Leadership Assessment

Robert Hogan   Hogan Assessment Systems
Robert Tett   Universtity of Tulsa

Leadership Assessment

1. Introduction. Common sense suggests that leadership is the most important topic in the social, behavioral, and organizational sciences. A trip through the business section of any bookstore will also reveal that it is the most popular—based on the number of books written on the topic (well over 7000). The literature on leadership falls into two discreet categories. The first, and by far the largest, contains books designed for the popular or mass market. This vast literature contains nuggets of wisdom and flashes of insights from entrepreneurs, military officers, historians, business school professors, and consultants; however, it is not systematic, empirical, or verifiable, and it lacks an assessment base. In short, it is not a foundation on which to build a reliable understanding of leadership; it is entertaining rather than deeply informative.

   The second literature, and by far the smaller of the two, comes from the empirical tradition of Academia. This tradition has the critically important characteristic of adhering to the standards of intellectual accountability that normally prevail in empirical research—publicly available data, standardized analytical methods, peer review, etc. By definition, then, the empirical tradition is more informative about leadership, at least in principle, than the vast collection of opinions contained in the literature designed for mass markets. But the empirical tradition suffers from four problems that limit its utility. First, it typically defines leadership in terms of the persons who happen to be in charge of the organizational unit being studied. But many, if not most, people who are in charge of organizational units attain their status for political reasons rather than because they have demonstrated significant leadership. In addition, by defining leadership operationally, the empirical tradition avoids the question of what leadership really is. Second, by defining leadership operationally, the empirical literature doesn't converge—because the characteristics of persons in charge of one organizational unit are typically different from the characteristics of persons in charge of a different unit, and it is nearly impossible to compare leadership (defined in this way) across organizations. Third, the empirical tradition has been strongly influenced by behaviorism, and as a result largely ignores the relationship between personality and leadership—from a behaviorist perspective, circumstances are more important than personality as an influence on leadership. And fourth, the empirical tradition has largely ignored the links between leadership and organizational effectiveness— arguing that effectiveness is too hard to define.

   We believe that some attention to these definitional and conceptual issues will substantially clarify the assessment of leadership.

   A. Leadership Defined. Data and common sense indicate that people are naturally selfish; nonetheless, all significant human achievement is the result of collective effort. Consequently, leadership should be defined in terms of the ability to persuade people to set aside

their personal concerns and support a larger agenda—at least for a while. Leadership differs from management—subordinates respond to leaders because they want to, they respond to managers because they are paid to. Good managers, nonetheless, are often able leaders and vice versa.

B. Leadership and Organizational Effectiveness. All real world groups compete with other groups for desired resources—money, land, energy, food, loyal supporters, official patronage, league championships. Because leadership is a resource for a group, affecting its ability to attain its goals, leadership should be evaluated in terms of the effectiveness of the leader's team: Did it win the prize, the goal, the race, or the war? Sometimes a good leader's team loses because it is overmatched, and sometimes a bad leader's team wins because it has superior resources; nonetheless, leadership and organizational effectiveness are codependent.

C. Leadership and Personality. To clarify the links between leadership and personality, we need first to define personality. Personality has two definitions, they are quite distinct, and they concern personality from the inside and from the outside. Personality from the inside, which we call identity, is composed of a person's values, goals, aspirations, and self-image; identity can be assessed by asking a person about his/her goals, aspirations, and self-image. Leadership assessment from the inside consists of assessing the identities of leaders and comparing them with the identities of non-leaders.

Leadership from the outside, which we call reputation, is composed of the images and evaluations of a person, as held by those people with whom that person interacts. Leadership reputation is assessed using observer ratings from whatever source a researcher might prefer. This would include assessment center ratings and 360 feedback evaluations. Relevant research questions here include whom to ask to provide ratings, what rating dimensions should be used, and how leaders differ from non-leaders in terms of these ratings. The bottom line of this discussion is that leadership needs to be assessed from the inside and from the outside.

