, Needs, and Intentions Before the Interview

- Interviewers and candidates bring their own expectations, beliefs, needs, and intentions with regard to the job, interview, organization, and each other—these influence all three stages of the interview
  - Preinterview stage—factors have a significant impact on impressions held by the interviewer and candidate which in turn influence interviewer evaluations of those candidates, interviewers seek out and recall, even distort, information that supports or confirms their preinterview impressions
- Interviewer's evaluation of the candidate should not be influenced (or biased) by examining applicant information before the interview – access to the Internet has compounded this problem (e.g., applicant's Facebook page)
  - A lab experiment that found hypothetical job candidates posting negative work-related activity or alcohol abuse and drug use or sexually inappropriate pictures or the use of profanity were less likely to be recommended by the recruiter
- Candidates who thought the interviewer was likely to extend them a job offer after the first few minutes of the interview were less vigilant and motivated to effectively manage their image during the remainder of the interview

### *Individual Differences*

- Interviewer evaluations and hiring decisions have been shown to be influenced by:
  - A candidate's physical attractiveness
  - The candidate's handshake at the start of the interview
  - The candidate's personality—a direct effect during and after the interview
  - Whether a candidate has participated in a coaching program before the interview
- Candidate demographic characteristics have small and inconsistent effects on actual hiring decisions

### *Social Interactions During the Interview*

- Candidate's incentive to manage one's image
- Attitudinal similarity leads to more favorable ratings
- Interviewers who are more personable, competent, informative increased candidate's attractiveness to the job or firm
- Ratings tend to be higher when the interview is longer and the interviewer talks more during the interview
- Ratings had greater accuracy when the interviewers had a high level of verbal cues and moderate to high levels of nonverbal cues
- Candidate's impression management can create positive or negative images of the applicant—ingratiation and self-promotion positively affect interviewer evaluations
- Interviewers often disagree about which topics should be covered in an interview—these differences relate more to individual preferences than to the perceived importance of job requirements

### *Information Processing and Decision-Making Factors*

- Decision makers use a two-stage process
  1. Decision maker tries to reduce choices by screening out unsuitable candidates—interviewers categorize applicants
  2. Interviewers evaluate choices more rigorously in an attempt to choose the “best” candidate—characterization
- Initial information carries considerable weight in the interview—first impression establishes an anchor for the interviewer
- Interviewers give more weight to negative information than to positive information in decision making—unfavorable ratings on only one of several characteristics resulted in rejection in more than 90% of cases
- Ratings not accurate due to the amount of information an interviewer can recall after an interview—notes should be taken

## **RECOMMENDATIONS FOR USING THE INTERVIEW**

[slides 40–54]

### **Restrict the Scope of the Interview**

- One of the major weaknesses of the interview is that it often is used to accomplish too many purposes – recruitment and selection should be separated systematically

- Too many WRCs are evaluated in the interview—limit the scope of the interview to a much narrower band of applicant characteristics
- Table 10.7 shows an example of a job analysis that demonstrates the appropriate use of the interview in a selection program using various WRCs

### Limit the Use of Preinterview Data

- Assuming the interview focuses on a small number of WRCs, limit interviewers to two types of preinterview information
  1. Complete data about any of the WRCs to be covered in the interview
  2. Incomplete or contradictory statements reported on the application blank or other similar instruments
- Yet, access to data not directly relevant to the purposes of the interview only contributes to deficiencies in interviewer decisions

#### *Adopt a Structured Format*

- A set of questions should be formulated for each WRC identified as appropriate for the interview – these should be asked of each applicant
- Interviewers may go beyond these questions as they feel necessary—either to clarify a given response, seek other important details, or pursue a closely related area (this approach builds consistency which makes comparisons among applicants much easier)

#### *Use Job-Related Questions*

- Information gathered must be useful in measuring characteristics required in the job—three types of questions:
  1. Questions of job knowledge
  2. Questions of social interaction
  3. Questions of personality or habitual behaviors
- Questions that could be used in a selected interview for the job of maintenance supervisor are presented in Table 10.8 (from WRCs previously identified in Table 10.7)
- Use multiple questions for each WRC
  - The more items an assessment device possesses that measure the same WRC, the greater its reliability and predictive validity
  - The number of questions will depend on the information developed from the job analysis and the time available for each applicant
  - Consider first the relative importance of the WRCs being measured in the interview
- Rely on multiple independent interviewers
  - The use of multiple interviewers will enhance reliability if their ratings are independent of the influence of other interviewers (or managers)
  - Predictive validities for panel interviews are comparable to or somewhat lower than the correlations reported for single interviewers
  - Pooling or averaging interviewer judgments reduces the influence of measurement error, particularly if from different interviews of the same candidate
  - The same interviewers should be used across all candidates

- Apply a formal scoring format
  - An interview format that provides a formal, defined scoring system is superior in many ways to a format that does not—enhances legal defensibility as well as the reliability and validity of interviewer judgments
  - Interval measurement scales the most commonly used systems
  - Table 10.9 contains an example of a scoring form that can be used in interviews for the position of maintenance supervisor
- Train the interviewer
  - Training should be provided to interviewers regardless of whether the training is structured or unstructured
  - Focal skills of an interviewer are the abilities to
    - a. accurately receive information
    - b. critically evaluate the information received
    - c. regulate his or her own behavior in the delivery of questions
- Receiving information
  - Instruction has concentrated on factors that influence
    - a. hearing what the respondent has said
    - b. observing the applicant's behavior
    - c. remembering the information received
- Evaluating information
  - Focuses on improving the decision-making process by pointing out common decision errors and providing methods to overcome these errors
- Interviewing behavior
  - The interview is fundamentally a social and interpersonal process—the behavior of the interviewer can affect the candidate's responses
- Results of training
  - Training programs, for the most part, have reduced some of the more common rater errors
  - Training has been found to enhance the reliability of interviewer judgments
- Account for social dynamics and multiple purposes
  - The interview is based, in part, on a social exchange between two people with conflicting objectives
  - The organization should explicitly address the multiple purposes it has to deal with in a systematic way
  - Structured interviews can provide some of the most valuable predictor information about the candidate while also being viewed favorably by the applicant

## CHAPTER 11

# *Ability Tests for Selection*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand what is meant by the term ability tests and what work-related characteristics (WRCs) they can measure.
2. Know what cognitive ability tests measure and how they can be used in selection.
3. Know the broad range of jobs for which cognitive ability tests are valid selection measures.
4. Understand that cognitive ability tests demonstrate adverse impact in selection but are also valid.
5. Understand the issues involved with using cognitive ability tests in unproctored testing situations that include computers and other electronic devices.
6. Know what mechanical ability tests measure and how they can be used in selection.
7. Know what clerical ability tests measure and how they can be used in selection.
8. Know what physical ability tests measure and how they can be used in selection.
9. Know what information should be reviewed before using or purchasing an ability test for selection and how to evaluate that information.

### HISTORY OF ABILITY TESTS IN SELECTION

[slides 4–7]

- In 1908, tests were developed for use in selection of street car operators for the Paris Transportation Society—measures included reaction time, ability to estimate speed and distance, ability to choose correct driving behavior in reaction to street incidents
- World War I became a major impetus in the development of other tests used in selection—a committee was formed, they decided that the development and use of tests was the greatest contribution that psychology could offer to military efficiency
  - Developed paper-and-pencil tests that would provide scores to be used as a basis for rejecting recruits thought to be unfit for military service
  - First test developed was a mental ability or intelligence test with objective scoring methods—format that could be rapidly scored, alternate forms, required a minimum of writing, structure that permitted economical use of time
  - These same requirements have characterized industrial ability tests ever since
- The next two decades saw the development of mechanical, motor, clerical, and spatial relations ability tests, among others

- WWII provided another boost to test development—all three military organizations had psychological testing programs with emphasis on development of specialized tests to assist in placing recruits in appropriate jobs
- Similar tests used extensively by industrial organizations after the war
- Use of some ability tests later halted in the late '60s and '70s because of EEO laws and early Supreme Court decisions
- Recently, use of ability tests in selection has increased substantially

### Definition of Ability Tests

- Written or physical performance tests that primarily measure knowledge in specific topics or physical processes in specific actions
  - We discuss and provide examples of mental, mechanical, clerical tests—types of test that are by far the most numerous ability tests, almost always paper-and-pencil tests that can be administered in a relatively short time
  - We also discuss physical ability tests which measure muscular strength, cardiovascular endurance, movement, coordination, similar physiological or motor characteristics—these tests usually require special equipment and are given individually
- Frequently ability tests called either *achievement* or *aptitude* tests
  - *Achievement* measure effects of formal learning on knowledge
  - *Aptitude* indicate how much knowledge or skill acquired “naturally”
- We now know that distinction between achievement and aptitude tests are unnecessary—*all tests measure what a person has learned up to the time he or she takes the test*
  - *Aptitude* and *achievement* have been replaced by the term *ability*

