

PART ONE

Leadership Development and Selection

COPYRIGHTED MATERIAL
<http://www.pbookshop.com>

<http://www.pbookshop.com>

Best Practices in Leader Selection

Ann Howard

Getting the right leader in the top position stimulates organizations to prosper and grow. Chief executive officers (CEOs) account for 14 percent of the variance in organizational performance,¹ which means that there is a huge payoff if selection is done right. Moreover, it can cost millions of dollars if it's done wrong.

Unfortunately, there are a lot of CEO failures; estimates range from 30 percent to 50 percent.² A Booz Allen Hamilton study found that the rate of CEO dismissals in the world's 2,500 largest public companies increased by 170 percent from 1995 to 2003. Nearly one-third of the CEOs departing in 2003 (3 percent of a total of 9.5 percent) were fired for poor performance.³ Given these failure rates, it is not surprising that confidence in leaders is often shaky. In one national survey of public opinion based on 1,300 interviews, the average level of overall confidence in business leaders was 2.78 on a 4-point scale.⁴

The problem of poorly selected leaders could worsen as the Baby Boom generation retires, the supply of quality candidates dwindles, and the competition for talent heats up. Surveys have found that human

resource (HR) professionals anticipated greater difficulty filling leadership positions in the future. The higher the management level, the more difficulty expected: 66 percent of respondents expected more problems filling senior leadership positions compared to 52 percent for mid-level and 28 percent for first-level leader positions.⁵

There are multiple reasons why senior-level positions are so difficult to fill. The skill requirements for top-level jobs are high, as are the risks, evidenced by the excessive CEO failure rate. Detracting from the job are competitive pressures from a fast-moving global economy and elevated visibility and surveillance. CEOs and boards are now scrutinized intensely by shareholders, regulators, politicians, and the legal system, and their specific decisions are being second-guessed.⁶ At the same time the pool of qualified, well-prepared candidates for top-level jobs has shrunk with the evaporation of many preparatory mid-level positions and organizations' neglect of thoughtful succession planning.

This chapter describes how to get leader selection right. It reviews the objectives of selection, describes current selection techniques and evidence about their efficacy, and looks at how individual selection methods can be combined into an effective selection system. The chapter draws from general selection research and provides specifics for leaders where available.

OBJECTIVES OF LEADER SELECTION

Purposes of Selection

Although selection is usually thought of as hiring from the outside, internal selection (hiring from within) is just as prevalent for leaders. In addition to promotions, candidates are selected into positions or programs for career development and succession planning. Figure 1.1 shows some of the points at which leader selection occurs.

Recruitment of candidates varies by purpose and by management level. Entry-level leaders are usually a mix of outside hires and internal promotions. Organizations often place recruits from college campuses in first-level positions as an introduction to management roles. A classic pitfall of internal promotions is the selection of the best producer or technical performer, who is not necessarily the best manager. Such an ill-considered promotion leaves the organization with a mediocre leader and without a top performer.

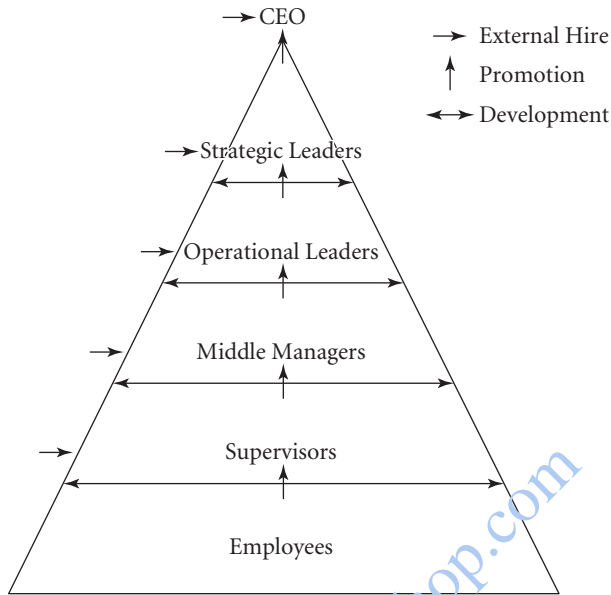


Figure 1.1. Sites for Leader Selection.

Middle managers are traditionally brought up from the lower-management ranks. Selection for career development can occur at any management level but succession management programs are usually aimed at higher levels. External hiring is common for a CEO, particularly if the organization is in trouble or is moving in a new direction. Outsiders run more than a third (37 percent) of the Fortune 1,000 companies, according to public affairs firm Burson-Marsteller, while insiders preside over the other two-thirds.⁷

Criteria for Selection Systems

Some might believe that the ultimate measure of a selection system's value is organizational effectiveness. However, such a criterion confuses performance or behavior with results. Organizational effectiveness is determined by multiple factors that are beyond the control of an individual leader. These factors can be internal, such as production delays or a labor dispute, or external, such as competition and market conditions. As noted in the introduction to this chapter, a leader, particularly at high levels, can have substantial impact on organizational

performance. However, organizational effectiveness is determined by more than a leader selection system.⁸

Criteria for measuring selection system success are of two types. The first concerns the output of the system, the most important of which is the individual performance of those selected. Organizations want the selection process to produce high-quality people who are well suited to their positions, will perform their required tasks well, and will remain motivated and committed. The system should also provide information about selected candidates that will prepare them and their managers for the growth and development that will inevitably be needed.

Additional criteria concern the nature of the selection system. It must be fair and appear fair to the candidates. It must work efficiently and remain viable over time. Each of these criteria warrants further exploration.

INDIVIDUAL PERFORMANCE. How well selected candidates perform in their new positions is the most important measure of selection system success. But there is more complexity in measuring the performance of leaders than that of individual contributors. A leader gets things accomplished through other people, so an important consideration is how leaders affect their work team and others in the organization. Thus satisfaction, retention, and performance of leaders' direct reports can add important data to the evaluation of leader quality.

There are three primary categories of things needed for success on a job.⁹ These include declarative knowledge (knowledge about facts or things; knowing what to do), procedural knowledge and skill (knowing how to perform a task), and motivation (whether to expend effort, how much effort to expend, and persistence in that effort). The first two components are often called "can do" factors, while the latter is called the "will do" factor.

Traditional research has focused more on the "can do" than the "will do." Yet high-quality hires will have little impact on organizational effectiveness unless they are motivated to stay with the organization long enough to make a difference. On average, managers stay in one organization 9.9 years,¹⁰ although this rate varies with economic conditions.