2. Assessing Leadership from the Inside. As noted above, personality from the inside concerns a person's values, motives, and self-image—identity. Various methods and instruments are available to assess leadership from the inside and many of them yield valid results. For example, Sparks (1990) reports on research conducted at EXXON in the 1950s, designed to identify managerial talent. Using a combination of measures of cognitive ability, personality, biodata, and interviews, in a sample of 443 managers, and a composite success criterion (salary, level, and rated effectiveness), Sparks reports a cross-validated multiple R of .70. Howard and Bray (1990) report roughly comparable results using multiple methods in a longitudinal study at AT&T. Thus, it is possible to predict leadership performance using measures of identity, and in many cases with an admirable degree of validity.

The most robust procedures for assessing leadership from the inside fall into four

categories: (A) projective measures of personality; (B) objective measures of personality; (C) specialized measures of personality; and (D) mixed measures of personality.

A. Projective Measures. McClelland (1975) and his associates have used the Thematic Apperception Test (TAT) to study leadership in a wide variety of organizations and countries. They report reasonable validity coefficients with scores for "socialized power" (desiring power in order to bring about organizational change, not self aggrandizement) and leadership performance.

B. Objective Measures. Objective measures of personality with demonstrated validity for predicting leadership fall naturally into four sub-groups. The best-known exemplars of the first group are the 16 PF (Cattell & Eber, 1961), and the Guilford-Zimmerman Temperament Survey (Guilford & Zimmerman, 1956). These inventories were constructed using internal consistency indices and factor analysis; the goal was to define reproducible factors, and predictive validity was a side issue. The second group contains objective measures of personality developed in an empirical manner and designed to maximize validity. The best-known example of this approach is the California Psychological Inventory (CPI; Gough, 1987); .the well-respected CPI is widely used in management consulting around the world. The third group of measures concerns values and occupational interests. The best known of these is the Myers-Briggs Type Indicator (MBTI: McCaulley, 1990); the MBTI is not well regarded by many psychometricians, but McCaulley provides clear data that a pattern of MBTI scores typifies executives world wide. A second important measure in this category is the Campbell Work Orientations Surveys (Campbell, 1990); Campbell shows that his inventory of values, interests, and preferences predicts a wide variety of leadership outcomes. The final category of objective personality measures contains inventories based on the Five-Factor Model (FFM: Wiggins, 1996). A substantial body of research shows that: (a) most existing measures of personality can be reconfigured in terms of the FFM, and (b) measures of normal personality based on the FFM are also robust predictors of leadership (cf. Hogan & Hogan, 1995). The FFM is the new paradigm for measures of normal personality, although there is strong resistance to this notion among some personality researchers (e.g., Block, 1995).

C. Specialized Measures of Personality. There are far too many specialized measures of personality, used to predict leadership, to cover responsibly here. There are thousands of individual scales the best known of which concern authoritarianism, machiavellianism, self-monitoring, and dogmatism; these measures predict specific aspects of leader performance. Recent developments in theory and research suggest that these specific scales can be placed in the larger context of the FFM. Two recent special measures are important here—Emotional Intelligence (EQ) and Transformational leadership. The measurement base for the EQ movement (Goleman, 1995) is ad hoc and not well regarded by measurement experts. Nonetheless, a

mounting body of evidence indicates that the ad hoc dimensions of EQ (self-awareness, self-management, social awareness, and social skill, as identified by Goleman, can be reliably measured and predict leadership performance fairly well.

The Transformational leadership movement begins with Sigmund Freud and Max Weber, who argued that successful leaders have charisma, the ability to attract and develop a following. Robert House (1977) turned Weber's list of charismatic traits into a rating scale, and showed that the scale predicted leadership performance. House's results, combined with Burns' (1978) book, created a surge of interest in charismatic leadership, now termed neo-charismatic or transformational leadership, around which a substantial body of empiricism has developed, much of it based on a measure called the Multifactor Leadership Questionnaire (MLQ; Bass & Avolio, 1991). The results of this movement can be summarized in terms of six points. First, there is considerable consensus regarding the components of transformational leadership—the key components include integrity, conviction, commitment, vision, optimism, openness to new ideas, and consideration of and concern for subordinates. Transformational leadership is contrasted with transactional leadership, which emphasizes goals, accountability, performance management, and compensation. Second, these characteristics—which are desired by subordinates regardless of cultural context—resemble the components of EQ. Third, transformational leadership as a gestalt suggests that: (a) there is a moral component to leadership; (b) leadership depends on the ability to develop a relationship with subordinates; and (c) there is one best way to behave as a leader. Fourth, considerable data support these claims. Fifth, transformational leadership (and EQ) are clearly related to personality. Consequently, sixth, transformational leadership seems to be a syndrome of normal personality and should be captured by components of the FFM.