### COGNITIVE ABILITY TESTS

[slides 8–15]

- At the center of many of the early Supreme Court decisions regarding discriminatory effects of the use of tests in selection
- Recent work in selection indicates that for almost all jobs, cognitive ability tests are related to job performance
- Cognitive ability tests widely used in selection—principles governing appropriate use of ability tests have been developed for cognitive ability tests

### Development of Cognitive Ability Tests

- Thought to be the first work on mental ability or intelligence tests done by French psychologists Alfred Binet and Theodore Simon from 1905–1911 – attempted to develop tests that would identify mentally retarded children
- Binet and Simon sought to develop an age scale for each year between age 3 and adulthood—contained a sample of curriculum questions appropriate for instruction at each academic grade level

- A child's mental age based on correct answers to various grade-level scales, mentally retarded students identified as those whose calculated mental age was substantially below their chronological age
- In 1916, this test was translated for use in the U.S.—*Stanford–Binet Intelligence Scale*—modified periodically and still used today

### What is Measured

- Important to understand three points about these early tests
  1. Close association between the content of these tests and academic achievement
  2. Measure several distinct abilities—verbal, mathematical, memory, reasoning (Table 11.1 shows a list of various mental abilities)
  3. A variety of scores can be obtained from tests—measure different mental abilities and report scores on all items as one total score. This total score indicates overall cognitive ability. Other tests provide separate scores on each tested ability, then adds these score to report overall cognitive ability score. Another test measures separate abilities and does not combine scores—each reported individually.

### The Wonderlic Personnel Test

- Developed in 1938, used extensively by the U.S. military and private companies
- A multiple-choice test that consists of 50 items and allows 12 minutes for completion
- Items cover vocabulary, “commonsense” reasoning, formal syllogisms, arithmetic reasoning and computation, analogies, perceptual skill, spatial relations, number series, scrambled sentences (Table 11.2 contains items similar to those used in the *Wonderlic*, but not part of the test itself)
- Primary factors measured by the test are verbal comprehension, deduction, numerical fluency
- Can be administered via computer, mobile phone, or other devices that can be linked to the Internet, or in paper-and-pencil form during proctored administration
- Extensive set of norm scores developed through its long history
- Test publisher provides tables indicating distribution of scores by education level of applicants, region of the country, gender, age, ethnicity, also specific positions within industries—construction, hospitality, professional services, healthcare, manufacturing
- Normative data updated and published approximately every 10 years
- Common for multiple forms of the text to be used simultaneously for both online and proctored administration—parallel form reliability among forms ranges from 0.73 to 0.95, test-retest reliability is 0.82 to 0.94

### A Summary of Cognitive Ability Tests

- *Cognitive ability tests* and *intelligence* or *I.Q. tests* indicate the same type of tests—selection specialists can more appropriately conceptualize these tests as cognitive ability tests
  - The term *cognitive ability* makes explicit that these tests measure various mental abilities
  - Measures of an individual's ability to do cognitive tasks—mentally manipulate words, figures, numbers, symbols, impose logical order to solve problems
  - Tests measure academic ability, but does not mean that mental ability is useful only for academic selection

## THE VALIDITY OF COGNITIVE ABILITY TESTS

[slides 16–23]

- Extensive research on the validity of cognitive ability tests shows evidence that they are related to job performance

### Project A

- A multiple-year effort to develop a selection system appropriate for all entry-level positions in the U.S. Army
  - One major task was the development of 65 predictor tests that could be used as selection instruments
  - Another major task was the development of categories of work performance across entry-level jobs—five categories were determined
  - Selection specialists conducted a giant validity study – Table 11.3 presents only a small part of the results

### Validity Generalization Studies

- A validation study is necessary for each selection program being developed
- Differences in validity coefficients among studies that have used mental ability tests for similar jobs are due to methodological deficiencies in the validation studies themselves
- When these deficiencies are corrected, the differences among these validity coefficients across various organizations are close to zero

#### *Validity Generalization for the Same Job*

- Studies have analyzed data from validity studies conducted for the same job that use the same type of test as a predictor—Table 11.4 summarizes the corrected mean validity coefficients of some of these studies

#### *Validity Generalization for Different Jobs*

- Studies conclude that mental ability tests are valid across a large variety of jobs and can serve as useful selection instruments—Table 11.5 presents some results representative of these studies

#### *Comparison of Cognitive Ability and Other Selection Tests*

- Schmidt and Hunter compared validity of mental ability tests to that of 18 other selection procedures – Table 11.6

#### *Implications for Selection*

- No longer necessary to conduct validity studies within each organization
- Cognitive ability tests are valid predictors for a wide variety of jobs, however the magnitude of validity correlation between cognitive ability and task performance varies from job to job
- Scores obtained from a general cognitive ability test may be as good a predictor of job performance as are scores from a composite test of specific abilities related to the job

## COGNITIVE ABILITY TESTS AND ADVERSE IMPACT

[slide 24]

- Large differences exist in mean scores between white and black respondents and between white and Hispanic respondents
- Using these tests in making selection decisions usually results in a higher percentage of whites being selected than are members of the other two groups—the definition of adverse impact
- A Pareto analysis led to improvement in diversity hiring without significantly reducing the criterion-related validity

## MENTAL ABILITY TESTS AND THE INTERNET

[slides 25–28]

- Many organizations use electronic delivery systems for gathering data about job candidates
  - application blanks and résumés often submitted online
  - mental ability tests and other knowledge-based tests increasingly used in online forms
- Two issues of much concern
  - equivalence of the electronic and traditional forms of the tests
  - cheating

### *Equivalence of Paper-and-Pencil and Electronic Tests*

- Recent research has found that electronic and paper-and-pencil tests are usually not equivalent in terms of test taker reactions and scores
- Potosky and Bobko used two cognitively based tests—one timed, the other untimed—timed tests much more problematic to translate into an electronic version than untimed tests and probably do not result in equivalent forms – the mixing of scores on an paper timed test and an online timed test would be uninterpretable
- Table 11.7 presents their main conclusions

### *Unproctored Internet Testing*

- Scores upon which decisions about applicants are made can be affected by several sources of error:
  - Someone other than the applicant completing the test
  - The use of books or Internet material
  - Aid from accomplices
- General agreement that brief biodata tests, situational judgment tests, personality tests the most appropriate for Internet testing

## EFFECTS OF PRACTICE AND COACHING

[slide 29]

- Studies of the effects of coaching have found that such training has a minimal effect of test scores
- Studies of the effects of practice (reapplying and taking the same test two or more times) found that scores on mental ability tests increased between the first and second test completions and also between the 2nd and third completions—improvement does not necessarily translate into improved job performance
- Evidence that coaching and test preparation can lessen the gap between white and black test takers on a job knowledge test necessary for promotion

## MECHANICAL ABILITY TESTS

[slides 30–31]

- Written or physical manipulation tests that mainly measure spatial visualization, perceptual speed and accuracy, and mechanical information acquired by experience in an industrialized society
- Two testing methods generally used in mechanical ability tests—manual performance and written problems
- Mechanical ability tests can measure either general or specific abilities
- We discuss one of the most frequently used general mechanical ability tests—the *Bennett Mechanical Comprehension Test* (BMCT)

### The Bennett Mechanical Comprehension Test

- Has been the most widely used mechanical ability test for more than 50 years and has been utilized for a large number of different jobs—carpenter, engine lathe operator, welder, electrician, other skilled crafts
- Two versions of the BMCT—BMCT and BMCT II
  - BMCT consists of 68 multiple-choice questions with a 30-minute time limit—can be administered on paper or online
  - BMCT II consists of 55 multiple-choice questions with a 25-minute time limit—administered online only
- Questions measure respondent's ability to perceive and understand relationship of physical forces and mechanical elements in practical situations
- Reported reliabilities range from 0.81 to 0.93

## CLERICAL ABILITY TESTS

[slides 32–33]

- Written tests that predominately measure perceptual speed and accuracy in processing verbal and numerical data

### The Minnesota Clerical Test

- Developed in 1993, generally regarded as the prototype of clerical ability tests and most widely used
- Written at a second-grade reading level and to completed in 15 minutes
- One form consisting of two separately timed and scored subtests—number checking and name checking
- Table 11.8 contains items similar to those used in these subtests

## PHYSICAL ABILITY TESTS

[slides 34–38]