There is less research on the relationship between selection methods and attachment, whether measured as turnover, absences, or commitment. Factors other than the accuracy of selection come into play

with these outcomes. Common causes of turnover are personal reasons, such as getting married or returning to school, and undesirable behavior by one's manager. In fact, satisfaction has been equated to satisfaction with one's supervisor.¹¹ Research is sparse on selection methods and leader satisfaction, although this is an important precursor to retention.¹²

INFORMING INDIVIDUAL DEVELOPMENT. A chosen leader will seldom be perfect, and a sound selection system should also identify individuals' relative strengths and development needs. For example, a leader might be strong in business management skills like operational decision making or financial acumen but need development in interpersonal skills such as building strategic relationships. For internal selection, information about what characteristics need strengthening are an essential part of the process, not only for those who are selected, but also for those not selected who want to try again. The shoring-up process, for knowledge as well as skill development, can come in the form of training, coaching, or a critical assignment.

Many organizations also want an external hire's on-boarding process to include a development plan to work on needed skills and abilities. This requires the selection method to provide fodder for development—specific information that the new leader and his or her manager can follow to establish development steps. Jump-starting development could be an important factor in retention. When asked to choose the one most important reason employees leave, respondents most often cited a lack of growth and development opportunities (chosen by 25.3 percent). Only 8.3 percent chose a poor relationship with the manager.¹³ This suggests that employees endure a certain level of dissatisfaction with their managers as long as there are opportunities for growth.

LEGAL DEFENSIBILITY. Civil rights legislation and subsequent court cases have emphasized the importance of equal opportunity and the need for selection methods to be unbiased. Selection methods that produce adverse impact—defined as a selection rate for protected groups that is less than four-fifths (80 percent) of the rate for the highest group—must have clear evidence of job relevance and demonstrate that alternative methods are not feasible. This does not negate the use of methods with high adverse impact, but it makes them more subject to scrutiny.

CANDIDATE ACCEPTANCE. Selection is a two-way relationship, and there are consequences if a method affects candidates negatively. Candidates want to feel that their true skills, abilities, and potential are being evaluated and that they are being treated fairly relative to other candidates. Negative reactions are a particular concern to organizations because good candidates might withdraw from the competition and/or harbor negative feelings about the organization.

Unfortunately, research on candidate acceptance has seldom included leaders.¹⁴ While those at lower levels expect and accept a more high-tech, high-volume selection approach like online screening and testing, C-level (chief or highest-level executives) candidates often feel that they are above standard methods of testing or assessment and that their prior performance should speak for itself.¹⁵

In the past few decades, boards of directors often employed executive search firms to locate and screen new CEOs. The exact methods for selection were secret and probably idiosyncratic, but search firms commonly use unstructured interviews along with reference checks. As will be shown, these methods, though acceptable to candidates, are not very accurate, despite the fact that the top job has the highest consequences of any in an organization.¹⁶

The benefits to organizations of this type of selection are ripe for challenge. Although outsiders are increasingly sought to fill CEO positions, insiders have better performance records. The Booz Allen Hamilton study mentioned earlier found that the median shareholder return in 2003 among companies run by insiders was 3 percent compared to -5 percent for companies run by outsiders. Among North American CEOs who left their jobs in 2003, their boards forced 55 percent of outside hires and only 34 percent of insiders to resign. In Europe, 70 percent of departing outsider CEOs were dismissed, compared with 55 percent of insiders.¹⁷ This kind of evidence, combined with increased scrutiny of the practices of boards of directors, has laid the groundwork for acceptance of more sophisticated practices for selecting C-level leaders, including CEOs.

EFFICIENCY. Organizations should naturally favor selection methods that cost less and can be administered quickly and easily. However, HR professionals rarely track cost per hire.¹⁸ The level of investment in a selection system needs to be weighed against its potential payoff.

For many organizations the cost of selection may have less of a bearing on evaluations of efficiency than speed, particularly for external

hiring. It typically takes nearly 10 weeks to fill a management vacancy (compared to 6 weeks for nonmanagement staff), and 25 percent of the selection forecast HR professionals described hiring as slow or cumbersome.¹⁹

Leader selection in the future will likely be increasingly dependent on computer technology, which enhances not just efficiency but reach. Recruiting has already benefited from technological advances such as e-recruiting expanded pools of candidates, applicant tracking systems, online screening tools, and electronic job previews. Biographical data can also be collected with questionnaires or scored electronically from résumés. Tests and inventories are easily put in digital form and are increasingly delivered via the Internet.

Audio and video technology can deliver structured interviews with no apparent loss of reliability or additional adverse impact.²⁰ Assessment center simulations are also being automated. In-basket items can be delivered via e-mail, voice mail, or video on electronic desktops. These items can be supplemented by telephone or videoconference role plays. An advantage of online simulations is that communicating at a computer desktop better represents what modern leaders do.

Personnel Selection Paradigms

Selection works. Evidence accumulated through meta-analyses has shown that various selection methods have higher validity than might be expected.²¹ That is, across studies—once researchers removed errors of small samples, restricted range, and unreliability—statistical relationships between scores on selection methods and performance were usually strongly significant. Because there is large variance in leader performance, utility ratios based on almost any selection technique with modest validity can be justified.

A problem with the traditional paradigm is that its lack of a theoretical basis made it difficult to map predictors to performance constructs across different measures, contexts, and samples. For example, determining that cognitive ability tests predict leaders' job performance better than personality tests does little to advance the understanding of leadership. A new personnel selection paradigm, which has emerged in the past two decades, focuses on the nature of constructs and their interrelationships in order to enhance understanding and build practical applications. That is, different dimensions of job performance are related to variations in the validity of selection

methods across different contexts.²² For example, the trait of conscientiousness might be related to work standards in technically oriented leadership positions but not to the most important competencies for sales leadership.

