Evidence-based leadership development: A case study on 360-degree feedback.

Kai Externbrink[1] & Ilke Inceoglu[2]

[1] Fresenius University of Applied Sciences, Cologne; [2] Surrey Business School, University of Surrey, Guildford

ABSTRACT

We argue that effective leadership development should be evidence-based, i.e. that it combines the best available scientific evidence with research in the specific organizational context. To illustrate our proposition, we report findings from a case study in a multinational organization. The goal was to examine which rater source in the company's 360 degree feedback would provide the most valid information about leadership competencies. Therefore, we explored relationships between 360 degree ratings and assessment center (AC) ratings of the same leadership competencies (N=151). It was predicted that AC ratings show higher overlap with 360 degree ratings for behaviors that specific rating sources can more easily observe in the ratees' work life. Results showed that peers were the most accurate observers of leadership competencies in 360 degree assessments, compared to managers and subordinates. This corroborates our argument for an evidence-based instead of an intuitive handling of 360 degree feedback results. Practical implications and avenues for future research are discussed.

Keywords: *Evidence-based Management, Case-Study, 360 degree Feedback, Assessment Center, Leadership Competencies*

1   Introduction

Developing future leaders is one of the most important tasks of today's human resource management. It is a strategic need and crucial to build organizations' dynamic capabilities for competitive advantage (Teece, 2009). When searching for suitable interventions, human resource practitioners are, however, often confronted with "dangerous half-truths and total nonsense" (Pfeffer & Sutton, 2006). Unsurprisingly, the many leadership development programs in organizations are influenced by implicit theories, personal experiences or simply the latest trend (Walshe & Rundall, 2001).

Evidence-based Management (EBM) on the contrary, refers to organizational practices that are informed by the best available scientific evidence (Rousseau, 2006; Briner & Rousseau, 2011). Evidence means reliable, objective and valid information on the effectiveness of organizational practices generated through an explicit and systematic research process (Externbrink & Dormann, 2014). Decision makers can either refer to evidence reported in scientific journals, or to evidence that is based on action research in their own organization (Reason & Bradbury, 2001): The organization is explored systematically to inform initial action. The results are then analyzed and evaluated to inform follow-up action and so forth. In doingso, so-called "best practices" can be applied in due consideration of a given organizational context. Accordingly, Briner, Denyer and Rousseau (2009) suggest an integration of research evidence and local information for managerial decision making. EBM leads to better results, since lower evidence-orientation can result in inappropriate decisions and thus impair organizational performance

(Schmidt & Hunter, 1998; Terpestra & Rozell, 1993; Siddique, 2004).

In this paper we report findings from a case study in a multinational organization, where we applied evidence-based practice to answer the question which rater source in the company's 360-degree feedback would provide the most valid information about leadership competencies. Although many intuitive answers to this question may exist, less scientific evidence is available. This is partly due to the fact that predictive validity of 360-degree ratings depends on the organizational context. Against this background, our case-study contributes to the literature on EBM and 360-degree feedback in the following way: It demonstrates how consultants and HR managers can apply 360-degree feedback more effectively by systematically using local research evidence that shows how well these tools work in the organization and how results may be influenced by the local context.

2   360-degree Feedback and Evicende-based Practice

Many principles of EBM apply to 360-degree feedback. If well-conducted it is a reliable, objective and valid feedback process, which describes how subordinates, peers, and managers perceive the focal manager on a wide range of job-related competencies (Seifert & Yukl, 2010). Feedback allows individuals to adjust the level and direction of their effort as well as performance strategies to match competency requirements of their role (Locke & Latham, 2002). There is ample evidence for the value of 360-degree feedback as an integrated part of leadership development programs: for example, the positive performance impact of feedback is shown in the meta-analysis

of Stajkovic and Luthans (2003). Similarly, the longitudinal study of Walker and Smither (2003) over a period of five years showed significant increases in managerial performance in response to upward feedback programs. The meta-analysis of Smither, London and Reilly (2005) over 24 longitudinal studies implies that performance improvement following 360-degree feedback is more likely, when results are used to derive personal development goals. Consequently, the use of competency-based 360-degree feedback in combination with development planning, training and individual coaching has advanced to a state-of-the-art human resource practice in leadership development.