- Three reasons for this type of testing:
  1. EEO legislation has prompted an increase in women applicants for traditionally male-dominated physical labor jobs
  2. The use of appropriate selection devices for physically demanding jobs can reduce the incidence of work-related injuries

3. Because the Americans with Disabilities Act prohibits pre-employment medical examinations, the most feasible way to collect data about the physical status of applicants is through the use of specific physical ability tests that measure the worker characteristics required by the job

### The Fleishman Taxonomy

- Edwin Fleishman et al developed a taxonomy of 52 different abilities—both physical and nonphysical – necessary for performing work activities. We discuss only the measurement of the nine physical:

1. Static strength
2. Dynamic strength
3. Explosive strength
4. Trunk strength
5. Extent flexibility
6. Dynamic flexibility
7. Gross body coordination
8. Gross body equilibrium
9. Stamina

- Validity coefficients for specific jobs among the results:

- Pipeline workers (0.63)
- Correctional officers (0.64)
- Warehouse workers (0.39)
- Electrical workers (0.53)
- Enlisted army men (0.87)

### The Hogan Performance Test

- The extensive work of Joyce Hogan has produced the three components of physical performance described in Table 11.9. She combined two lines of research in the development of this taxonomy

1. Data about physical requirements derived from job analysis
2. Data based on physical ability tests already developed for selection

- By examining these two sources of information about physical work performance, a comprehensive model of physical abilities could be developed

- Results indicated that these tests were significantly correlated with both supervisors' ratings of physical performance and work simulation of critical job tasks

### Legal Issues in Testing Physical Abilities

- Selection specialists working with physically demanding jobs must be especially concerned with three groups of applicants—females, disabled workers, older workers

- Adverse impact for scores a common occurrence for each group

- A major study presented meta-analytic estimates of sex differences in physical abilities and the effects of selection system design, specificity of measurement and training in possibly reducing sex differences on physical ability test scores

- Height/weight standards and state laws prohibiting women from lifting items that exceed certain weight have been hard to defend on the basis they reflect strength – if strength needed, measure that directly

- No presumptions should be made that a disabled or older worker is unable to perform a job

## RECOMMENDATIONS FOR THE USE OF ABILITY TESTS IN SELECTION

[slides 39–40]

### Review Reliability Data

- Was it calculated correctly on similar workers?

### Review Validity Data

- Does the ability test measure what it is said to measure?
- Have the test scores been correlated with some measure of training or job performance

### Stuff to Remember About the Use of Ability Tests

- Ability tests are:
  1. Useful – valid predictors for all forms of job performance
  2. Cheap—can be purchased for a reasonable price
  3. Fast –most take 30 minutes or less to complete
  4. Easy—can be administered individually or group settings
  5. Versatile –many come in several languages
  6. Scorable—publishers usually provide a scoring key
  7. Understandable—what the test is measuring
  8. Sometimes falsely marketed—look at the test development data!

## CHAPTER 12

# *Personality Assessment for Selection*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Know and define the five personality traits of the Five-Factor Model.
2. Be aware of different measurement methods to assess these personality traits.
3. Understand why faking may be an issue and how to minimize its influence.
4. Know when a personality trait will be a useful predictor.
5. Apply steps to use personality and to establish job relevance.
6. Be aware of legal concerns.

### A BRIEF HISTORY

[slide 4]

- In the late 1960s, the utility of personality was questioned in the field of psychology by Mischel and in the work setting by Guion and Gottier—both critiques concluded that personality traits have little to no predictive relationship to actual behavior, consequently nearly a two-decade break in the use of research into personality at work
- In the late 1980s, research emerged to revive interest in the usefulness of personality in selection settings
- After positive findings were supported in research in the late 1990s, more research followed
- We review what was learned and which concerns should be considered in adopting personality testing as part of the selection approach

### DEFINITION AND USE OF PERSONALITY IN SELECTION

[slides 5—11]

- Personality is often equated with social skill—psychologists view this differently as they consider traits to be linked to motivational work habits
- Personality, therefore, would seem to be critically important in selection
- Researchers agree that personality characteristics can be grouped into five broad dimensions—the Five-Factor Model
- Managers intuitively believe personality traits matter at work—nearly as important to the candidate as general mental ability
- Recent meta-analytic data show these traits can be relevant predictors of the broad criteria of work effectiveness, retaining productive employees

- Personality traits contribute incremental validity to the prediction of success at work
- There is little or no adverse impact as mean scores tend to be smaller at the Five-Factor level than at the facet level of each trait
- Even if the effect is small to modest, the employer obtains this benefit each day the employee shows up for work
- Personality tends to change in a consistent trajectory during adulthood

### Personality Traits and the Five-Factor Model

- Personality data in selection requires classifying and measuring individuals according to some set of personality characteristics—*traits*
- Traits are used to explain the consistency of an individual's behavior over a variety of situations
- The essence of traits has been conceptualized in many ways—one common element is that they are viewed as the dispositional or relatively stable and enduring ways people tend to think, feel, act
- Personality data in selection requires specification of job tasks and identification of traits linked to these tasks
- Table 12.1 (derived from several studies) provides examples of personality traits that have been used in selection
- Five general factors thought of as the five core traits that influence behavior — the Five-Factor Model
  - *Extraversion*—sociability, gregariousness, assertiveness, optimism, ambition, activity
  - *Neuroticism* (viewed from emotional stability pole)—calmness, security, confidence, resistance to anxiety, lack of emotion
  - *Agreeableness*—courtesy, flexibility, good-naturedness, supportiveness, generosity, cooperativeness, forgiveness, tolerance
  - *Conscientiousness*—responsibility, organization, dependability, decisiveness, hard work, achievement orientation, perseverance
  - *Openness to experience* (also referred to as *intellect* or *culture*)—imagination, culture, curiosity, intelligence, artistic sensitivity, originality, broad mindedness
- The Five-Factor Model depicts personality at its broadest level of abstraction as nearly all of a person's characteristics are described using just these five traits
- Each of these traits is composed of a number of more narrowly defined specific traits—facets
- These facets are homogeneous in nature and are able to differentiate the multiple aspects of which a broad Five-Factor Model trait is composed
- A widely used Five-Factor personality test has differentiated the five broad factors into 30 more specific facets, each factor consisting of six distinctive facets—Figure 12.1

### Other Personality Traits

- *Core self-evaluations* consist of four frequently studied traits—self-esteem, generalized self-efficacy, locus of control, emotional stability
- *Emotional intelligence*—a broad construct that appears to be composed of attributes that go beyond just personality as it also includes interpersonal ability, self-efficacy, cognitive ability
- *Proactive personality* reflects a dispositional approach toward taking initiative at work and making changes to the work environment

## PERSONALITY MEASUREMENT METHODS

[slides 12—21]

### Inventories in Personality Measurement

- *Self-report questionnaires*—consist of a series of brief items asking respondents to use a multiple-choice answer format to indicate personal information about thoughts, emotions, past experiences—the five personality dimensions measured are presented in Table 12.2
- *Other self-report personality inventories*—some frequently used and names of the original authors are reported in Table 12.3
- *Forced-choice inventories*—require test-takers to choose the most liked item of two to four equally desirable items—advantage is that faking or response distortion tends to be reduced
- *Projective techniques*—similar to self-report questionnaires in that they require verbal responses that are scored to obtain measures of personality characteristics, but intentionally ambiguous

### Observer Ratings of Personality

- People interacting with the person (applicant) can rate the way that person tends to think, feel, act surprisingly well
- Ratings predict job performance better than self-report assessments, particularly when the observer knows the individual and is able to observe behavior or reactions relevant to the personality trait

### The Interview in Personality Measurement

- An interview frequently is used as a convenient way to determine how an applicant would typically act on the job
- One thorough review found that traits are the most frequently assessed constructs in the employment interview—conscientiousness is the single most frequently measured WRC in the interview
- Accurate assessment of a job candidate's personality may actually increase when using an unstructured interview rather than a structured interview—unstructured interview relies on questions that are open-ended and can be answered with a wide variety of responses resulting in more accurate assessment of the candidate's personality

### Evaluating the Validity of Personality Tests as a Predictor

#### *The Validity of Self-Report Inventories*

- A study quantitatively summarized the findings from 15 prior meta-analyses of the Five-Factor Model's predictive validity across five occupational groups and four specific work criteria, however did not focus on productive criteria that can be predicted at hire
  1. Conscientiousness and emotional stability were valid predictors of overall work performance
  2. Extraversion was a valid predictor for some occupational groups
  3. Agreeableness and openness to experience demonstrated modest validity overall
- A number of other meta-analyses recently conducted to examine unique criteria—leadership, expatriate success, etc.
  1. When selecting effective leaders from ineffective ones, results reveal that those candidates high in extraversion, openness to experience, and emotional stability experience the greatest success on the job