The construct-oriented selection paradigm has led to various attempts to understand the multidimensionality of job performance. For example, John Campbell and his colleagues suggested eight general factors of performance across jobs (job-specific task proficiency, non-job-specific task proficiency, written and oral communication task proficiency, demonstrating effort, maintaining personal discipline, facilitating peer and team performance, supervision/leadership, and management/administration).²³

Studies have shown that decision-making and problem-solving competencies relate to one's early managerial performance, whereas interpersonal skills come into play several years later in the career.²⁴ Executives reviewing this evidence suggested that it takes more time to impact human systems than physical resources. That is, general managers can quickly diagnose and address problems or opportunities where raw materials or capital assets could enhance organizational performance, but it takes much longer to manage relationships with people or implement a new vision that affects trust or corporate culture.²⁵ Another explanation for the later impact of interpersonal skills might be visibility. Lower-level managers' problem solving and decisions will show up in productivity figures, but their interpersonal skills might not be evident to anyone but their direct reports. As managers move into higher-level positions, they interact with many more people and reveal their interpersonal skills to a wider audience. The greater stress of higher-level positions also might bring out underlying personality factors (such as arrogance) that become derailers in visible interpersonal situations.

Criteria of Effective Leadership

Before selecting leaders, organizations need to define what they expect them to do. These expectations can be stated in terms of personal competencies, often grouped into performance domains. For example, operational decision making is a competency in the business management domain whereas developing strategic relationships is in the interpersonal domain. Leadership is not only multidimensional, but is also moderated by various situational factors, such as management level, cultural context, and specific types of business challenges.

DOMAINS OF LEADER PERFORMANCE. Broad domains should include interpersonal and communication skills, leadership of others, administrative or business skills, and motivation or effort. Beneath these broad rubrics, however, is a long list of more specific competencies.

Competencies. To define the behavioral requirements for jobs, organizational practice has shifted dramatically away from job analysis, which identifies task details or activities that differentiate jobs (such as inspecting or investigating), to competency modeling, which identifies individual-level competencies required for groups of jobs (for example, decision making or influence). Assessment centers always had competencies (often called *dimensions*), so these constructs are not new.²⁶

One of the advantages of competency modeling is that it focuses on *how* work is accomplished (worker characteristics), whereas traditional job analysis concentrates on *what* is accomplished (job and task characteristics). Another advantage, particularly evident in recent years, is that competencies can provide a direct link to business goals and strategies. Competencies relate to behavioral repertoires—what people can perform and outcomes they can attain rather than tasks.²⁷ They are thus more appropriate for describing jobs that are changing.

Differences by Management Level. Leader requirements vary significantly by management level. There are pronounced differences among entry-level supervisors (leaders with one or possibly two direct reports), middle-level managers (leaders of leaders), operational leaders (those responsible for large business units), and strategic leaders (those who set organizational direction). The implication is that leader selection does not happen one time in a career. People are selected into the initial level of leadership, but even if they perform well there, there is no guarantee that they will be effective at a higher level of leadership. Thus, selection methods are usually reapplied at the major transition points shown in Figure 1.1.

Different abilities are needed for success at various levels.²⁸ For example, first-level leaders need skills in coaching, empowerment, and routine decision making. Mid-level leaders must make broader operational decisions and balance the needs of those above, beyond, and across from the subsystem they manage. Executives manage multiple units and have profit-and-loss responsibility. As leaders climb the management ladder, they are faced with challenges of increasing scope, complexity (scale), and ambiguity.

Cultural Differences. The extensive GLOBE (Global Leadership and Organizational Behavior Effectiveness) research found differences in the desirability of leader behaviors across cultures,²⁹ which complicates leader selection. However, multinational organizations often want a common model across geographical units. This is not out of the question, given that organizational cultures are often seen as having a stronger pull on behavioral styles than country cultures.

Another issue with cross-cultural applications is that there are country preferences for particular selection techniques. For example, assessment centers are popular in the United Kingdom, Germany, and the Netherlands but are seldom used in France.³⁰

Business Challenges. Global competition and environmental changes have focused management's attention on strategy. Organizations often want to know whether a leader is able to tackle specific business challenges, such as a turnaround, start-up, rapid growth, or strategic change.

Some competencies are more critical than others for managing different business challenges and thus rise in importance as measures of effective leadership. For example, a business trying to cultivate innovation might emphasize competencies such as change leadership, selling a vision, and establishing strategic direction as key criteria for sizing up its executives. Specific market segments may also need to address common business challenges. For example, a recent survey of hospital CEOs identified the three most critical leadership skills for organizational success over the next three years as strategic thinking, team building, and internal and external communication.³¹

INDIVIDUAL SELECTION TECHNIQUES

Selection methods can be arrayed across a continuum that ranges from signs of behavior (predispositions to act in a certain way, as from a personality test scale of extraversion) to samples of behavior (demonstrations of complex behavioral responses, such as coaching a direct report). Figure 1.2 provides examples of leader selection methods that take three positions along that continuum.

1. *Inferences* are made about how people will behave in leadership situations from their answers to tests (which have correct and incorrect answers), inventories of their personal qualities or

Signs → Samples

| Inferences about Behavior | Descriptions of Behavior | Demonstrations of Behavior |
|--|---|---|
| <ul style="list-style-type: none"> • Cognitive tests • Situational judgment tests • Personality inventories • Integrity tests • Leadership potential inventories • Motivational fit • Projective techniques | <ul style="list-style-type: none"> • Biographical data • Career achievement records • Reference checks • Interviews | <ul style="list-style-type: none"> • Administrative simulations (in-basket, planning, fact finding, analysis) • Interactive simulations (interaction role play, media interview, group discussion, business game) |

Figure 1.2. Leader Selection Techniques.

beliefs, or other techniques. These methods answer the question “Who am I?” For example, a test might identify a leader as “conscientious” or “smart.”

2. *Descriptions* of knowledge or experience are expressed in written or oral form. These include factual information about the candidates’ backgrounds as well as their perspectives on past or future behavior. These methods answer the questions “What have I done?” and “What do I know?” For example, a biographical data form might describe a candidate as experienced in hospital administration.
3. *Demonstrations* of leader behaviors are elicited from work samples and simulations. These methods answer the question “What can I do?” For example, a candidate to head a hospital might demonstrate in a simulated interaction with a physician (a role player) that he or she can gain the physician’s cooperation to save hospital costs.

Techniques that make inferences about behavior are usually closed-ended (multiple choice), while demonstrations of behavior are always open-ended (free response). Closed-ended tools lend themselves to computer scoring and are more efficient, while demonstrations of behavior provide the best information for individual development. Candidate acceptance is highest with job-relevant demonstrations of behavior and lowest with inference-making tools that are less well understood.

Inferences about Behavior

COGNITIVE TESTS. Tests of general mental ability (often called *g*) are very strong predictors of performance on jobs of all types, in large part by affecting the acquisition of job knowledge. Although *g* is derived from items measuring several specific abilities (such as verbal, numerical, or spatial), it represents a common factor that emerges regardless of specific content—a general property of the mind that reflects human differences in intellect.

