

The Impact of Over- and Underrating on Leadership Performance

Darrin S. Kass

Bloomsburg University of Pennsylvania, USA

Paul F. Rotenberry

West Chester University of Pennsylvania, USA

Shiloh D. Erdley-Kass and

Bloomsburg University of Pennsylvania, USA

Steven Welch

Bloomsburg University of Pennsylvania, USA

Abstract

This study examined the notion of leadership over- and underrating: self-assessments of one's managerial competencies relative to assessments made by others. Specifically, we examined the relationships between over- and underrating and performance on an assessment center. Our results showed that underraters were more likely to have higher performance scores than overraters. Multiple regression results indicated that as underrating increased scores on the objectively-rated competencies rose. These findings suggest that the mechanisms that underlie underrating and overrating – perhaps humility and narcissism, respectively - have important implications for personal and professional development.

Keywords: overrating, underrating, leadership performance

1. Introduction

The goal of this research is to expand the understanding of the relationship between over- and underrating, self-assessments of one's managerial competencies relative to assessments made by others, and work performance. This study improves on previous research by including measures of objective performance on a standardized assessment center, rather than 360-feedback evaluations, which are subject to bias and inaccuracy.

2. Literature Review

Research suggests that self-perceptions of ability rarely match external perceptions, and that this degree of self-awareness may not matter with regards to leadership effectiveness and personal and professional success (Atwater and Yammarino, 1997; Fleenor, Smither, Atwater, Braddy and Sturm, 2010; Harris and Schaubroeck, 1988; Mabe and West, 1982; Moshavi, Brown and Dodd, 2003; Zenger and Folkman, 2015). In general, people's view of themselves does not fit with how other's view them likely because of various contextual factors and the notion that self-perceptions of ability are often based on comparisons to others, an "ideal self," or a past or future self. Zenger (2014) suggests that self-perceptions are only likely to match external reviews from peers and managers 50 percent of the time because of cognitive distortions, even when objective measures of

leadership are utilized. Fleenor et al. (2010) reported that extensive internal and external factors contribute to the varying degrees of agreement between self and other ratings including biographical, personality and individual characteristics, and job-relevant experiences.

Self and other ratings of performance often involve various degrees of overrating and underrating of performance and ability. Individuals who underrate their skills and evaluate themselves with lower scores than other raters are categorized as underraters, while those who overrate their abilities and evaluate themselves with higher scores are categorized as overraters. Research has shown that there are significant differences in the outcomes associated with over- and underrating.

There are conflicting outcomes associated with overrating. Some findings suggest that overraters have negative attitudes, lack an awareness of their strengths and weaknesses, and are less likely to utilize training and development opportunities (Woo, Sims, Rupp, & Gibbons, 2008). On the other hand, research has also found that overrating is correlated with high achievement, high self-esteem and social desirability, and low anxiety levels (Brutus, Fleenor and McCauley, 1999; Goffin and Anderson, 2007; Nowack and Mashihhi, 2012). Nowack and Mashihhi (2012) suggested that the contrasting outcomes of overrating may be due to an overraters' extensive focus on future achievement and their ability to selectively focus on positive behaviors. Findings also suggest that overrating skills is more common among males and older executives. Unfortunately, it can lead to detrimental personal and social outcomes such as poor organizational effectiveness, career derailment, and negative physical health outcomes (Goffin & Anderson, 2002; Vecchio and Anderson, 2009; Woo, Sims, Rupp, and Gibbons, 2008).