2. Teamwork is an important criterion
3. Two other meta-analyses show how these traits relate to expatriate job performance as well as entrepreneurial status
4. Two other recent studies examine deviant behavior and turnover
5. A final set of meta-analyses investigate the process through which personality affects job performance

## The Validity of Ratings from Other Observers

### *Validity of Projective Techniques*

- Various issues arise concerning scoring and use of information obtained from projective instruments that lead to questions about their general usefulness in selection
  - Reliability of an individual's responses at two different times
  - The impact on an individual's score due to the total number of responses given
  - The scoring of the information provided, for which the proposed benefit of projective instruments becomes a liability

## Legal Issues in the Use of Personality Tests

### *Faking in Personality Inventories*

- Do candidates intentionally alter their responses to increase the likelihood of receiving a job offer? Empirical evidence is contradictory
- What are the implications if candidates do fake? Little evidence that faking affects predictive validity of personality inventories
- Evidence that some applicants harmed by those who fake—this effect is greatest when only a small proportion of applicants are selected
- Research suggests that instructions should include a warning that faking may be detected

## Role of Technology and Global Trends

- Rapid changes in computer technology and web-based assessments continue to provide greater efficiency in testing, offer access to a larger, more diverse applicant pool, and provide more precise measurement of constructions, through application of item-response theory
- Web-based personality tests result in high levels of internal consistency—scores are equivalent to paper-and-pencil personality scores and have similar correlations between items
- Responses to personality tests seem to be influenced by cultural norms—differences in scores across cultural groups resulting from different response tendencies and nonequivalence across cultures
- Interpret scores by comparing to norms based on a cluster of countries that are culturally similar

## RECOMMENDATIONS FOR THE USE OF PERSONALITY DATA

[slides 22—27]

### Define Personality Traits in Terms of Job Behaviors

- Definition of important worker attributes necessary for all selection measures, however lack of a definition seems to be more common with personality measurement than other types of worker attributes

- Job analysis information could supply information to generate adequate definitions in two ways
  1. Through task approach
  2. Job analysis methods that produce worker attributes

### Define Work Effectiveness as Training Productive Employees

- Possible to identify and predict a broad overall work effectiveness criterion *at hire*, but only if we simultaneously consider those WRCs that predict retaining employees who are also highly productive
- Predictors need to be comparably broad to the overall work effectiveness measure and each personality trait should be psychologically meaningful yet capture the general factor common to the many diverse facets of that trait
- Generally need to focus on a small set of predictors to ensure significant gains in incremental validity

### The Appropriateness of the Selection Instrument

- Two important characteristics of the personality selection measure should be considered
  1. Breadth of the personality trait being measured—more effective in selection to use traits that affect a wide set of behaviors simultaneously rather than a narrow set
  2. Whether respondents can learn “correct” responses to the instrument—instruments for which appropriate responses are not apparent are preferable to those for which answers are apparent

#### *Source of Measurement*

- Observer ratings of personality shown to predict performance better than self-reports, thus considerable value to considering ways to assess personality in the interview (letters of reference, headhunters)

#### *Incremental Validity over Other Selection Measures*

- Low correlations among various predictors reduce the extent of overlap or redundant information being assessed and thus result in incremental gains in predictive validity when combining multiple predictors

#### *Trade-offs between Validity and Disparate Impact*

- Practitioners should continue to examine mean score differences between different subgroups of applicants to ensure disparate impact is reduced

#### *Faking “Good” during Selection*

- We know that including warnings in the instructions that faking on the personality test can be detected, and may require retesting, tends to discourage faking—warnings work

#### *Cheating during Selection*

- The Internet provides for more efficient testing more cheaply—such tests should be proctored to enhance standardization of the testing procedures as well as to ensure security

#### *Using “At-Work” Personality Tests*

- Contextualizing personality items so that they encourage the test-taker to think about their behavior “at work” can lead to superior predictive validity—linking responses through at-work items leads to significant gains in predictive validity

*Realizing Prediction Gains over Time*

- Hiring candidates based on their personality can lead to substantial long-term gains for both the individual and the organization—reveals the value of using personality ensure effective selection

**CONCLUSIONS**

[slide 28]

- Traits vary greatly in the extent to which they influence behavior
  - A small number of core traits have a strong influence on behavior when aggregated over many instances of that behavior
  - A larger number of narrower facets will achieve higher predictive validity, but only for relatively limited or more precise outcomes
- The situation has an important influence on an individual's behavior
- Measuring personality is challenging—steps must be taken to reduce “faking good” or cheating on web-based assessments

## CHAPTER 13

# *Simulation Tests*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand what a simulation test is.
2. Know the advantages of using a simulation test in selection.
3. Know the limitations of using a simulation test in selection.
4. Understand what a work sample test is.
5. Understand what an assessment center is.
6. Understand what a situational judgment test is.

### SIMULATION TESTS

[slide 4]

- Selection devices that assess applicants by means of testing situations that resemble actual parts of the job being considered—the terms *performance test* and *work samples* also frequently used
- Examples of simulation tests are requiring applicants to
  - write a simple computer program to solve a common work problem
  - fix a plumbing leak in a house
  - work together to design a marketing strategy for a new product

### CONSISTENCY OF BEHAVIOR

[slides 5–7]

- Wernimont and Campbell stated that selection decisions are most accurate when “behavioral consistency” is the major characteristic of the selection program
- To clarify their point they categorized all selection devices as either signs or samples
  - *Signs*—selection tests used as secondary indicators of an individual’s behavior in job activities
    - Examples: application forms that ask for degrees, former job titles, years of work experience
  - *Samples*—selection tests that gather information about behaviors consistent with the job behaviors being predicted
    - Examples: instruments that gather information about work experience and educational history related to specific job behaviors, simulation exercises requiring applicants to complete a set of actions that replicate job behaviors

## Limitations of Simulations

- Difficult to construct simulation tests representative of job activities
- Simulation tests usually developed on the assumption that the applicants already have the knowledge, ability, skill to complete the job behavior required in the test, if specialized knowledge required, training in that specialized knowledge should be included before the test begins—makes the test longer and more difficult to administer
- Simulations much more expensive than other selection devices

## Types of Simulation Tests

- By content of the test:
  - *Motor*—simulation that replicates physical activities of the job (e.g., operating a machine, installing a piece of equipment)
  - *Verbal*—simulation that replicates the verbal, cognitive, or people-interaction activities of the job (e.g., simulating an interrogation, demonstrating how to train subordinates)
- Fidelity:
  - Refers to the degree to which the simulation matches or replicates the physical and/or psychological demands and activities of the job for which it is designed

## WORK SAMPLES

[slides 8–12]

- The oldest type of simulation used in the early 1900s to select clerical staff (tests of typing and editing a business letter and skilled tradespeople (automobile repair, carpenters, masons)
- Samples of job tasks that applicants are asked to perform—performance is scored
- High-fidelity simulations—either motor or verbal
- Table 13.1 presents examples of various work-sample tests and the jobs for which they have been used—a wide variety of tests have been used, even for the same job

## The Development of Work-Sample Tests

- Steps taken to develop these tests:
  1. Perform job analysis
  2. Identify important job tasks to be tested
  3. Develop testing procedures:
    - Select tasks
    - Specify testing procedures
    - Establish independent test sections
    - Eliminate contaminating factors
    - Select the number of test problems
  4. Develop scoring procedures (Table 13.3)
  5. Train judges

## The Validity of Work-Sample Tests

- Studies that examine the validity of work-sample tests in selection have been consistently positive—many single-sample studies
  - Authors determined only a small difference in validity when objective (measured outcome of the task) or subjective (supervisor's judgment) job performance criteria measures were used
- Work-sample tests demonstrate much smaller differences between groups of black and white applicants than many written tests – especially cognitive ability tests
- Other benefits—no complaints lodged about their appropriateness
- Can also serve as a way of displaying the actual job to the applicant

## ASSESSMENT CENTERS

[slides 13–28]

- When simulation tests are used for selection of managers, professionals, executives, they are commonly referred to as *assessment centers* (ACs)
- An AC is a procedure for measuring WRCs in groups of individuals (6-12 people) using a series of devices, many of which are verbal simulation tests
- ACs have been used for both selection and career development—we concentrate on those used for selection
- The fidelity of simulations in an AC can vary from high to fairly low depending on the specific makeup of the AC—some designed for a particular job, others general appropriate for a group of jobs, some include traditional tests
- Because of this variation we think about ACs as having moderate fidelity