D. Mixed Measures of Personality. Several leadership assessment procedures combine measures of leadership from the inside with measures of leadership from the outside, and form a bridge to the next section of this essay. The two best known and thoroughly validated of these mixed measures are the Managerial Practices Survey (MPQ; Yukl, Wall, & Lepsinger, 1990), and the Leadership Practices Inventory (Posner & Kouzes, 1990). The MPQ is better regarded by academic researchers; it asks a manager to evaluate him/herself in terms of 11 categories of managerial behavior; then subordinates evaluate the manager using the same categories. The Leadership Practices Inventory is based on a critical incidents survey of experienced managers ("What did you do when you were at your best?"); the dimensions of this assessment line up with the model of transformational leadership discussed in the preceding section.

3. Assessing Leadership from the Outside. Subjective ratings of others' reputation as leaders began the 1950s using: (a) on-the-job behavioral description; and (b) assessment centers. Researchers at Ohio State University analyzed 1,800 behavioral descriptors and developed the Leader Behavior Description Questionnaire (LBDQ; Stogdill & Coons, 1957) and the

Supervisory Behavior Description Questionnaire (SBDQ; Fleishman, 1957) which contained two broad dimensions: Consideration and Initiating Structure. Researchers at the University of Michigan created scales of leader Support, Goal Emphasis, Work Facilitation, and Interaction Facilitation (Bowers & Seashore, 1966). More recent on-the-job leadership rating scales are psychometrically improved, broader in scope, and use multiple rating sources. For example, *The Profilor* (Hezlett, Ronnkvist, Holt, & Hazucha, 1997) assesses 38 managerial and leadership competencies using multisource ratings. A growing number of 360-degree measures of leader and managerial behavior are available for use in developmental feedback (London & Smither, 1995).

The second major source of leadership ratings from the outside is the assessment center. Originally developed in World War II to select officers and spies (Murray & Mackinnon, 1946), modern assessment centers use job simulations (e.g., in-basket exercise, leaderless group discussion) to evaluate leadership potential. The advantage of assessment centers over on-the-job ratings include control of situational factors and better assessor training and accountability. Assessment centers are valid predictors of leadership (Gaugler, Rosenthal, Thornton, & Bentson, 1987), but there are many questions about what they measure (e.g., Bycio, Alvares, & Hahn, 1987; Sackett & Dreher, 1982). Specifically, measures of such themes as planning and organizing, and interpersonal skills correlate higher within exercises than individually between exercises. Nonetheless, assessment centers are a major source of data on leader and managerial behavior.

4. Major Findings and Research Challenges. Leadership assessment is intimately linked to leadership research. Key findings include the following.

Leadership is multidimensional. Leadership has often been conceptualized in terms of dichotomies, for example, Consideration and Initiating Structure, participative and autocratic styles, and person- and task-orientation. But leadership is a more nuanced concept, composed of an array of narrower facets. More specific leadership assessment allows (a) greater precision in developmental feedback and matching people to jobs; (b) richer conceptual frameworks for comparing alternative leadership perspectives, and (c) stronger tests of nomological networks guiding validation efforts (Tett, Guterman, Bleier, & Murphy, 2000). A multidimensional approach also requires lengthy testing time for high fidelity data, and poses logistic problems in multisource feedback systems (Graddick & Lane, 1998).

Leadership is an evolving construct. Leadership assessment tools may need to be updated or replaced to reflect changes in the meaning of leadership as research progresses over time.

