CHAPTER 2
Foundations of Recruitment and Selection I: Reliability and Validity

If Nothing Else, Students Should Learn That:

- personnel recruitment and selection strategies based on information obtained through scientific methods are more likely to benefit an organization than decisions based on impressions or intuition.
- the basic concepts of reliability and validity underlie contemporary recruitment and selection practices.
- different research methods and psychological measurement tools assist HR professionals in personnel selection.

Learning Outcomes:

- Understand the basic components that make up a traditional personnel selection model. [Understand]
- Explain the concepts of reliability and validity. [Understand]
- Recognize the importance and necessity of establishing the reliability and validity of measures used in personnel selection. [Apply]
- Identify common strategies that are used to provide evidence on the reliability and validity of measures used in personnel selection. [Remember, Understand]
- Discuss the requirement for measures used in personnel selection to evaluate applicants fairly and in an unbiased fashion. [Apply]
- Describe the practical steps needed to develop a legally defensible selection system. [Remember]

Key Concepts: Why Is This Chapter Important to HR Students?

- Although students may not be immediately excited about statistical topics, the concepts of reliability and validity are critical in recruitment and selection testing because they are scientifically derived.
- The accuracy of scientifically derived statements can be observed, critiqued, and used by others to predict which applicants will do well on the job. Such procedures are much more likely than gut feelings or intuition to produce results that meet legal requirements.
- Being familiar with measurement, reliability, and validity issues, and using only procedures that will withstand legal scrutiny, helps HR professionals ensure that their selection procedures meet acceptable professional standards.
- Students should know that they do not necessarily need to be statistical experts to be able to value the importance of reliable and valid instruments; there are practical steps in place that professionals can follow to develop a legally defensible recruitment and selection system.

Student Motivation: Why Should Students Care?

- Students need to realize that it is important to hire the best people based on the validity and reliability of the testing instruments used in the interview process, and that failing to do so may negatively impact organizational performance (and may also have negative legal consequences for the company).
- Students need to be aware that recruiters should ensure that applicants are judged in a fair and unbiased manner, not only for moral reasons, but also for legal reasons.

Engagement Strategies: What Can I Do in Class?

Classroom Discussion

- Ask the class to discuss the following ethical issue: Do individuals making staffing decisions have an ethical responsibility to know about measurement issues? Why or why not? (Responses will vary.)
- Ask the following questions: What is your perception of fairness? Does an organization have an obligation to make an enterprise as profitable as possible on behalf of its owners, or should it meet the objectives of society by providing equal employment opportunities for members of different population groups?

Group Discussion

- Short opening vignette: Ask students to review the "Sham Psychometric Test" controversy at the beginning of the chapter. Break up the class into groups of three or four and have each group discuss the consequences of using a test that may or may not be valid. Review the issues of validity and reliability and discuss how they relate to the hiring process. Each group should choose a leader to communicate the responses that were discussed in the small groups.

- In pairs or small groups, have students discuss a time when they experienced bias. Ask them to think of the first week of class and their first impressions of their teachers and/or peers. What biases were operating at that time? You can also discuss your first impressions of your students.
- In groups of three or four, ask students to discuss an ethical situation they have encountered at work. Have them describe the situation and how the situation was handled. When students share their experiences in a large group discussion, interesting discussions will follow.

Text Review

- Review **Figure 2.5, Validation Strategies**, and compare the different types of validation strategies. It is important to be able to distinguish among these strategies to make informed HR decisions about testing instruments.
- Review **Figure 2.6, An Example of Range Restriction**, and discuss.

Student Activities

- Direct students to **Figure 2.1, Job Analysis, Selection, and Criterion Measurements of Performance: A Systems Approach**, to review the steps involved in hiring a police constable. Ask the following questions: What are some of the major elements of the process? What questions do you have? Also ask students to review each of the steps in the selection process for the Toronto police in **Recruitment and Selection Today 2.1**.
- Provide students with three columns of made-up data in a spreadsheet: one for cognitive ability scores, one for emotional intelligence scores, and one for performance. Here is an example:

IQ	EQ	Performance
102.00	93.00	92.00
101.00	64.00	75.00
107.00	90.00	79.00
101.00	79.00	88.00
134.00	71.00	98.00
112.00	82.00	86.00
100.00	78.00	75.00
98.00	77.00	63.00
114.00	58.00	85.00
108.00	63.00	94.00
99.00	68.00	74.00
104.00	96.00	85.00
103.00	47.00	96.00
96.00	93.00	72.00
95.00	39.00	68.00
97.00	88.00	82.00
102.00	65.00	87.00
111.00	93.00	86.00
99.00	61.00	71.00
105.00	49.00	99.00

Using spreadsheet software (e.g., Excel) ask students to: 1. create a scatterplot of the data; 2. calculate a correlation between IQ scores and EQ scores; 3. calculate a correlation between IQ scores and performance; 4. calculate a correlation between EQ scores and performance; and 5. interpret the scores.

Assessment Tools: What Other Resources Are Available?