## The Beginning of Assessment Centers

- The Management Progress Study of AT&T marked the beginning of the use of the AC for industrial organizations
  - Begun in 1956 to study the career development of men hired for managerial positions
  - One major focus of the study was to identify the WRCs thought to be related to the successful career progress of managers—one challenge was developing measuring devices
    - A three-and-a-half day assessment center was devised
    - Managers brought together in groups of 12—several ways tried to measure personal characteristics
    - Data obtained from these measures were related to subsequent movement through managerial levels

## What Is Measured and How

- Dimensions
  - Starts with a job analysis to identify clusters of job activities that are important parts of the job of interest
  - Each cluster should be specific and observable, comprising job tasks that are related in some way

- These job clusters (*dimensions*) are measured by the AC devices – think about dimensions as the WRCs measured in ACs
- Table 13.4 provides brief definitions of nine dimensions commonly used in assessment centers—note that these dimensions are defined based on actual job activities
- Traditional assessment devices
  - The AT&T ACs often used various types of traditional tests and interviews—mental ability, projective personality, paper-and-pencil personality tests
  - Because of the expense of using and interpreting such tests, no longer used very often
  - However ACs do commonly use an in-depth interview that is similar to the background interview—gather information from the candidate about job activities that represent the behavioral dimensions being evaluated
  - Each interview is structured and focuses on previous job behaviors—formal scoring system used to evaluate each behavioral dimension, shown to be effective at arriving at a final evaluation
- Simulation tests
  - *In-Basket*—a paper-and-pencil test designed to replicate administrative tasks of the job under consideration (Table 3.5)
  - *Leaderless group discussion* (LGD)—designed to represent those managerial activities requiring interaction of small groups of individuals in order to solve a problem successfully, participants tested in groups of six, no one member designated as the official leader or supervisor of the group (Table 13.6)
  - *Case analysis*—each participant provided with a long description of an organizational problem that changes according to the job being considered, case frequently describes history of certain events in a company with relevant financial data, marketing strategy, organizational structure, case focuses on a dilemma that the participant is asked to resolve (must give specific recommendations, present supporting data, detail any changes in company strategy)

## The Training of Assessors

- If an AC is to be useful as a selection device, it is crucial that *assessors* (staff members who have the responsibility of observing and evaluating behaviors of participants) have the necessary training (Table 13.7)
- The major duty of an assessor is to record the behavior of a participant in an exercise and use that data to rate the participant on each behavioral dimension appropriate for the exercise
- Assessors are usually managers within the organization
- After all exercises have been completed, assessors come together to discuss their observations
- Ratings and data gathered by each assessor are used to develop group or final scores of each participant on each dimension
- Understanding the behavioral dimensions
  - The first step in training assessors is to familiarize them with the behavioral dimensions – all assessors should have a common understanding of the dimensions
  - Example definition that could be used for an AC dimension focused on first-line supervisors:  
*Tolerance for stress: Stability of performance under pressure and/or opposition. A first-line supervisor finds himself/herself in a stressful situation because of three main factors: (a) multiple demands on the work unit that must be completed at approximately the same*

*deadline, (b) the joint roles he/she must play as both a representative of management to nonmanagement employees and a representative of nonmanagement employees to management, and (c) confrontation by employees who are angry or hostile because of a work situation.*

- Observing the behavior of participants
  - Initial tendency of assessors is to immediately make evaluative judgments about the performance of participants in AC exercises—instead they should focus on recording the behavior of the participant
- Categorizing participant behavior
  - Assessor must record the behavior of the participant under the proper dimension
- Determining the rating of participant behavior
  - Assessors must be consistent in the use of rating scales to evaluate the behavior of participants on the dimensions
- Determining dimension and overall evaluation ratings
  - Combine the data on the same dimension across two or more exercises—focuses on using PEDRs (post-exercise dimension ratings) to develop DRs (dimension ratings)
  - Last step in training concerns how to use the DRs to form the OAR (overall assessment rating) of the participant's ability in the complete AC to determine the most critical dimensions which are weighted more heavily in producing the overall rating
  - Assessor groups are often required to complete a mock assessment or a small group of candidates under observation of experienced assessors

## The Validity of Assessment Centers

*Phase 1: We Love Assessment Centers!*

- Ratings predictive of movement into middle-level management
- Reduced or eliminated both racial and gender differences—evidence ACs could produce adverse impact
- For about 25 years after AC development at AT&T, selection specialists loved ACs

*Phase 2: This Is Not Working and It's Your Fault, Not Mine!*

- 1982 research study focused on the failure of ACs to demonstrate the pattern of correlations among dimension ratings they were predicted to produce
- If the dimension of planning and organizing measured with in-basket test, interview, and case analysis, then correlation of ratings for participants should be quite high—referred to as *convergent validity*—but evidence show that ratings were very low

*Phase 3: Well I Might Have Made a Mistake and We May Be Okay After All.*

- This phase of thinking and research on ACs started in the early 2000s
- Its major theme is that the measurement model that expected PEDRs to correlate for the same dimension across exercises and not to correlate among different dimensions within an exercise is an incorrect way of thinking about ACs—AC OAR scores found to consistently correlate with various measures of work performance

## SITUATIONAL JUDGMENT TESTS

[slides 29–30]

- Situational judgment tests (SJTs) are verbal simulations and regarded as low-fidelity because they are almost exclusively descriptions of work situations or behaviors rather than actual replications
- They ask the respondent to choose among multiple-choice alternatives that are various behaviors that may be taken to address the work situation described in the question
- Table 13.8 contains examples of item stems (questions) and possible responses that may be contained in a situational judgment test

## CHAPTER 14

# *Testing for Counterproductive Work Behaviors*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Know what is meant by the term *counterproductive work behaviors* (CWBs).
2. Understand why testing for CWBs has become so important for organizations.
3. Understand how polygraph testing is carried out and its limitations.
4. Know how both overt and personality integrity tests are written and the differences between these two.
5. Understand why integrity tests are related to both job performance and CWBs.
6. Know the validity of integrity tests when used in selection.
7. Understand the legal issues when organizations use integrity tests.
8. Understand the differences among the various types of drug tests and what is measured in each.
9. State the legal issues when organizations use drug tests.
10. Understand what is measured in genetic testing and its limitations.
11. State the legal issues when organizations use genetic testing.
12. Know the basis for neuroscience lie detection and its limitations.

### TESTING FOR COUNTERPRODUCTIVE WORK BEHAVIORS

[slide 4]

- Counterproductive work behaviors (CWBs) is the term used to include a wide array of illegal or unethical forms of purposeful employee conduct that harms organizations or people—*theft, fraud, illicit drug use, violence, cursing, sabotaging, purposely doing work incorrectly*
- In this chapter we cover the following topics:
  1. Integrity testing using polygraph devices (lie detectors)
  2. Integrity testing using written tests
  3. Testing for drug and alcohol use by gathering physical evidence
  4. Genetic testing of individual to find indicators of disease or physical deficiencies
  5. Neural imaging of brain activity to detect various activities and physical characteristics

## POLYGRAPH TESTING

[slides 5–7]

- A polygraph is a machine that measures the physiological responses of examinees as they give verbal responses to the direct questions of a polygraph operator
- Although once a mainstay for hiring retail personnel in large “big-box” stores, the use of polygraphs for selection is now illegal—a U.S. federal law restricts the use of polygraph for selection to a very narrow set of circumstances (most legal applications involve organizations engaged in police work or domestic security)

### *Procedures*

- The examiner conducts a pretest discussion that covers all questions used in the test—purpose is to make sure examinee understands the wording and meaning of the questions and can answer with a simple yes or no
- After this discussion, the polygraph is attached to the examinee, and the actual interview is conducted—list of questions usually repeated once or twice to obtain more reliable data
- Examination typically entails three types of questions
  1. Irrelevant, nonemotional question (“Are you six feet tall?”)
  2. Emotional control question (“Did you ever lie ...?”)
  3. About behavior of interest—concerns issues such as prior thefts
- To detect lying, operator looks for evidence of autonomous disturbance associated with answers to the last type of question

### *Limitations*

- One difficulty with polygraph testing is that other reactions besides guilt can trigger an emotional response
- Another difficulty is the variety of countermeasures that examinees can use to avoid detection
- The major drawback to using the polygraph is the frequency of *false-positive* results—results indicating individuals are lying when in fact they are not
- *False negatives* have also received attention—occur when test results are judged as truthful when in fact they are not

## INTEGRITY TESTING

[slides 8–18]