In practice it can be quite difficult, however, to decide on development goals and plans following 360-degree ratings on specific competencies, given the low agreement between raters (e.g. Borman, 1997). Not all rater sources are equally predictive of actual job performance; Beehr, Ivanitskaya, Hansen, Erofeev and Gudanowski (2001) for example reported positive relationships for competency-based 360-degree ratings and performance appraisals only for manager and peer ratings. Research has shown that such patterns depend on the source and content of the feedback and especially on the specific characteristics of the recipient's organization (Brett & Atwater, 2001; DeNisi & Kluger, 2000; Warr & Bourne, 2000; Yukl & Seifert, 2005). It is therefore not clear whether all competency ratings should be averaged across rating sources or more weight should be placed on ratings by particular rating sources. In practice systematic approaches are often lacking and average ratings, for example, are applied out of convenience.

Given these challenges, an evidence-based strategy to make optimal use of 360-degree feedback, is to gather further evidence within the specific organizational context. The objective of this study is therefore to examine relationships between 360-degree ratings and assessment center (AC) ratings of the same leadership competencies in a global Fortune 500 company. Competencies are defined here as "sets of behaviors that are instrumental in the delivery of desired results or outcomes" (Bartram, Robertson & Callinan, 2002, p. 7).

ACs can be defined as "a variety of testing techniques designed to allow candidates to demonstrate, under standardized conditions, the skills and abilities that are most essential for success in a given job" (Coleman, 1987, p.3). Candidates go through a series of standardized tests, interviews and simulation exercises. Such job simulations are usually based on job-analyses and critical incidents in order to design content valid situations that are realistic and representative for the managerial role in the specific organization. Behaviors assessed in ACs are observed (e.g. group exercises) or measured through other means (e.g. in-basket exercises) by trained assessors who are not familiar with the assessed individual. Overlap between 360-degree and AC ratings are therefore expected for behaviors that specific rating sources can more easily observe in the ratees' work life.

3 Method

The study was conducted as part of a leadership development program for junior and senior managers. The objective of this program was to develop relevant leadership competencies (as defined by the company's top executives) for career progression using 360-degree feedback (competencies rated by managers themselves, their managers, peers and direct reports) and ACs, measuring the same competencies as in the multisource feedback. Competencies relevant for the junior manager role were Analyzing, Deciding & Initiating Action, Adapting & Responding to Change and Working with People. For senior managers these were Formulating Strategies & Concepts, Enterprising & Performing, Leading & Supervising and Making an Impact. The process of defining relevant competencies together with top executives was based on the meta-analytically validated universal competency framework (Bartram, 2005).

Data was collected with 151 international managers from an organization in the fast moving consumer goods sector who participated in an internal leadership program. Ninety of these were junior managers, who were in their first managerial role, and 61 were senior managers, leading a whole function.

4 Sample

In the sample of junior managers 53% were male. Twenty-three percent was below the age of 30. Thirty-three percent was between 31 and 40 years old while one percent was over the age of 41. For 43 percent of junior managers information on age was missing. Seventy-five junior managers were from Europe, seven were from Asia, three from Australia, two from South America and three from North America.

In the senior manager sample 71% were male and 73 % were between 30 and 40 years old. Fifteen percent were between 41 and 45 years old, while two percent were older than 46 years. For ten percent of the senior managers information on age was missing. Forty-five participants were based in Europe, while eight were from Asia, six from North America, and two from Australia.

5 AC rating process

All observers received training prior to the AC. Teams of three observers (one assessment consultant and two internal top executives) rated each manager. Each assessor assessed at least two AC exercises. The junior manager AC consisted of a case study and presentation exercise, a role-play and a group discussion. Senior managers underwent the same exercises (specific to the senior manager role) but instead of a group discussion they took part in a competency-based interview. Competency ratings were agreed through a judgmental process for each exercise based on standardized rating sheets, using a 5 point-rating scale ranging from "poor" to "excellent". Observers rotated after each exercise and the overall rating was agreed for each competency in a final observer conference. As part of the assessment process, all managers also completed an online verbal and a numerical reasoning test (SHL Online Verbal Test & Online Numerical Test,

ONT & OVT: SHL, 1998) as well as a personality measure (Occupational Personality Questionnaire, OPQ32i: SHL, 2006). Results of the cognitive ability and personality measures were integrated into the final competency rating. The AC matrices with mappings of exercises to constructs are given in the appendix.

6   360-degree feedback

Each competency was rated (self, manager, peers and direct reports) using four items each. Means, standard deviations and Cronbach's alpha of the competencies as rated in the 360-degree assessment and in the AC are provided in Tables 1 and 2.

Raters were asked: "How well do the following statements describe the person you are rating?". A five-point rating scale was applied ("not at all" to "extremely well"). For each manager about five peers and three direct reports provided ratings in addition to manager- and self-

assessment. Ratings were averaged for each rater source and analyzed separately.