Other assessment tools include a Test Bank of multiple choice questions, true/false questions, short-answer questions; PowerPoint® slides; and MindTap, a fully online learning solution that combines all student learning tools—readings, multimedia, activities, and assessments—winto a single Learning Path that guides the student through the curriculum.

Reflections on Teaching: How Can I Assess My Own "Performance"?

Good teaching requires ongoing self-assessment and reflection. At the completion of this lesson, you may find it helpful to reflect on the following, and consider whether you want or need to make any adjustments for subsequent lessons.

- What worked in this lesson? What didn't?
- Were students engaged? Were they focused or did they go off on tangents?
- Did I take steps to adequately assess student learning?
- Did my assessments suggest that they understood the key concepts?
- What (if anything) should I do differently next time?
- How can I gather student feedback?
- How can I use this feedback for continuous improvement of my teaching?

SUGGESTED ANSWERS TO DISCUSSION QUESTIONS, EXERCISES, AND CASE STUDY

DISCUSSION QUESTIONS

1. **Discuss why it is better to base a selection system on science than on a "gut feeling."**
   Science produces information that is based on accepting as true only objective information that can withstand continued attempts to cast doubt on its accuracy. The accuracy of scientific statements is examined empirically through methods that can be observed, critiqued, and used by others, unlike gut feelings. Scientific information is dynamic and constantly evolving.

2. **Can an invalid selection test be reliable? Can an unreliable selection test be valid?**
   Yes, an invalid test can be reliable. For example, repeated results of a personality test are expected to be reliable; however, they are not a valid measure of cognitive ability. In terms of an unreliable test being valid, reliability of a measure places an upper limit on validity. Mathematically, the size of a validity coefficient cannot exceed the reliability of the measures used to obtain the data. Validity coefficients obtained from perfectly reliable measures of the predictor and criterion will be higher than those obtained with less-than-perfect measures. The decrease in magnitude of the validity coefficient associated with measurement error of the predictor, the criterion, or both is called *attenuation*.

3. **When should you use content, construct, or criterion-related validation strategies?**
   Validity is a unitary concept. Content, construct, and criterion-related validity are different but interrelated strategies commonly used to assess the accuracy of inferences based on measurements or tests used in the workplace. Sometimes these different strategies are mistakenly viewed as representing different types of validity. To overcome this misinterpretation, the older terms of construct validity, content validity, and criterion-related validity are no longer used in the measurement literature, although they are still sometimes used in assessing selection systems.

4. **Does an organization have an obligation to make the enterprise as profitable as possible on behalf of its owners, or does it have an obligation to meet the objectives of society by providing equal employment opportunities for members of different population groups?**
   There are no easy answers to this question. One resolution is to compare the fairness of the test in question with the fairness of an alternative that might be used in place of the test.

5. **What is your perception of fairness**
   The concept of fairness in measurement refers to the value judgments that people make about the decisions or outcomes that are based on measurements. An unbiased measure or test may still be viewed as being unfair either by society as a whole or by different groups within it. Issues of fairness cannot be determined statistically or empirically. Fairness involves perceptions. Fairness has no single meaning. There are three meanings of fairness that are relevant in selection: equitable treatment in the testing process; lack of bias; and fairness in selection and prediction.

6. **Do individuals making staffing decisions have an ethical responsibility to know about measurement issues? Why or why not?**
   Yes, it is important from business, ethical, and legal standpoints to have tests that are scientifically sound. It is also important to have procedures that are perceived as fair. From an ethical view, the perceived fairness of the testing procedures may negatively affect the unsuccessful candidates.

EXERCISES

1. **A marketing company is evaluating a new employment test that measures advertising aptitude of potential employees. You have used the new measure on a trial basis over the past year while you have continued to make selections using your established procedures. You have developed the following database, which includes information on gender, a score from the first administration of the test given during the selection process before the applicant was hired (Test 1), a score from a second administration of the test given at the end of the first year on the job (Test 2), and a performance score assigned by the supervisor at the end of the employee's first year of employment (Performance). You have been asked to evaluate the reliability, validity, and any gender bias of the new test. (Note: This exercise**

requires you to calculate correlation coefficients. If you do not know how to do that, your instructor will provide you with the coefficients.)

Employee	Gender	Test 1	Test 2	Performance
1	Male	24	18	20
2	Female	18	13	29
3	Male	21	17	17
4	Male	7	13	8
5	Female	14	28	25
6	Male	20	21	26
7	Male	8	6	7
8	Female	13	9	12
9	Male	15	13	18
10	Male	19	15	22
11	Female	25	22	23
12	Female	23	16	27
13	Male	18	13	10
14	Female	12	14	6
15	Female	17	12	13
16	Female	6	9	12

a. What is the reliability of the new test?
ANS: Test–retest: 0.55 (correlation between T1 and T2).

b. What is the predictive validity of the new test?
ANS: 0.67 (correlation between T1 and Performance).

c. What is the concurrent validity of the job?
ANS: 0.64 (correlation between T2 and Performance).

d. Is the test biased toward either males or females?
ANS: 0.16 (males = 1; females = 2, correlated with Performance). This suggests a slight positive bias toward females, but likely not a substantial one.

e. Would you recommend that the company adopt the new test as part of its hiring procedures?
ANS: Although the tests demonstrate good validity coefficients, their reliability is low, meaning the tests likely only have limited applicability. You may wish to recommend the adoption of a new test.