### Paper-and-Pencil Integrity Tests

- Because the Employee Polygraph Protection Act of 1988 greatly restricted the use of polygraph testing, the attention of both companies and test developers turned to paper-and-pencil integrity tests—no federal law prohibiting such testing so permissible unless a state has passed a law specifically limiting or prohibiting the use of such a test
- Two forms of these tests developed
  - a. Overt integrity test—printed test that requires answers about behaviors typical of individual predisposed to exhibit critical work-related behaviors (CWBs)
  - b. Personality-oriented measures—printed test that requires answers to broad-based personality questions

*Overt Integrity Tests*

- Rationale underlying this type of test is to measure job applicants' attitudes and cognitions towards theft that might predispose them to steal at work—especially when the need and opportunity are present
- Past research has shown that the “typical” employee-thief
  - a. is more tempted to steal
  - b. engages in many common rationalizations for theft
  - c. would punish thieves less
  - d. often thinks about theft-related activities
  - e. attributes more theft to others
  - f. shows more inter-thief loyalty
  - g. is more vulnerable to peer pressure to steal

*Personality-Oriented Measures*

- Employee theft is just one element in a larger syndrome of antisocial behavior or organizational delinquency—dishonesty, theft, drug and alcohol abuse, vandalism, sabotage, etc.
- The assumption is that there is a common personality pattern underlying organizational delinquency—feasible to identify those with this pattern by using personality inventories
- Frequently these measures will contain a subset of items identified as an integrity scale – two advantages
  1. Items on a personality-based integrity test not as transparent in their intention to measure theft and related transgressions—reduces the possibility of fake responses
  2. Applicants not selected need not be rejected due to failing an integrity test—rather they are rejected for a mismatch between their personality profile and that of successful employees
- The Hogan Personality Inventory is one personality-based integrity measure—a 206-item, true or false response, Big-Five type of personality survey to be taken in 15 to 20 minutes
- Developed specifically for use in employment situations—contains seven primary scales and six occupational scales (Table 14.1)
- Primary scales are personality dimensions measured by the 206 items
- Occupational scales are composed of various combinations of the 206 items of the primary scales—developed to provide information about specific employment topics
- One occupation scale is the Reliability scale—composed of 18 items that measure honesty, integrity, organizational citizenship—high scores are good, low scores are bad (Table 14.2)

**What We Know About Integrity Tests**

- *Constructs measured*—may be hierarchical in nature with conscientiousness, agreeableness, and emotional stability representing the general factor
- *Variables related to integrity tests*—the relationship between integrity measures and cognitive ability depends on which factors of integrity are measured by a test

- *Relationship between overt and personality-oriented tests* – a study using test items from 3 overt and 4 personality-oriented tests concluded that:
  1. there were positive correlations – showed a great deal of similarity in what they measured
  2. there were some differences—overt tests correlated more highly with honesty and supervision attitudes, personality tests correlated more highly with self/impulse control, home life/upbringing, risk taking, diligence, emotional stability
- *Validity*—determining the validity of integrity tests is difficult because of the problems of measuring a criterion variable with which to correlate integrity scores
  - Table 14.3 compares two sets of meta-analyses integrity test results from studies for both job performance and counterproductive behavior criteria, shows overall true validity estimates as well as findings for operational validity

#### *Usefulness of Integrity Tests*

- Harris et al calculations revealed that use of an integrity test in pre-employment decision making would lead, on average, to a savings of over \$200,000 per 1,000 job applicants for a typical organization
  - *False positives*—all selection tests have false positives, the use of personality-based integrity tests may be more socially acceptable than the use of overt measures
  - *Faking*—research on faking on integrity tests comes to the conclusion that faking does not seem to make a difference in the validity of the test

### Legal Issues in Integrity Testing

- Harris et al observed that over a 60-year period, only about 30 formal complaints were filed against integrity tests—all were dismissed indicating that integrity tests do not elicit many legal complaints and that complains involving such tests are likely to be defensible
- *Other stuff*—new formats of integrity testing have been tried but are not used extensively given the demonstrated validity of integrity tests

## DRUG TESTING

[slides 19–25]

- There are a number of reasons for testing job applicants for drug use or abuse, which include the following:
  1. Deter employees from abusing alcohol and drugs
  2. Prevent hiring individuals who use illegal drugs
  3. Provide a safe workplace for employees
  4. Protect the general public and instill consumer confidence that employees are working safely
  5. Comply with state laws or federal regulations
  6. Benefit from Workers' Compensation Premium Discount program
  7. Reduce costs of healthcare claims, particularly short-term disability claims
- Additionally, research has found that drug use is negatively associated with job performance and positively correlated with accidents, injuries, absences, involuntary turnover, job-withdrawal behavior

## Drug Testing Methods

- Paper-and-pencil tests – simplest, and least controversial, drug test
- Urine tests—used more often
  - Basic urine test relatively inexpensive
  - Results returned quickly (usually within 24 hours)
  - Testing usually convenient for the applicant
- Hair analysis – becoming more frequently used, can detect drug use over a longer period of time than urine testing
- Oral fluid tests—oral swab rubbed on the inside of the mouth, saliva on the swab analyzed for the presence of drugs
  - Barring lab error, this test is always accurate

## Accuracy of Chemical Tests

- Because the physical properties of individual drugs are invariant, chemical tests should be completely accurate
- False positives may occur in an initial test, however these errors should be eliminated by follow-up confirmation tests
- Errors in confirmation tests occur only when the lab conducting the testing does not use appropriate, standardized procedures

## Legal Issues in Pre-Employment Drug Testing

- In general, employers testing job applicants for drugs in the pre-employment phase are less legally exposed than organizations testing current employees
- Major questions about drug testing have centered on six legal arguments (most applicable to current *employees*)
  1. Testing represents an invasion of privacy
  2. Testing constitutes an unreasonable search and seizure
  3. Testing is a violation of due process
  4. Drug users are protected under the Americans with Disabilities Act
  5. Testing may violate the Civil Rights Act
  6. Testing may violate the National Labor Relations Act
- Drug testing has caused negative reactions among applicants as well as employees—important for selection and HR specialists to understand reactions to a drug-testing program
- In general, employees and applicants react more favorably to drug testing when
  - a. advance warning of the testing is given
  - b. company uses rehabilitation of employees rather than termination when drugs detected
  - c. drug testing adheres to fair detection procedures and explanation of results
- Also, employees and applicants are more positive about drug testing when there is perceived to be a need for the test
- Drug users had much more negative reactions to all forms of testing than did nonusers—most do not regard the negative reactions as a bad thing

- One current drug-testing issue facing companies is pre-employment drug testing of applicants for jobs located in states where marijuana has been legalized—no apparent legal cases challenging an employer's refusal to hire someone who has tested positive for marijuana and the applicant showed the employer a medical marijuana card
  - As long as marijuana is considered to be an illegal drug under federal law, its use is not protected by the ADA

### Guidelines for Drug-Testing Programs with Job Applicants

1. A company is in the most legally defensive position when it limits testing to those positions that have major safety implications or a history of poor performance in specific areas that might be linked to drug usage
2. Private employers have more latitude in conducting drug testing as they choose unless an employer is subject to federal regulations
3. Private employers have more flexibility in their drug-testing programs with applicants than with existing employees
4. Combination of screening and confirmatory tests necessary for valid drug testing
5. Company should obtain written consent of the applicant before testing and provide the individual with the test results afterward
6. Standardized procedures used in the testing program should be applied to all job applicants tested
7. Program should be designed and reviewed periodically to ensure privacy is afforded to individuals being tested

## GENETIC TESTING

[slides 26–29]

### What Is Genetic Testing?

- An analysis of human DNA, RNA, chromosomes, proteins, or metabolites that detects genotypes, mutations, or chromosomal changes
- Genetic testing produces *genetic information* on individuals, family members of those tested, individuals' family members' health histories
- In some cases, insurers having access to genetic information denied coverage based on pre-existing conditions, in other cases, employers possessing knowledge of genetic information denied jobs to certain applicants or fired existing employees
- The rationale behind these decisions is that applicants or employees having the wrong genetic profile might develop expensive diseases or become injured on the job thus increasing workers' compensation costs

### Legal Concerns

- A key legal case (*Norman-Bloodsaw v. Lawrence Berkeley Laboratory*) raised the consciousness of the federal government regarding genetic testing and discrimination against certain racial minorities
  - Without the individual's knowledge or consent, the lab tested for sickle-cell genetic markings linked to sickle-cell anemia—the plaintiffs in the case won based on the lab's invasion of applicants' privacy

- The decision led Congress to recognize that genetic discrimination could occur in the workplace through genetic testing – could produce adverse effects against gender, racial, and ethnic groups as well as applicants having certain diseases not related to the applicants' ability to perform their jobs