Tests measuring *g* have their highest predictive validity for complex jobs. Positions of leadership, particularly high up in an organization, are unquestionably complex and are strongly predictive from cognitive tests. In practice, however, there is likely to be a restriction of range in *g* as leaders move up the management hierarchy.

While arguably industrial/organizational (I/O) psychology's most powerful tool, cognitive tests incur the largest adverse impact against minorities. Candidates do not react as favorably toward these tests as to interviews or work samples³² and may be particularly uneasy if they anticipate adverse impact. A practical problem with using cognitive tests for selection is that they measure capabilities that are not readily amenable to change.

SITUATIONAL JUDGMENT TESTS. These tests of decision making and judgment in work settings are primarily used at lower levels of management. Items typically describe a scenario and respondents identify the most appropriate response from a list of alternatives. Other versions of these tools do not present a situation but ask respondents to indicate their level of agreement with statements about the appropriateness of various work behaviors. These tests may not have incremental validity beyond cognitive ability tests.³³

PERSONALITY INVENTORIES. Personality inventories measure candidates' attitudes, motivations, and psychological character. They get at the "will do" aspect of individual performance. They also predict style of leadership: "Who we are determines how we lead."³⁴

There is extensive research on the clusters of personal traits known as the "Big 5": extraversion, emotional stability, agreeableness, conscientiousness, and openness to experience. Yet these measures have shown low validity for management jobs. Large personality domains are better at predicting global performance criteria than particular

performance areas.³⁵ Researchers have recently focused on mapping more specific personality traits to aspects of a job.

Personality inventories can also be used to forecast potential derailment of leaders. An old saw in executive search is “Hired on experience, fired on personality.”³⁶ The problem is that the dark side of personality can coexist with well-developed social skills, and potential derailers may lurk undetected. Stated more starkly, the bright side is the person you meet in an interview while the dark side is the one who comes to work.³⁷

Some have questioned whether personality traits should be linked to outcomes in a linear fashion.³⁸ For example, you can be too conscientious, conventional, and rule bound. Or take impulse expression: if you’re too high, you blurt things out; if too low, you’re fearful and rigid.³⁹ Research on leader derailment has shown that strengths taken to extremes can become weaknesses.⁴⁰

Another complexity in scoring personality tests is consideration of score profiles, which can be difficult to work with.⁴¹ In line with the new selection paradigm, combinations of personality test constructs need to be mapped to relevant aspects of performance. The more traditional approach has been to use an inventory that covers the major dimensions of personality and then determine empirically which dimensions are relevant.⁴²

Personality inventories are limited in that there is no direct translation into performance outcomes. Rather, you must translate personality into behavior and then into outcomes.⁴³ Personality doesn’t create business results; behavior does.

Another problem with personality questionnaires is that applicant acceptance is lower than that for interviews, work samples, or even cognitive tests.⁴⁴ Some personality tests reject high scorers as fakers, raising potential legal issues. Sex differences can also bring charges of adverse impact. Added to this brew is the understanding that personality can’t readily be changed: it has been speculated that traits like emotional stability and extraversion might have neuropsychological roots.⁴⁵ However, people who understand the nature of their personality can take steps to mitigate the biggest problems it might cause them.⁴⁶

Faking is another potential problem with personality tests. Although some studies show that this effect is overblown,⁴⁷ there is still a concern that it is easy to distort these instruments if you are motivated to do so.

INTEGRITY TESTS. One type of personality instrument that has come into vogue, particularly in light of the many recent corporate scandals, is the integrity test. Integrity tests use facets of conscientiousness and emotional stability in their construction. Parallel surveys over time showed more than double the use of integrity tests between 1999 (7.9 percent of respondents) and 2004 (16.4 percent of respondents).⁴⁸

LEADERSHIP POTENTIAL INVENTORIES. Some inventories directly measure leader characteristics and potential. Older tests based on the global factors of consideration and initiating structure included the Leadership Opinion Questionnaire, Leader Behavior Description Questionnaire, and Supervisory Behavior Description Questionnaire. However, these tests have no established validity.⁴⁹ There is considerably more evidence to support transformational and transactional leadership theory, particularly from the Bass Multifactor Leadership Questionnaire, or MLQ. Transformational leadership appears to be more related to follower satisfaction than performance of the leader or group.⁵⁰ Although some relationships with performance have been shown, it is not yet clear whether transformational leadership measures can aid top-level selection.⁵¹

MOTIVATIONAL FIT. Motivational fit measures get at the “will do” factor by addressing whether the candidate will be personally satisfied with the work (job fit), the company culture (organization fit), or the company’s location (location fit). Inventories allow candidates to compare their preferences and values with the nature of the job and organization. Although this approach is seldom used to reject candidates, it can help them to screen themselves out.

Motivational fit can also be evaluated with interview questions. The interviewer asks the interviewee how he or she would react to certain job characteristics. For example, if a leadership position required working with diverse groups of people, an interviewer might ask a candidate, “Tell me about a time when you worked with people with a wide range of backgrounds or perspectives. How satisfied were you with that and why?”

PROJECTIVE TECHNIQUES. Projective techniques present ambiguous stimuli and ask people to fill in what’s missing. The theory behind projectives is that personality structure influences the way individuals interpret their environment.

A fundamental problem with all projective measures is that the level of inference is so great that they require a trained scorer. This constraint poses a challenge for selection system efficiency.

Descriptions of Behavior

Descriptions of behavior can be oriented toward the past (what you did) or the future (what you would do). The premise behind describing what you have done is behavior consistency: past behavior is the best predictor of future behavior. The premise behind describing what you would do in the future is that behavioral intentions predict future behavior.

BIODATA. Biographical information, or biodata, quantifies descriptions of past activities and accomplishments, such as degrees earned or specific work experiences. It can be collected in various ways, such as checklists of skills and experiences, forms requesting background information, and computerized résumé screening.

Several types of research suggest that these measures relate to leader effectiveness. For example, a review of seven studies found that personal history correlated .38 with success in management.⁵² Other studies have explored how personal experiences relate to leader effectiveness. For example, in AT&T's Management Progress Study, college major and extracurricular activities predicted interpersonal skills in an assessment center.⁵³ Biographical data lends itself to automated screening, as when applicants fill in forms on the Internet that are then scored against the organization's criteria.