7   Results

Data was analyzed separately for junior and senior managers. Results showed that AC ratings correlated positively with 360-degree ratings for the same competency but only if rated by peers (Tables 3 and 4). This pattern was especially consistent for the sample of senior managers indicating that peers are a valid source for observing various leadership competencies. In the sample of junior managers AC Ratings of Deciding & Initiating Action correlated significantly with the corresponding 360-degree rating of their manager, showing that managers of junior managers are more likely to observe behaviors such as showing initiative and decision making skills of their direct reports. Self-assessment of competencies did not correlate significantly with their corresponding AC rating, except for Formulating Strategies & Concepts in the sample of senior managers.

Table 1: Overview of assessed competencies, descriptive statistics and reliabilities for junior managers

Competencies	Self			Manager			Peer			Direct Report			AC	
	M	SD	α	M	SD	α	M	SD	α	M	SD	α	M	SD
Analyzing	4.08	.41	.57	4.14	.46	.72	4.06	.38	.85	4.08	.53	.88	3.29	.85
Deciding & Initiating Action	3.81	.44	.58	4.05	.46	.64	3.87	.36	.60	3.85	.56	.80	3.03	.67
Adapting & Responding to Change	3.88	.46	.60	4.94	.68	.85	3.65	.51	.84	3.78	.64	.93	2.88	.74
Working with People	3.94	.49	.60	4.03	.60	.85	3.88	.44	.84	3.88	.56	.93	2.84	.77

Table 2: Overview of assessed competencies, descriptive statistics and reliabilities for senior managers

Competencies	Self			Manager			Peer			Direct Report			AC	
	M	SD	α	M	SD	α	M	SD	α	M	SD	α	M	SD
Formulating Strategies & Concepts	4.09	.54	.74	4.13	.47	.75	4.03	.35	.60	4.00	.46	.82	3.26	.70
Enterprising & Performing	4.05	.50	.98	4.25	.51	.68	4.03	.36	.60	4.11	.54	.84	3.38	.61
Leading & Supervising	4.07	.42	.47	3.89	.88	.82	3.46	.88	.85	3.92	.82	.96	3.23	.57
Making an Impact	3.99	.46	.62	4.22	.47	.70	4.03	.36	.75	3.91	.52	.83	3.10	.70

Table 3: Correlations between AC ratings with 360-degree ratings of the same competencies for junior managers

Competencies	Correlations			
	AC - Self	AC - Manager	AC - Peer	AC - Direct Report
Analyzing	.15	.18	.22*	-.06
Deciding & Initiating Action	.17	.24*	.01	-.06
Adapting & Responding to Change	.12	.05	.21*	-.14
Working with People	.08	.01	.27**	.14

*Note: *p < .05, **p < .01, N = 90*

Table 4: Correlations between AC ratings with 360-degree ratings of the same competencies for senior managers

Competencies	Correlations			
	AC - Self	AC - Manager	AC - Peer	AC - Direct Report
Formulating Strategies & Concepts	.43**	.14	.37**	.10
Enterprising & Performing	.02	.13	.32*	.12
Leading & Supervising	.02	.15	.40**	.01
Making an Impact	.03	-.20	.29*	.01

*Note: *p <.05; **p <.01; N = 61*

Table 5: Correlations between ability scores and 360-degree ratings of analytical thinking for junior managers

Ability Measures	Descriptives		Correlations with 360-degree Ratings of Analyzing			
	M	SD	Self	Manager	Peer	Direct Report
			N = 89	N = 87	N = 89	N = 51
Numerical Reasoning	58.33	28.99	.12	.20	.17	-.15
Verbal Reasoning	61.83	27.60	.13	.13	.27*	.19

*Note: *p <.05*

Table 6: Correlations between ability scores and 360-degree ratings of Strategic Thinking for senior managers

Ability Measures	Descriptives		Correlations with 360-degree Ratings of Strategic Thinking			
	M	SD	Self	Manager	Peer	Direct Report
			N = 57	N = 57	N = 57	N = 45
Numerical Reasoning	64.58	27.37	.16	.14	.11	-.15
Verbal Reasoning	60.68	28.51	.35**	.27*	.10	-.15

*Note: *p < . 05, **p < . 01, N = 61*

In order to test how accurate the peer ratings were, 360-degree ratings on the competency Analyzing were correlated with results of the ability tests for junior managers (Table 5). Peer ratings showed statistically significant relationships with the verbal ability test results (r = .27).

Correlations between verbal reasoning and the managerial rating of strategic thinking may indicate that managers of senior managers are more likely to interact with their direct reports in discussions about strategy, which are influenced by rhetoric skills and analytical thinking.