#### *The Genetic Information Nondiscrimination Act of 2008 (GINA)*

- After 13 years of congressional debate, President George W. Bush signed this Act into law—consists of several parts (titles), we focus on Title II because it specifically addresses issues most relevant to employers
- Title II prohibits an employer, employment agency, labor organization, joint labor-management committee from discriminating against employees and individuals because of genetic information
- Relatively few cases tried under the provisions of GINA, however formal genetic discrimination complaints have been filed with the EEOC—involving companies asking for medical history rather than genetic testing
  - GINA also prohibits gathering such medical history
- Table 14.4 summarizes some job applicant protections and employer requirements of Title II

## **WHAT'S NEXT? NEUROSCIENCE-BASED SELECTION**

### **[slide 30]**

- *Possibly*—at least some point in the future when enough scientific evidence is available and assuming the evidence as well as state/federal laws are supportive
- Neuroscientists have begun to examine relationships between neural circuits of the brain and development of diseases as well as behaviors and personality characteristics
- Like genetic testing neuroimaging does not mean perfect predictions can be made—although significant progress made, still developing
- However, with current and future research, possible to describe in terms of probabilities the likelihood that certain neural wirings are linked to specific outcomes such as disease and behavior—concerns with the ethical impact
- An attorney in New York thinks a GINA-like federal statute should be adopted similar to Title II—would prohibit employers from requesting, acquiring, disclosing neuro information and from discriminating on the basis of neuro information

## CHAPTER 15

# *Strategies for Selection Decision Making*

### LEARNING OBJECTIVES

After studying this chapter, you will be able to:

1. Understand the importance of using mechanical methods for collecting and combining selection procedure data for making selection decisions on job applicants.
2. See by example how two mechanical methods are used for combining selection procedure information prior to making selection decisions.
3. Decide which strategies for making employment decisions are most appropriate in particular organizational contexts with specific employment needs.
4. Understand the importance of auditing an organization's selection decisions and learn from successes and failures.

### STRATEGIES FOR SELECTION DECISION MAKING

[slides 4–13]

- By following sound, systematic decision-making procedures and studying decision outcomes, managers can improve their effectiveness
- Selection decisions come in two basic varieties—simple and complex
  - Simple selection decisions involve one position opening with several applicants who are assessed on WRCs important to job success
  - Complex selection decisions involve a number of applicants and several position openings—here, the decision is who to select and which candidate is best for each job
- Answering the following three questions will enhance the quality of managers' employment decision making:
  1. For a specific selection situation, what are the best methods for collecting predictor information on job applicants?
  2. Because we often collect application information using more than one predictor, how should we combine scores on the predictors to arrive at an overall score selection decision-making purposes?
  3. Once an overall or total score on two more predictors is obtained, how should this overall score be used to make selection decisions?

## Modes for Collecting Predictor Information

- We all make selection decisions in our daily lives, but we sometimes make errors—*measurement error*
- In personnel selection, decision makers collect information from job applicants using predictors that generally reflect one of two different measurement philosophies
- Table 15.1 shows examples of modes used to *collect* predictor information representing these philosophies

## Modes for Combining Predictor Information

- Once selection procedures' data have been collected, the information must be combined to reach a selection decision
- Table 15.2 summarizes two modes—*judgmental* and *mechanical* for *combining* predictor information for selection decision-making purposes
- When predictor data are combined using human intuition (gut instincts) to reach an overall assessment, then predictor data have been combined judgmentally
- Predictor information combined by entering applicant test and interview scores into a statistical equation developed to predict work performance involves use of a *mechanical* mode

## Methods for Collecting and Combining Predictor Information

- We describe six methods used for collecting and combining predictor information from job applicants
- These methods describe how selection decision makers collect and combine predictor information mechanically, judgmentally, or both
- Table 15.3 lists these methods and an example of each—some methods more effective than others

## Which Method Is Best?

- In 1954, Paul Meehl's work concluded that clinical experts' intuitive predictions of human behavior were significantly less accurate than predictions made using more formal, mechanical means
- A later review found that *pure statistical* and *mechanical composite* methods for collecting and combining data were always either equal or superior to all other methods
- A more recent view concluded that the major issue is *how* data are *combined* in making a prediction rather than how the data are collected
  - Mechanical methods of combining predictor data relative to clinical or judgmental combination methods improved the ability to predict work performance by more than 50 percent
  - The rate of identifying acceptable hires was *reduced* by more than 25 percent when judgmental data combination methods were used rather than mechanical means

## Implications for Selection Decision Makers

1. Be careful about relying too heavily on résumés and other initial information collected early in the selection process
2. Use standardized selection procedures that are reliable, valid, and suitable for the specific selection purpose
3. When feasible, use selection procedures that minimize the role of selection decision maker judgment in collecting information

4. Avoid using judgment in *combining* data collected from two or more selection procedures used for determining applicants' overall scores
5. When clinical judgments are used in combining data, there is a significant loss of information, and decision quality is substantially reduced

## METHODS FOR COMBINING PREDICTOR SCORES

[slides 14–22]

- There are many different ways of mechanically combining predictor scores
- We describe and give an example of two methods—*multiple regression* and *unit weighting*—that conform to our prescription of using mechanical methods for combining predictor information.
- Each method has a long history of use in a variety of disciplines including human resource selection
- Table 15.4 presents data for the job of patient account representative

### Method One: Multiple Regression

- Shows the maximum linear association between two or more selection procedures or predictors and a criterion for a sample of individuals
- Applicants' scores can be entered into an equation—prediction equation

$$\hat{Y} = 5 + 2X_1 + 1X_2$$

where

$\hat{Y}$  = predicted work performance

5 = a constant or intercept value of the regression line

2, 1 = regression weights for test  $X_1$  and test  $X_2$

$X_1, X_2$  = applicants' scores on the tests

- Because it is possible to compensate for low scores on one predictor by high scores on another, multiple regression is sometimes referred to as a *compensatory* method
- Multiple regression makes two basic assumptions
  - a. Predictors are linearly related to the criterion
  - b. Because predicted criterion score is a function of the sum of the weighted predictor scores, predictors are additive and can compensate for one another
- The multiple-regression strategy has several advantages:
  - It minimizes errors in prediction and combines the predictors to yield the best estimate of applicants' future performance
  - It is a very flexible method—can be modified to handle nominal data, nonlinear relationships, both linear and nonlinear interactions
  - Regression equations can be constructed for each of a number of jobs using either the same predictors weight differently or different predictors
- The multiple-regression approach has disadvantages too:
  - There are statistical issues that are sometimes difficult to resolve
    - When a relatively small sample size is used to determine regression weights or the predictors are correlated with one another, the weights will not be stable from one sample to the next and standard errors of the weights increase—cross validation of the prediction equation essential prior to application

- If no preliminary screening of applicants is done, the multiple regression strategy requires assessing all applicants on all predictors—costly with a large applicant pool

### Method Two: Unit Weighting

- Multiple-regressions models are referred to as a proper linear model because predictors are identified to maximize the relationship between the predictors and the criterion
- An improper linear model is chosen by some nonoptimal means—various predictor weights could be chosen based on a theory, job analysis results, SME opinions, or in the case of unit weighting given equal weights
- Unit weighting of predictors has been used successfully in a wide variety of disciplines and is appropriate in certain human resource selection applications
- With sample sizes of up to 200, unit weights and regression weights perform about equally as well, however when sample sizes are less than 75, unit weights perform better in prediction than regression weights
- How is unit weighting of predictors calculated? Robyn Dawes and Bernard Corrigan noted: “The whole trick is to decide what variables to look at and to know how to add.”
- To calculate unit weights, see Table 15.5 where we show just the math test scores and structured interview ratings for each of five applicants
  - Raw math and interview scores are converted into standardized  $z$  scores
  - Add all applicants’ predictor  $z$  scores to obtain their composite  $z$  score
  - A  $z$  score shows the relative standing that an individual has on a predictor or composite predictor relative to others in the sample

## STRATEGIES FOR MAKING EMPLOYMENT DECISIONS

[slides 23–41]

- Numerous strategies concerned with selecting individuals for jobs have been reported in the human resource selection literature—we look at the following:
  - a. Top-down selection
  - b. Cutoff scores
  - c. Multiple cutoff scores
  - d. Multiple hurdles
  - e. Combination method
  - f. Banding

### Strategy One: Top-Down Selection

- Job applicants are rank-ordered from highest to lowest based on their predictor scores
- Job offers are made to the applicants beginning with the applicant with the top score, the applicant ranked second, etc., until all openings are filled
- As far as work performance is concerned, maximum utility is gained from a valid predictor when top-down hiring is used
- The biggest problem with top-down selection is the possibility of adverse impact against legally protected racial/ethnic groups