CAREER ACHIEVEMENT RECORDS. Career achievement records capture key experiences that demonstrate effective performance of competencies. They are usually completed by applicants and reviewed for accuracy by their managers. They require descriptions of actual behavior, organized by important competencies.

Career achievement records have several advantages, particularly in restructuring (and often downsizing) situations when internal candidates are applying for positions in an organization pointing in a new strategic direction. Candidates are given a voice in the selection process, which helps them perceive it as fair. These measures often replace poorly constructed performance data for experienced candidates who are being screened for transfer or promotion. Their meta-analytic predictive validity with overall job performance is .45.⁵⁴

On the downside, career achievement records can be time consuming and costly to construct and score. Because the process seldom includes follow-up questions, the documented behaviors can be difficult to verify.

REFERENCE CHECKS. Despite their small but statistically significant correlation with job performance, problems with lawsuits in the United States have made many organizations reluctant to give out very meaningful data. The popularity of reference checks also varies in other countries. For example, the British use them a lot (74 percent of companies surveyed), but the French seldom do (11 percent).⁵⁵

INTERVIEWS. One clear research result is that unstructured interviews don't work nearly as well as structured. One estimate, based on reliability analyses, of the upper limits of validity was .67 for highly structured interviews and .34 for unstructured.⁵⁶

Perhaps because of their prevalence, interviews are the focus of the majority of legal court challenges to selection.⁵⁷ Another disadvantage of interviews is that they are time consuming and thus not very efficient compared to other selection methods. Moreover, interviewers need training in conducting and evaluating interviews, particularly those that are structured around competencies. Video or audio delivery of structured interviews offers hope for enhancing interview efficiency.⁵⁸

One controversy concerns the relative effectiveness of interviews structured around past experiences (behavior description interviews, or BDI) or hypothetical situations (situational interviews, or SI). Both types are usually competency based. However, BDI ask about what you did ("Describe a time when you weren't sure what a customer wanted. How did you handle the situation?"), and SI ask about what you would do (A situation is described where it's unclear what a customer wanted. "What would you do?").

There is sufficient empirical data to support both SI and BDI as predictors of job performance, although three studies found the BDI more valid for higher-level jobs.⁵⁹ It was speculated that the addition of descriptively anchored rating scales to the BDI significantly enhanced its validity. Not only do such scales enhance reliability, but they also improve construct validity by more clearly specifying the performance domain.⁶⁰

Situational interviews have an advantage when organizations want to directly compare the responses of candidates for future jobs. But

they are measures of maximal performance, whereas BDI are measures of typical performance.⁶¹

Another controversy is whether interviews lose validity if conducted by telephone rather than face to face. Some data showed lower relationships to later job performance for phone-based interviews. Regardless of delivery mode, interviews have been criticized for measuring social skills, experience, and job knowledge rather than the constructs or competencies they are purported to measure.⁶²

Demonstrations of Behavior

Candidates being evaluated for jobs as individual contributors are sometimes asked to provide work samples that show what they can do. The validity of work samples should be enhanced to the extent that there is point-to-point correspondence between predictor elements and criterion elements. Their predictive validity with overall job performance is quite high (.54),⁶³ and they add significantly to prediction from cognitive ability tests.

For complex leadership roles, candidates are usually asked to demonstrate their capabilities in several diverse simulations of key managerial roles. Managerial simulations range in complexity from brief mini-simulations conducted with interviews to lengthy and involved analyses used for executive assessments. Factors that determine the complexity of a simulation include the length of preparation time needed, richness of detail provided, difficulty level of the issues, and number of competencies covered. Leader simulations are of many types but usually fall into two broad categories, administrative and interactive.

ADMINISTRATIVE SIMULATIONS. These simulations involve individual problem solving.

In-Basket. These exercises have participants review information similar to that found in a manager's in-basket. Participants respond to correspondence or voice mail messages involving, for example, productivity, morale, public relations, and training needs.

Planning. In a planning or scheduling exercise, participants are given resources with which to accomplish specific tasks. Leader candidates might be asked to schedule others' activities or schedule people to jobs.

Fact-Finding Exercise. The participant is given a brief description of a hypothetical situation. The task is to seek information from a resource person and make a decision about that situation within a limited time period.

Analysis. In these exercises participants analyze quantitative and narrative data and make short- and/or long-term recommendations to improve matters such as productivity, quality, profitability, organizational structure, and morale.

INTERACTIVE SIMULATIONS. In these simulations candidates interact with role players or participate in groups.

Interaction Role Plays. Participants review background information on a peer, internal or external customer, direct report, or prospective client. They then meet with a trained role player to resolve a problem or gain commitment to a course of action. Role plays are often structured to bring out a primary competency such as coaching, influencing others, or managing conflict.

Media Interview. Typically confined to executive assessments, these exercises provide participants with background information about a situation that has attracted media attention. Candidates then meet with a media representative (a trained role player) to answer probing questions and defend their organization's actions.

Presentation or Visionary Address. Participants make formal presentations or an inspirational talk on some selected problem or topic to one or more role players or external observers.

Group Discussion. A group of candidates reviews background material on a situation or set of issues, and then meets to discuss the issues and decide how to resolve them. Group discussions vary in that they may or may not assign roles, appoint leaders, or inject competition into the discussion.

Business Game. In a business game, a group is assigned a problem and members must work together to solve it. Business games usually have a competitive element and require some organization of effort and division of labor among team members.

VALUE OF SIMULATIONS. Simulations are traditionally associated with assessment centers, and the majority of evidence about their validity comes from this context. The assessment center itself has a long track record of predictive validity,⁶⁴ although research evidence for individual types of exercises is sparse. Significant relationships with performance criteria, including advancement into higher levels of management, have been found for such simulations as leaderless group discussions, in-basket exercises, and business games.⁶⁵

Simulations offer organizations several advantages. Because participants demonstrate live behavior, there is little opportunity to fake and no need to speculate about what behavior would be like. Simulations can be directed at future jobs or challenges, which gives them an advantage over current descriptions of performance. Applicants usually react positively, believing the technique to be fair and job related. Adverse impact is less for administrative simulations than for cognitive tests and negligible for exercises measuring interpersonal skills.⁶⁶

A significant advantage of simulations over other selection methods is their usefulness for development. They are mostly aimed at skills and abilities that can be learned, and demonstrations of live behavior in simulations provide excellent material for credible feedback.