8   Discussion

We examined which rater source in a multinational company's 360-degree feedback would provide the most valid information about their junior and senior managers' leadership competencies. Therefore, we explored relationships between 360-degree ratings and AC ratings of the same leadership competencies. The positive relationships found between AC and 360-degree ratings were consistently higher across all competencies for peer ratings, except for the competency Deciding & Initiating Action for junior managers. Correlations were of moderate size but clearly exceeded those with other rater sources.

For senior managers, ability test results were correlated with 360-degree ratings of the competency Formulating Strategies & Concepts which requires ability (Table 6). Here, relationships were only statistically significant for manager and self ratings (.27 and .35 respectively).

The pattern of results was interpreted in the specific context the organization operated in, namely the highly dynamic, fast moving consumer goods sector. To increase the organization's responsiveness to external changes, work is organized in flexible, project related systems. Here, peers may have a more comprehensive perspective on their performance as they are likely to have more opportunities to observe each other compared to subordinates, and superiors, especially when superiors are in more senior roles.

Every study has its limitations. A limitation of this study was that although 360-degree feedback and AC assessments were conducted very thoroughly, AC ratings were used as a proxy for external, impartial ratings and AC ratings are well known to have measurement challenges of their own (e.g. Lance, 2008). Moreover, using forced choice ratings for 360-degree assessment may have achieved better discrimination (Bartram, 2007) and there-

fore provided deeper insights into the rating process. Thirdly, our study provides a criterion validation of the various dimensions in the 360-degree feedback. According correlations suffer from an upper limit as given by reliability of the criterion as well as of the assessment itself. As seen in Table 1 & 2, Cronbachs alpha values differ quite considerably between judgments, which might contribute to differential levels of correlations.

Besides these limitations, the study allows several evidence-based recommendations. In many executive development contexts it is assumed that performance ratings by managers and direct reports are the most important or reliable source of information. It appears, however, that for the organization participating in our study more weight should be placed on peer ratings to follow up with more differentiated development measures. Their judgment might be a more accurate representation of observable behaviors (as assessed in assessment center exercises). As part of the feedback process scores from different rater sources should be discussed separately and explored by considering the observability of specific competencies by these rater sources. Manager ratings may be less accurate for behaviors they are less likely to observe because they interact less with their subordinates than peers. In general, peer ratings could be utilized more widely in, for example, development centers or as part of performance management programs.

Putting these results in the context of evidence-based leadership development, we suggest that leadership programs should be informed by the best available scientific evidence, which is applied in due consideration of local evidence. This study showed that in the specific organizational context in which the data was collected, peer ratings were more highly concordant with impartial observers' ratings from assessment centers. Development measures could consider this by, for example, following up with more frequent and shorter feedback ratings, based on separate rater sources for specific competencies. Only relying on peer ratings can cause issues as well, especially in highly competitive work cultures, which needs to be taken into account. We argue that HR and line managers can apply 360-degree feedback tools in leadership development programs more effectively by systematically using local research evidence that shows how well these tools work in their organization and how results may be influenced by the local context. It is not always possible to collect sufficient data for a quantitative analysis, but there are other ways of collecting local evidence which Briner and Rousseau (2011) discuss in detail.

Future research is warranted addressing the positive consequences of evidence-based leadership development. Our research agenda on "evidence for evidence based leadership development" includes qualitative as well as quantitative studies. First, we need more case studies that investigate the changes that go along with an introduction of evidence based practice in leadership development. Secondly, we need more longitudinal research that relates HR managers' focus on evidence-based practice with organizational performance indicators.

9   References

Bartram, D. (2005). The Great Eight Competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 90,* 1185-1203.

Bartram, D. (2007). Increasing validity with forced-choice criterion measurement formats. *International Journal of Selection & Assessment, 15,* 263-272.

Bartram, D., Robertson, I. T. & Callinan, M. (2002). Introduction: A framework for examining organizational effectiveness. In Robertson, I. T., Callinan, M., & Bartram, D. (Eds.), *Organizational Effectiveness: The role of Psychology* (pp.1-10). West Sussex: John Wiley & Sons.

Beehr, T. A., Ivanitskaya, L., Hansen, P. C., Erofeev, D. & Gudanowski, D. M. (2001). Evaluation of 360 degree feedback ratings: relationships with each other and with performance and selection predictors. *Journal of Organizational Behavior, 22,* 775-788.

Borman, W. C. (1997). 360 degree ratings: An analysis of assumptions and a research agenda for evaluating their validity. *Human Resource Management Review, 7,* 299–316.

Brett, J. F. & Atwater, L. E. (2001). 360 feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology, 86,* 930−942.