## Strategy Two: Cutoff Scores

- A cutoff score represents a score on a predictor or combination of predictors below which job applicants are rejected
- Judgment will play a role in choosing the method for setting the cutoff score and for determining the actual cutoff score value to be employed in selection—when a cutoff score is developed and used, the rationale and specific procedures for identifying that particular score should be carefully documented
- Cutoff scores can be established in two general ways
  - a. How job applicants or others performed on a selection procedure
  - b. Using judgments of SMEs regarding the appropriateness of selection procedure content to set the cutoff score

### *Basing Cutoff Scores on Applicants' or Others' Performance*

- Cutoff scores can be determined by administering the selection procedure to individuals (job incumbents) other than applicants and using that information as a basis for score development
- When a group of individuals other than applicants serves as a basis for deriving cutoff scores, care should be taken in determining that they are comparable to the applicant pool for the job
- When the groups are not comparable, legal questions concerning the fairness and meaningfulness of such scores can arise

### *Using Experts' Judgments*

- The *Ebel method* is based on an analysis of the difficulty of test items (Table 15.6)
- In the *Angoff method*, judges or SMEs estimate the probability that a minimally qualified applicant could answer a specific test item correctly—these estimates are used to establish cutoff scores for the test (Table 15.7)

### *Legal and Psychometric Issues*

- The *Uniform Guidelines* states that “...the degree of adverse impact should be considered”
- The *Principles for the Validation and Use of Personnel Selection Procedures* states “if based on valid predictors demonstrating linearity...throughout the range of prediction, cutoff scores may be set as high or as low as needed to meet the requirements of the organization”
- Table 15.8 lists guidelines that should be considered in using cutoff scores in selection

## Strategy Three: Multiple Cutoff Scores

- Each applicant is assessed on each predictor
- All predictors are scored on a pass-fail basis
- Applicants are rejected if any one of their predictor scores falls below a minimum cutoff score
- This method makes two important assumptions about work performance
  - a. A nonlinear relationship exists among the predictors and the criterion
  - b. Predictors are not compensatory
- Advantages of this method:
  - Narrows the applicant pool to a smaller subset of candidates who are all minimally qualified for the job
  - Conceptually simple and easy to explain to managers

- Two major disadvantages:
  - Requires assessing all applicants using all predictors—with a large applicant pool, the selection costs may be large
  - Identifies only those applicants minimally qualified for the job—no clear-cut way to determine how to order those applicants who pass the cutoffs

### Strategy Four: Multiple Hurdles

- Each applicant must meet the minimum cutoff or hurdle for each predictor before going to the next predictor—applicants must pass the predictors sequentially
- Failure to pass a cutoff at any stage results in the applicant being dropped from further consideration
- In a variation of the multiple-hurdle approach—the *double-stage strategy*—two cutoff scores are set: C1 and C2 (Figure 15.1)
  - Those whose scores fall below C2 are accepted unconditionally
  - Those whose scores fall below C1 are rejected terminally
  - Those whose scores fall between C1 and C2 are accepted provisionally

### Strategy Five: Combination Method

- The combination method is a hybrid of the multiple-cutoff and multiple-regression approaches
- Two major assumptions
  1. The more restrictive assumption is derived from the multiple-cutoff approach—a minimum level of each predictor attribute is necessary to perform the job
  2. After that level has been reached, more of one predictor attribute can compensate for less of another in predicting overall success—derived from the multiple-regression approach
- The combination method has the advantages of the multiple-cutoff strategy, but additionally provides a way to rank-order acceptable applicants
- A major disadvantage is that it is more costly than the multiple-hurdle approach—all applicants are screened on all predictors
- Most appropriate when the assumption of multiple cutoffs is reasonable and more of one predictor attribute can compensate for another above the minimum cutoffs
- More appropriate when the size of the applicant pool is not too large and costs do not vary greatly among procedures

### Strategy Six: Banding

- An alternative in dealing with the problem of selection procedure adverse impact is the use of *banding* of selection procedure scores—has received legal support
- Involves establishing ranges of selection procedure scores where applicants' scores within a range or band or treated the same —within a band, any score differences are seen as due to measurement error and thus equivalent
- Several banding procedures have been developed to take into account measurement error and incorporate that information in selection decisions
- For those scores within a band, other means or attributes are used to choose among applicants within a band

### *Establishing Bands*

- Two forms of statistical methods used for determining whether selection procedure scores differ—the standard error of measurement method and the standard error of differences method

### *Fixed and Sliding Bands* (Figure 15.2)

- *Fixed bands* use the top applicant score attained as the starting point
- *Sliding bands* also based on the top applicant's score, however once that applicant is selected the band is recalculated using the next highest applicant score – each decision based on those applicants still available

### *Selecting Within a Band*

- Once bands have been created, other selection specifications are used in choosing individual within the same band—examples include interpersonal skills, job experience, work performance, training, work habits, professional conduct

### *Advantages of Banding*

- Employer has more flexibility in making hiring decisions
- Employer takes into account those factors not measured by traditional selection methods

### *Concerns Regarding Banding*

- Debate as to whether using banding to identify equally qualified individuals and then choosing on the basis of minority status within the band will withstand legal challenge
- Any departure from top-down hiring based on applicants' ranked test scores leads to a loss of economic utility as reflect in work performance
- The value of banding in reducing adverse impact if preferential treatment cannot be given to minority group members
- May involve an incomplete system for weighting factors other than test scores

## **A PRACTICAL APPROACH TO MAKING SELECTION DECISIONS**

[slides 42–43]

- When choosing among selection decision-making procedures, consider:
  - a. Should the procedure be sequential or nonsequential?
  - b. Should the decision be compensatory or noncompensatory?
  - c. Should the decision be based on ranking applicants or on banding acceptable applicants?
- Begin by assessing the job and nature of work performance
- What determines success on the job?
- What factors contribute to success and how are they related?
- Having decided on a strategy, implement the strategy systemically in order to reap potential benefits
- Because of the sporadic nature of selection and the small number of applicants employed, it is often not possible to build objective strategies – especially true for small businesses
  - Build a systematic way of combining selection procedure data
  - *Bootstrapping* is based on the assumption that even though people can make sound judgments, they are not typically able to articulate how they made those judgments—through regression analysis, possible to infer weights used by the decision maker to arrive at a particular ranking
  - Procedural fairness—whichever method is used, critical that the procedures be perceived as fair

## AUDITING SELECTION DECISIONS: LEARNING FROM YOUR SUCCESSES AND YOUR FAILURES

[slide 44]

- The only way you can improve your selection decision making is to learn from what you have done in the past—both at the organizational level and the individual level
  - *Organizational level*—validation studies can help improve decision making by identifying scientifically those factors that predict job success and eliminating those factors that do not
  - *Individual level*—most managers do not think about their success and failure rate when making decisions
    - Several bad selection decisions can lead to lower productivity and high separation and replacement costs as well as to potentially damaging litigation
    - A simple box-score tally of successes and failure can improve decision making over time

## RECOMMENDATIONS FOR ENHANCING SELECTION DECISION MAKING

[slides 45–48]

1. Use standardized, reliable, and valid selection procedures for collecting information on job applicants whenever possible
2. Encourage decision makers to participate in the data-collection process but discourage them from combining scores on selection instruments or making decisions based on intuition
3. When combining scores, use a mechanical means for doing so—multiple regression, unit weighting
4. Train managers and others to make systematic decisions, preferably using one of the objective (mechanical) strategies described in the chapter
5. Although difficult for small organizations to adopt some decision-making strategies, they can specify (in advance)
  - a. The weight of standards to be used in evaluating candidates
  - b. The procedures used for judging whether applicants meet those standards
  - c. The procedures for combining the standardized weights multiplied by the ratings given in order to arrive at an overall applicant score
6. Decide whether a compensatory or noncompensatory method for selecting job applicants is going to be used
7. For organizations using cutoff scores, the modified Angoff procedure is acceptable
8. Assuming a selection procedure predicts work performance and that all applicants extended an offer of employment accept the offer, maximum work performance among the group hired will occur when top-down selection is used – however, disparate impact against racial minorities is likely to occur if the selection procedure is highly correlated with general mental ability
9. Banding of selection procedure scores has been supported in the courts, however using minority status alone for selecting within bands is probably not legal in most employment settings
10. For jobs in which selection has taken place, decide on a standard for defining a successful hire and an unsuccessful hire, then have managers keep track of their hits and misses

11. Periodically audit selection decisions throughout the organization to identify areas or individuals needing improvement

*These do not guarantee that you will always make correct decisions in selecting personnel for your organization but they are guaranteed to tilt the odds in your favor*