But simulations are labor intensive, adding to time and costs. In recent years Web delivery has cut down considerably on administrative burdens, but simulations don't lend themselves to automatic scoring such as that for tests and inventories. Moreover, interrater reliability must be continually checked and reinforced, as assessors apply holistic judgment to scoring. Simulations don't readily measure all leadership competencies, in particular those that require actions that extend over long periods of time, such as customer networking.

Questionable Methods

Two popular but questionable methods should also be mentioned. For very different reasons, these techniques are *not* recommended for leader selection.

MULTIRATER SURVEYS. Multirater surveys (also known as *multisource* or *360-degree surveys*) attempt to capture the behavioral observations of those close to leaders. Respondents typically include direct reports,

peers, and supervisors, although customers or suppliers can also be tapped for their opinions. These surveys can be very useful for development. Leaders can learn from others about their reputations and consider how to optimize their favorable qualities and change undesirable aspects.

If used for selection, however, multirater surveys can become vulnerable to gaming the system or sources of work stress. Raters might have different motives for responding and undermine trust in a work group.⁶⁷

GRAPHOLOGY. Analyses of handwriting have no established validity. Data supporting the method came from essays, where content was evaluated and presumably influenced ratings. Graphology also incurs very negative applicant reactions. Nevertheless, the method is widely used in France and Israel.⁶⁸

SELECTION SYSTEMS

Although proponents of different selection techniques make various claims about their superiority, a single method is usually inadequate to meet organizational objectives. Thus, the best practice is to capitalize on the advantages and mitigate the disadvantages of any one method by combining it with others into a selection system.

The type of job affects how comprehensive the system can reasonably be. Practitioners can put together a more complete system for leaders, where the candidates are fewer and the stakes higher, than for lower-level positions, where the volume of candidates is high and the stakes lower.

Is Combining Selection Techniques Worthwhile?

Combining selection techniques is considered worthwhile if it can reduce the adverse impact associated with a single method. Multiple tools are also valuable if they raise the level of accuracy over one tool alone.

REDUCING ADVERSE IMPACT. The strong adverse impact of cognitive tests has led practitioners toward using them with personality tests, hoping to enhance the overall proportion of acceptable minority candidates. Generally this has not worked as well as hoped.⁶⁹ Combinations that

include cognitive ability almost never meet the four-fifths standard for selection rates described earlier.⁷⁰

Even noncognitive composites—groups of tests that emphasize personality traits, values, attitudes, or other human characteristics that don't depend on intellectual ability—can show adverse impact, especially if the selection ratio (proportion of candidates selected) is low (50 percent or fewer are chosen). A better approach to reducing adverse impact might be to vary forms of administration. Video-based tests can remove unnecessary variance due to reading comprehension.⁷¹ For example, a video-based situational judgment test showed less adverse impact than a written version.⁷²

INCREMENTAL VALIDITY. Even if adverse impact is not substantially reduced, it is clear that leadership requires more than *g*. Whereas cognitive ability affects thinking and knowledge, personality factors relate to interpersonal behavior and communication.⁷³ Using both methods together should produce incremental validity, or a stronger relationship with leader performance than using one method alone.

Cognitive and personality measures are fairly distinct, but if two selection methods are highly correlated, there is no advantage to adding the second. Where the lines between methods are murky, incremental validity research can be useful. For example, there is some evidence that personality inventories add incremental validity to assessment centers in the prediction of performance.⁷⁴ On the other hand, biodata can be oriented toward personality or *g* factors and may not add value beyond what is measured elsewhere. More complete studies of incremental validity are needed to sort out the relative contribution of different measures in order to construct a powerful yet efficient selection system.

The correlation of cognitive ability with employment interview evaluations (.40 in a meta-analysis) suggested that the interview might not add much to a selection system in which *g* was being measured by a test.⁷⁵ However, the correlation was less pronounced for high-complexity jobs, such as manager, perhaps because of a narrower range of test scores or better impression management skills. The correlation between cognitive ability and interview ratings was also smaller with more interview structure, perhaps because structured interviews are better at assessing other constructs that also relate to job performance.

Techniques for Combining Selection Methods

There are several common techniques for combining information from various selection methods into a decision about a candidate.

MECHANICAL OR JUDGMENTAL COMBINATION. Although previous research showed that statistical formulas were better for combining scores on various measures than clinical judgment,⁷⁶ these “mechanical” methods are seldom used in practice. One problem is the need to regularly cross-validate weighted formulas. Most often, practitioners take measures that are scored both clinically and statistically and combine them judgmentally.⁷⁷ This holistic method is based on the premise that the whole is greater than sum of the parts, that human judgment can discern patterns of behavior that should influence final decisions.

MULTIPLE HURDLES. Multiple hurdles are common in practice but receive little research attention. The idea behind multiple hurdles is to first screen out candidates with more efficient tools. In other words, use an inexpensive inventory to wash out those clearly unsuited to the position and reduce the number of candidates eligible for more expensive, labor-intensive methods, such as interviews or assessment centers. Figure 1.3 gives an example of what a selection funnel with multiple hurdles might look like.

After the important criteria have been established, a computer can screen candidates for relevant background and experience with a scorable biodata instrument. Tests and inventories can yield more information about fit to the position. Some or all of these results might be revealed to the candidate (for example, motivational fit results) to permit both a realistic preview of the job and self-selection. Remaining candidates can be put through simulations of the target job, and those that survive this process can be interviewed by hiring managers. An on-boarding process, including development plans, is launched as the candidate is hired.

An important consideration with multiple hurdles is where to set cutoff scores at each stage. A high bar might leave too few candidates. Moreover, quality candidates might be lost if a less valid predictor is used as an initial hurdle. A hurdle too low won't offer much gain in efficiency.