Briner, R. B. & Rousseau, D. M. (2011). Evidence-Based I-O Psychology: Not there yet. *Industrial and Organizational Psychology: Perspectives on Science and Practice. 4,* 3-22.

Briner, R. B., Denyer, D. & Rousseau, D. M. (2009). Evidence-based management: Construct clean-up time? *Academy of Management Perspectives, 23,* 19-32.

Coleman, J. L., (1987). *Police Assessment Testing: An assessment center handbook for law enforcement personnel.* Third edition, Springfield: Charles C Thomas Publisher.

DeNisi, A. S. & Kluger, A. N. (2000). Feedback effectiveness: Can 360-degree appraisals be improved. *Academy of Management Executive, 14,* 129−139.

Externbrink, K. & Dormann, C. (2014). Führen und Entscheiden: Evidence-based Management. In J. Felfe (Hrsg.), *Trends der psychologischen Führungsforschung – Neue Konzepte, Methoden und Erkenntnisse* (S. 429-441). Göttingen: Hogrefe.

Lance, C. E. (2008). Why assessment centers do not work the way they are supposed to. *Industrial and Organizational Psychology, 1,* 84−97.

Locke, E. A. & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57,* 705−717.

Pfeffer, J. & Sutton, R. I. (2006). *Hard facts, dangerous half-truths, and total nonsense: Profiting from evidence-based management.* Boston: Harvard Business School Press.

Reason, P. & Bradbury, H. (Eds.) (2001). *Handbook of Action Research: Participative Inquiry and Practice.* Thousand Oaks: Sage.

Rousseau, D. M. (2006). Is there such thing as "Evidence-based management"? *Academy of Management Review, 31*, 256-269.

Schmidt, F. L. & Hunter, F. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.

Seifert, C. F. & Yukl, G. (2010). Effects of repeated multi-source feedback on the influence behavior and effectiveness of managers: A field experiment. *The Leadership Quarterly, 21*, 856-866.

SHL (1998). *Online Verbal Test & Online Numerical Test (OVT&ONT).* Thames Ditton, UK: SHL Group.

SHL (2006). OPQ32: *Technical Manual.* Thames Ditton, UK: SHL Group.

Siddique, C. M. (2004). Job analysis: A strategic human resource management practice. *International Journal of Human Resource Management, 15,* 219-244.

Smither, J. W., London, M. & Reilly, R. R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings*. Personnel Psychology, 58,* 33-66.

Stajkovic, D. A. & Luthans, F. (2003). Behavioural management and task performance in organizations: Conceptual background, meta-analysis, and test of alternative models. *Personnel Psychology, 56,* 155-194.

Teece, D. J. (2009). *Dynamic Capabilities and Strategic Management: Organizing for Innovation and Growth.* Oxford University Press.

Terpstra, D. A. & Rozell, E. J. (1993). The relationship of staffing practices to organizational level measures of performance. *Personnel Psychology, 46,* 27- 48.

Walker, A. & Smither, J. W. (1999). A five-year study of upward feedback: What managers do with their results matters. *Personnel Psychology, 52,* 393-423.

Walshe, K. & Rundall, T. G. (2001): Evidence based management: From theory to   practice in healthcare. *The Milbank Quarterly, 79,* 420-457.

Warr, P. & Bourne, A. (2000). Associations between rating content and self-other agreement in multi-source feedback. *European Journal of Work and Organizational Psychology, 9,* 321-334.

Yukl, G. & Seifert, C. (2005). Facilitating multisource behavioral feedback to managers. In S. Reddy (Ed.), *Multi-source performance assessments: Perspectives and insights* (pp. 12−30). The ICFAI University Press.

Corresponding Author:
Dr. Kai Externbrink
Fresenius University of Applied Sciences
Im MediaPark 4c
D-50670 Köln
GERMANY
info@journal-bmp.de

Appendix

A 1: Assessment matrix for junior managers

Competencies	Case Study and Presentation	1:1 Role Play	1:Many Group Discussion	Ability Tests (ONT & OVT)	OPQ32
Analyzing	✓			✓	✓
Deciding & Initiating Action	✓	✓	✓		✓
Adapting & Responding to Change	✓		✓		✓
Working with People		✓	✓		✓

A 2: Assessment matrix for senior managers

Competencies	Case Study and Presentation	1:1 Role Play	Competency based Interview	Ability Tests (ONT & OVT)	OPQ32
Formulating Strategies & Concepts	✓			✓	✓
Enterprising & Performing	✓		✓		✓
Leading & Supervising		✓	✓		✓
Making an Impact	✓	✓			✓