Leadership means different things to different people. Agreement among peer, subordinate, and supervisor ratings in multisource systems is moderate at best on most dimensions (Conway & Huffcutt, 1997; Dalessio, 1998; Harris & Schaubroeck, 1988). Validation efforts need to be

sensitive to the diverse and sometimes competing values others hold about a given leader's role (Butterfield & Bartol, 1977), and to differences between raters in opportunity to observe leader performance.

Leadership assessment is a cognitive process and needs to be treated as such (Brown & Lord, 2001). Person perception holds promise as a framework for studying leadership reputation (Lord, De Vader, & Alliger, 1986; Mount & Scullen, 2001; Sessa, 2001). Greater attention needs to be given to the mental processes by which individuals form judgments of themselves and others as leaders (Church, 1997).

Leadership effectiveness depends on the context. Skills required in higher level leadership roles typically are different from those required at lower levels (Silzer, 1998). Leadership succession planning requires identifying the demands expected in future leadership roles.

Leader promotion can entail an assessment paradox. Despite the greater importance of leadership at higher levels, assessing senior leaders poses unique challenges due to greater ambiguity of the leader's role, increased political use of appraisal results, and reduced accountability for not undergoing appraisal (Gioia & Longenecker, 1994; Graddick & Lane, 1998; Longenecker & Gioia, 1992).

Responsible leadership assessment requires commitment from upper management that results will be used for stated purposes. Using developmental feedback data for promotion decisions can undermine assessment goals. Special efforts are needed to ensure adherence to stated purposes and confidentiality of results (Silzer, 1998).

References

Bass, B. M., & Avolio, B. J. (1991) *The multifactor leadership questionnaire*. Palo Alto, CA: Consulting Psychologists Press.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117,* 187-215.

Bowers, D., & Seashore, S. (1966). Predicting organizational effectiveness with a four-factor theory of leadership. *Administrative Science, 11*, 238-263.

Brown, D. J., & Lord, R. G. (2001). Leadership and perceiver cognition: Moving beyond first order constructs. In M. London (Ed.). *How people evaluate others in organizations.* Mahwah, NJ: Lawrence Erlbaum Associates.

Burns, J.M. (1978). *Leadership*. New York: Harper & Row.

Butterfield, D. A., & Bartol, K. M. (1977). Evaluators of leader behavior: A missing element in leadership theory. In J. G. Hunt & L. L. Larson (Eds.), *Leadership: The cutting edge*. Carbondale, IL: Southern Illinois University Press.

Bycio, P., Alvares, K. M., & Hahn, J. (1987). Situational specificity in assessment center ratings: A confirmatory factor analysis. *Journal of Applied Psychology, 72*, 463-474.

Campbell, D.P. (1990). The Campbell work orientation surveys. In K. E. Clark & M. B. Clark (Eds.) *Measures of leadership*. West Orange, NJ. Leadership Library of America.

Cattell, R.B. & Eber, H. W. (1961). *The sixteen personality factor questionnaire* (3rd ed.). Champaign, IL: IPAT.

Church, A. H. (1997). Do you see what I see? An exploration of congruence in ratings from multiple perspectives. *Journal of Applied Social Psychology, 27,* 983-1020.

Conway, J. M., & Huffcutt, A. I. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance, 10,* 331-360.

Dalessio, A. T. (1998). using multisource feedback for employee development and personnel decisions. In J. W. Smither (Ed.). *Performance appraisal: State of the art in practice.* San Francisco: Jossey-Bass.

Fleishman, E. A. (1957). A leader behavior description for industry. In R. M. Stogdill & A. E. Coons (Eds.), Leader behavior: Its description and measurement. Columbus, OH: Ohio State University.

Gaugler, B. B., Rosenthal, D. B., Thornton III, G. C., & Bentson, C. (1987). Meta-analysis of assessment center validity [Monograph]. *Journal of Applied Psychology, 72*, 493-511.

Gioia, D. A., & Longenecker, C. O. (1994). The politics of the executive appraisal. *Organizational Dynamics*, 47-57.