Another issue is the utility of later tools. To the extent that initial and later tools are correlated, there may be a restriction of range in

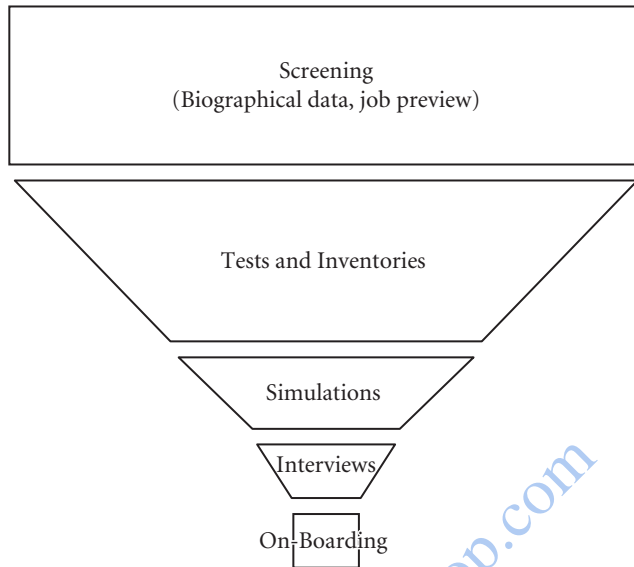


Figure 1.3. Multiple Hurdles.

the later methods that constrains their validity. Another concern is that information from an earlier screen might not be brought forward and integrated into the final evaluation of the candidate.

INDIVIDUAL ASSESSMENT. In individual assessment, one evaluator, usually a clinical psychologist, interprets results of an array of selection tools and makes a recommendation about a candidate's suitability. The method is frequently used for higher management, often in connection with succession planning.

Common selection tools include personal history information, ability tests, personality and interest inventories, and interviews.⁷⁸ Direct observations of behavior, as through simulations, are seldom used and have been one source of criticism of the practice.⁷⁹ Clinical psychologists are often unfamiliar with simulations and lack access to them. The approach to combining information from different tools is clinical and holistic.

One problem with individual assessment is that there are no standards or training to guide the practice, not even a systematic job analysis. Like simulations, job analysis is part of the standard

repertoire of industrial-organizational psychologists but not clinical psychologists. There is little research to support the efficacy of individual assessment, but small samples and other factors make such research difficult.⁸⁰

ASSESSMENT CENTERS. Assessment centers use information from a variety of selection tools to rate competencies and form an overall judgment about leadership potential. Best practices for assessment centers have been documented in the *Guidelines and Ethical Considerations for Assessment Center Operations*.⁸¹ The hallmarks of assessment centers are simulations, which are required for the process to be called an assessment center. Assessment centers require multiple assessors and an integration of data in order to avoid the biases of individual raters.

An assessment center can be a stage in a multiple hurdle. A decision-making committee brings together competency ratings from the assessment center and information from the other hurdles and makes a go/no-go decision. Alternatively, the entire selection system could be implemented as an assessment center. That is, the decision-making body uses information from all selection tools, including the assessment center simulations, and makes final ratings of competencies before deciding whether to accept the candidate. This model positions an assessment center within an assessment center.⁸²

Assessment centers originally included all sorts of selection methods, including projective tests, personality inventories, cognitive tests, and so forth.⁸³ However, when put into operational use with managers as assessors, the more psychological measures were dropped and the method became primarily dependent on simulations. Now that there is a return to professional assessors, personality tests are sometimes brought in to help interpret the behavior observed in the simulations in terms of personal traits.

A great deal of research has demonstrated the criterion-related validity of overall assessment ratings (OAR) with career progress, potential ratings, training, and performance. Estimates of the relationship between assessment center ratings and success in management range from .31 to .43,⁸⁴ with an upper bound of .63 under optimal conditions.⁸⁵

Assessment centers share the advantages reviewed about simulations, but as a system have other benefits as well. Particularly when

presented as an integrated “Day in the Life” format, assessment centers form a realistic job preview for candidates, which can guide those who are not a good fit to opt out of the process. This is especially important at the first level of leadership, where candidates might not realize what leadership entails. It can also be a revelation to senior-level candidates who don’t appreciate the difference between operational and strategic leadership.

Assessment center simulations are fundamentally plastic; that is, they can be designed to represent various organizational challenges. This is problematic for research, as wide variations in exercises make generalization more difficult. The method allows for personal treatment as administrators and role players interact with candidates—an appealing characteristic for aspiring senior managers. Adverse impact is minimal, and most candidates accept assessment centers as job relevant and fair.

MAKING LEADER SELECTION SUCCESSFUL

The next time you are asked to recommend the best system for selecting leaders, you won’t go wrong with one or both of the following answers: “It depends,” and “It’s complicated.” The variety of selection techniques, each with its advantages and disadvantages, is complicated enough. Add to that the myriad of circumstances that can influence leader effectiveness—organizational and company cultures, management level, purposes of selection, nature of candidate pools, market and competitive conditions of organizations, and so forth—and it’s clear that there is no one best system for leader selection. Practitioners need to be well informed and strategic to develop the best system for their organization’s particular circumstances.

Best Practices and Common Practices

Despite the accumulation of solid research support for the accuracy of various selection tools, it is dismaying to find that organizations still prefer unstructured interviews.⁸⁶ Even larger and more sophisticated organizations often fail to capitalize on the best techniques available. This suggests that the benefits of sound selection methods—and potential pitfalls of using poor ones—have not been adequately

communicated. In Development Dimensions International's selection forecast, 40 percent or more of HR professionals never used testing or assessment methods.⁸⁷ A survey of British organizations showed similar a similar pattern: between 30 percent and 41 percent never used tests or assessment. However, French organizations were considerably less scientific than the British, shunning cognitive tests, biodata, and assessment centers for more intuitive and interpretative approaches like graphology and personality measures.⁸⁸

Various psychologist researchers have lamented the gap between research and practice in personnel selection. Cropanzano pointed to the "justice dilemma," whereby practices that raise validity lower perceptions of fairness, while practices that raise perceptions of fairness may lower validity.⁸⁹ Employers are sensitive to the acceptability of selection instruments and thus gravitate away from cognitive tests and toward unstructured interviews. Anderson and colleagues argued that professionals should aim for methods with both high practical relevance (useful to organizations) and high methodological rigor (well grounded in research), an ideal they call "Pragmatic Science." In reality, there is much work in the academic literature that is rigorous but not very useful (Pedantic Science) and too many methods that organizations use that are not grounded in rigorous research (Populist Science).⁹⁰

Automation and Web delivery have inadvertently aided Populist Science. Though enabling rich enhancements to selection systems, the Web has also provided access to a seemingly endless supply of poorly developed selection devices. The Internet is flooded with snake oil. Personality inventories are particularly vulnerable to homemade nonsense. Never has it been more important for organizations to look for well-designed and researched tools from reputable practitioners.