2. **Form a small group. Choose a specific job held by one of the people in your group. After discussing the job, choose one characteristic that you think is crucial to performing that job. How would you measure both the characteristic and job performance? Use Figures 2.1 and 2.5 to help you specify the conceptual and measurement levels. How would you establish the validity of your characteristic as a predictor of job performance?**

ANS:
Job: Salesperson at men's clothing store.
Characteristic: Agreeableness (measured by a self-report survey such as the NEO-PI-R).
Performance: As rated by the customer in a survey instrument (1 = low agreeableness, 5 = high agreeableness) or with organizational sales figures.

- To establish the predictive validity of the personality characteristic of agreeableness as an important component of performance, you will have to measure the personality of the candidates before they start the job, and then correlate it with sales performance later.

3. **Contact two organizations and identify the type of selection procedure that they use. You may be able to do this on the web or by calling their HR department. Use the criteria in Table 2.1 to determine whether each organization's selection system is analytical or intuitive. Determine the validity of the selection procedures in use by each organization; that is, has the company conducted validation studies and, if so, what was the outcome? Collect any validation reports that the organizations might have produced and bring the reports back to class for review and discussion.**

ANS:
Student answers will vary. The big question for students to ask, based on Table 2.1, is "Does your organization use any type of analytical measurement of performance and its predictors?" Note that even if the answer is "yes," organizations may still not be using science-based selection practices. Their use of science-based selection practices can only be determined by finding out the validity and reliability of the tests that are currently being used.

CASE STUDY

Emotional Intelligence or Cognitive Ability?

Case Questions and Suggested Answers:

1. **What do you think? Should the hiring managers prefer EI over cognitive ability in predicting job performance? Why?**

   ANS:
   No, hiring managers should not value EI over cognitive ability when it comes to predicting job performance. The results of validity generalization studies show that IQ is a better predictor of performance across most jobs compared with EI. Current measures of EI may be too broad to fully capture the content domain of the construct, which is still not very clear, even after decades of study. Scales can be improved via scale-validation strategies at both the construct and the measurement level (see **Figure 2.5**).

2. **If you planned to use EI as part of your selection system, discuss the steps that you would take to ensure that you were able to make reliable and accurate inferences about job performance in your work situation. That is, what would you have to do to show that your measure was reliable and valid?**

   ANS:
   There are a number of steps to take to ensure the reliability and validity of our chosen measure of EI. As an HR manager, you should start by checking the existing reliability and validity data of your measure in the existing literature (these indices are sometimes easily obtained by conducting a literature review on the topic/construct at hand). In the absence of such data, or if the data are unclear, you should ensure that factors impacting the potential reliability, including chance, lack of standardization, and temporary characteristics, are minimized at test time. In terms of ensuring validity, you may take strategies outlined in **Figure 2.5** and, for small organizations, **Recruitment and Selection Notebook 2.2**. To avoid range restriction, you should look at your validity coefficients to ensure that scores from subgroups are similar to those from the overall sample. To avoid measurement error, you should ensure that the test is reliable over at least two points of data collection. To avoid sampling error, you should ensure that the tests are taken by the appropriate samples (e.g., applicants and job incumbents) in as large a population as possible (estimates from small samples will likely be quite variable, so corrections may need to be applied). You should also take steps to ensure that your test is perceived to be fair and unbiased by your employees, as described in **Recruitment and Selection Notebook 2.3**, because adverse reactions to tests can negatively impact, or attenuate, validity scores.

3. **EI tests are prone to applicant faking. What can you do to limit faking on an EI test? Would these procedures lower the reliability and validity of the tests?**

   ANS:
   Current research suggests several ways to try to reduce faking in applicant testing. For example, there are statistical methods that can be used to try to detect and correct "faked" items; however, these corrections have been shown to possibly lower the validity of the tests. Participants can also be instructed to answer as honestly as possible because their data will have important selection and performance implications for the organization. It is worth noting that some researchers question whether faking on these tests is that much of a problem at all. They say that faked answers can be interpreted as participants "correctly" responding to social desirability cues by proving to recruiters that they know what the organization expects for demonstrating good performance.

4. **What are the legal considerations of using EI tests with poor reliability and validity? What are the business costs?**

ANS:

From a legal perspective, the perception of unfairness may lead unsuccessful applicants to pursue discrimination charges against the prospective employer in various legal arenas, including courts and human rights tribunals. Recruitment and selection tests should also not have adverse impact on minority groups. The business costs of using tests with poor reliability and validity are possibly hiring the people with the wrong KSAOs for the organization. This reduces the usefulness of the tests and potentially leads to decreases in organizational productivity levels. Further, applicant reactions to selection procedures may affect applicants' decision to join an organization and the degree to which they trust the organization and its behaviour once they become an employee.