Goleman, D. (1995). *Emotional intelligence*. New York: Bantam.

Gough, H.G. (1987). *California Psychological Inventory* administrator's guide. Palo Alto, CA: Consulting Psychologists Press.

Graddick, M. M., & Lane, P. (1998). Evaluating executive performance. In J. W. Smither (Ed.). *Performance appraisal: State of the art in practice.* San Francisco: Jossey-Bass.

Guilford, J.P., & Zimmerman, W.S. (1956). Fourteen dimensions of temperament. *Psychological Monographs*, 70, no. 417.

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology, 41*, 43-62.

Hezlett, S. A., Ronnkvist, A. M., Holt, K. E., & Hazucha, J. F. (1997). *The Profilor technical summary.* Minneapolis, MN: Personnel Decisions, Inc.

Hogan, R., & Hogan, J. (1995). *Hogan Personality Inventory Manual.* Tulsa, OK: Hogan Assessment Systems.

House, R. J. (1977). A 1976 theoryu of charismatic leadership. In J.G. Hunt & L.L. Larson (Eds.), *Leadership.* (Pp. 189-207). Carbondale, IL: Southern Illinois University Press.

Howard, A. & Bray, D. W. (1990). Prediction of managerial success over long periods of time. In K. E. Clark & M. B. Clark (Eds.). *Measures of leadership.* West Orange, NJ: Leadership Library of America.

London, M., & Smither, J. (1995). Can multi-source feedback change perceptions of goal accomplishment, self evaluations, and performance related outcomes? Theory-based applications and directions for research. *Personnel Psychology, 48,* 803-839.

Longenecker, C. O., & Gioia, D. A. (1992). The executive appraisal paradox. *Academy of Management Executive, 6*, 18-28.

Lord, R. G., De Vader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between personaliyu traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology, 71*, 402-410.

McCaulley, M. H. (1990). The Myers-Briggs Type Indicator and leadership. In K. E. Clark & M.B. Clark (Eds.) *Measures of leadership.* West Orange, NJ: Leadership Library of America.

McClelland, D. C. (1975). *Power.* New York: Irvington.

Mount, M. K., & Scullen, S. E. (2001). Multisource feedback ratings: What do they really measure? In M. London (Ed.). *How people evaluate others in organizations.* Mahwah, NJ: Lawrence Erlbaum Associates.

Murray, H. A., & Mackinnon, D. W. (1946). Assessment of OSS personnel. *Journal of Consulting Psychology, 10,* 76-80.

Posner, B. Z., & Kouzes, J. M. (1990) Leadership practices. In K. E. Clark & M.B. Clark (Eds.) *Measures of leadership.* West Orange, NJ: Leadership Library of America.

Sackett, P. R., & Dreher, G. F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology, 67*, 401-410.

Sessa, V. I. (2001). Executive promotion and selection. In M. London (Ed.). *How people evaluate others in organizations.* Mahwah, NJ: Lawrence Erlbaum Associates.

Silzer, R. (1998). Shaping organizational leadership: The ripple effect of assessment. In R. Jeanneret & R. Silzer (Eds.). *Individual psychological assessment*. San Francisco: Jossey-Bass.

Sparks, C.P. (1990). Testing for management potential. In K.E. Clark & M.B. Clark (Eds.). *Measures of leadership*. West Orange, NJ: Leadership Library of America.

Stogdill, R. M., & Coons, A. E. (1957). *Leader behavior: Its description and measurement.* Columbus, OH: Ohio State University.

Tett, R. P., Guterman, H. A., Bleier, A., & Murphy, P. J. (2000). Development and content validation of a "hyperdimensional" taxonomy of managerial competence, *Human Performance*, *13*, 205-251.

Wiggins, J. S. (1996). *The five-factor model*. New York: Guilford.

Yukl, G., Wall, S., & Lepsinger, R. (1990). Preliminary report on validation of the Managerial Practices Survey. In K. E. Clark (Ed.), *Measures of leadership.* West Orange, NJ: Leadership Library of America.