Clearly there is much room for improvement in hiring practices. But there are signs of progress. Between 1984 and 1989, British organizations tripled or more the usage (in at least half of their selection processes) of cognitive tests, personality inventories, biodata, and assessment centers.⁹¹ Between 1999 and 2004, across an international sample, there was an increase of 40 percent or more in the proportion sometimes or extensively using integrity tests, computer-assisted interviews, and biodata.⁹² Another positive sign is the growing use of assessment centers for succession planning and their emergence for CEO selection.

Realizing Selection System Objectives

Fulfilling the promise of best practices in leader selection must go beyond popularizing more valid selection tools in organizations. Attention must be paid to meeting *all* the criteria for selection systems laid out at the beginning of this chapter. As Cropanzano suggested, there are “win-win” integrative solutions, such as work samples and assessment centers, that are both seen as fair and have high validity.⁹³

Where feasible, connections should be made between sound selection practices and important organizational outcomes like profits or turnover.⁹⁴ But there is need for caution in how these are presented. Utility formulas can estimate dollar returns from hiring better leaders, but when these are delivered as a “hard sell,” they can actually decrease management’s intentions to use a selection system.⁹⁵

Even a well-constructed selection system can still fail to accomplish its objectives unless attention is paid to how it is introduced to and embedded within the organization. To realize the selection system’s objectives, practitioners should attend to five components of execution.⁹⁶

1. *Communication.* Make a business case for investing in a selection system that will meet your criteria. The organization must be willing to commit the necessary resources, from time and money to manpower, including the attention of senior leaders.
2. *Accountability.* Complex selection systems require the commitment of many people across the organization. A senior executive should champion the initiative and hold others accountable for performing the roles that will lead to its successful execution.
3. *Skill Development.* Stakeholders using new tools and interviewing methods must be equipped with the knowledge and skills to use them. And leaders must have the skills to manage the operational changes that the new system will bring.
4. *Alignment.* The new selection system can’t realize its potential if other organizational systems dilute or undermine what it needs to be successful. It is particularly important to align selection with related systems and processes like performance management, leader development, succession management, and retention initiatives.

5. *Measurement.* To demonstrate the effectiveness of the system, it's important to measure what's happening. Measures pertaining to both individual performance and the nature of the selection system should be collected at the time of selection (lead measures) and after enough passage of time that the outcome of selection can become evident (lag measures).

EXECUTIVE SUMMARY

Selection of leaders is too important to leave to serendipity or chance. Ineffective leaders, particularly at the top, can be extremely costly to an organization, and unfortunately leaders fail all too often. Getting leader selection right can not only be a boon to organizational performance, but also give people an opportunity to excel in work they enjoy. The following six steps will help your organization realize the advantages of a leader selection system.

1. *Clarify your purpose.* Is the goal of your selection system external hiring, promotion, career development, succession planning, or some combination? What levels of leadership are you trying to fill? Your overarching purpose should guide expectations for your selection system and how they can best be met.

2. *Set selection system criteria.* Some of the criteria by which to judge the success of your selection system should focus on the outputs of the system. Most important will be the individual performance of selected candidates in their new positions. While organizational performance should also be affected, this is a less-than-perfect criterion because an individual leader controls only some of the factors that go into organizational effectiveness. Another important system output might be information that a leader can use to continue to grow and develop.

Other selection system criteria concern the nature of the system itself. These include (a) legal defensibility: if the selection system selects out a disproportionate number of members of protected groups (as cognitive tests do), you must be able to justify the practice; (b) candidate acceptance: if candidates don't believe their true skills and ability are being evaluated fairly (as often happens with graphology), their negative feelings might generalize to the organization; and (c) efficiency: if it takes too long to fill vacancies, particularly with external candidates, organizations won't look favorably on the selection system.

3. *Define leader success.* The characteristics of effective leaders must be defined in advance. Typically, these are competencies related to different aspects of interpersonal skills, leadership, and business management with additional attention to motivation. The particular configuration of competencies a leader needs depends on the key business challenges the leader must confront, the criteria management level of the position, and the culture in which the leader is embedded.

4. *Choose selection techniques.* A single selection technique is unlikely to accomplish all of your goals or provide all the information you want, so be prepared to include multiple methods in your system. There will be many valid selection techniques at your disposal, each of which will have advantages and drawbacks that need to be balanced.

Selection methods vary along a continuum from signs to samples of behavior. At the “signs” end of the continuum are tools like cognitive tests and personality inventories from which one can make *inferences about behavior* in leadership situations. These measures most easily meet the efficiency criterion, but they often lack candidate acceptance and have limited usefulness for guiding development. While cognitive tests are especially strong in terms of predicting individual performance, they incur the most adverse impact among selection measures.

A second type of selection method draws out *descriptions of behavior*, knowledge, and experience. These methods include biographical data, career achievement records, and interviews. Structure improves the validity of these methods for predicting job performance. While they can provide important information about work experience, the extent to which these methods are independent measures of key leadership constructs is still unclear.

The third category of selection methods, on the “samples” end of the continuum, represents *demonstrations of behavior*. These methods include diverse work samples and simulations, most often used in assessment centers, whose predictive validity is well established. Simulations have the added advantages of being able to address future jobs, provide information on trainable behaviors that are useful for developmental feedback, and engender positive reactions from candidates. However, simulations are labor intensive and inappropriate for competencies that roll out over extended periods of time, such as networking.

5. *Combine tools into a selection system.* The next step is deciding how to put your preferred individual techniques into a selection

system. How comprehensive you need to be will depend on the volume of candidates you expect, the potential impact of the leader on your organization's effectiveness, and available resources, including the time to fill positions, budget, and executives' and hiring managers' time and commitment. However, comprehensiveness needs to be weighed against incremental validity and usefulness. More tools are not necessarily better if you can forecast leader performance just as well with a smaller set.

Selection system efficiency can be enhanced with multiple hurdles. In other words, use efficient, computer-scored tools to eliminate lower-end candidates and reserve labor-intensive, expensive methods like interviews and assessment centers to differentiate among a smaller pool of more promising candidates.

6. *Execute your plan.* Last but not least, consider how you will introduce the system into your organization and assure its continued success. Important activities include communicating the business case for the system, assigning accountability for its execution, developing the skills of those who will carry it out, aligning other systems, and measuring the lead and lag indicators that will tell you if the system is meeting its objectives.

In case this all looks too complicated or tiresome, remember that effective leader selection can mean millions of dollars for your organization. Selection system failure has the same financial implications; they just go in the opposite direction.

Selection matters. Leader selection matters much more.