Similar to conflicting outcomes regarding overrating, research suggests that although underraters are more likely to be perceived as more effective leaders, they also tend to exhibit an inability to recognize and use their existing strengths. (Kaplan and Kaiser, 2009; Nowack and Mashihhi, 2012) While underraters are more likely to focus on weaknesses or areas of improvement, and are more likely to accept feedback and make changes, they are also more likely to be overly critical, blind to strengths, and overly focused on unrealistic personal and organizational goals. (Nowack and Mashihhi, 2012; Wilson and Ross, 2001) Perfectionistic tendencies are often a key driving force behind the self-perceptions and behaviors of underraters. (Nowack and Mashihhi, 2012) Still, recent research has indicated that the more individuals underrate themselves on a 360-degree feedback assessment, the more highly they are perceived as a leader and their employees are more engaged. (Zenger and Folkman, 2015) Additionally, underrating is associated with a higher probability of possessing strengths and a lower probability of possessing fatal flaws such as lack of energy and enthusiasm, acceptance of mediocre performance, lack of interpersonal skills, and poor judgment.

The purpose of this research is to expand the understanding of the outcomes associated with over- and underraters by determining if underrating is associated with greater success on objective measures of managerial performance. This study improves on previous research in two ways. First, research on over- and underrating has typically used 360-degree feedback assessments, which vary greatly in accuracy and effectiveness (for a review see Nowack and Mashihhi, 2012). 360-degree feedback is only moderately objective, as raters, "may have limited ability or motivation to provide accurate and constructive feedback." (Bommer, Rubin, & Bartels, 2005, p. 103) The current study includes an assessment center, which uses a standardized work simulation and trained raters with no previous knowledge of the participants. The present research includes both self-ratings and peer-reviewed external ratings on managerial skills like *Communication*, *Leadership Initiative*, *Teamwork*, *Organizing*, and *Decision-Making*, and thus allows for the most unbiased, objective determination of over- and underrating. Second, the assessment center is a simulation that provides a realistic management performance scenario that has been shown to be predictive of managerial success. (Gaugler, Rosenthal, Thornton, & Bentson, 1987; Howard, 1997; Whetzel, Rotenberry, and McDaniel, 2014) Therefore, it allows for the evaluation of very objective and unbiased performance outcomes associated with over- and underraters.

1. Method

We collected data from students enrolled in a part-time MBA program at a regional university in the Pennsylvania State System of Higher Education. The program has been accredited by AACSB since 2004 and typically graduates an average of 25 students per academic year. The sample was drawn from 205 students over a 6-year span. Students were excluded if they did not complete the Iliad Assessment Center managerial exercise, which is discussed below, near the conclusion of their degrees. In addition, missing data on the control variables (also discussed below) led to participant exclusion.

The final sample consisted of 115 participants: 44 women and 71 men. Ages ranged from 22 to 60 years, with an average age of 27.2 years. Sixteen participants identified a language other than English as their primary language. All of the participants included in the sample completed the GMAT prior to being admitted into the MBA program and scores were considered in admissions decisions; GMAT score ranged from 320 to 700, with an average of 478.

2. Measures

2.1 Assessment Center

The Iliad Assessment Center (Bommer & Bartels, 1996) was used to identify skill levels in five dimensions: Active Communication, Teamwork, Decision Making, Leadership Initiative, and Organizing. The Iliad has been validated and employed in other published studies, and Rode, Arthaud-Day, Mooney, Near, Baldwin, Bommer, and Rubin (2005) discuss it in greater detail. It is a 145-minute simulation of a day in the life of a manager and requires participants to complete the following tasks: 1) attend two 20-minute leaderless group discussion meetings; one to discuss and develop customer service initiatives, and the other to identify a potential CEO successor for the company, 2) deliver a 3-minute persuasive speech about the participant's vision for the organization, and 3) work on an in-basket exercise in the role of a top manager in a publishing company.

The in-basket is a work document that includes the times and content of the leaderless group discussion meetings and speech times, administrative information (contact information, organizational charts, etc.), and a series of managerial tasks that require some sort of action on the part of the participant (e.g., financial and operational decisions, communications, delegation). The activities included in the simulation require more than the time allotted to complete, and therefore, participants need to prioritize their workloads and manage their time and schedule. Participants are responsible for maintaining their schedules during the simulation, as well as determining how much time, effort, and energy to allocate to the managerial tasks.

The leaderless team discussions and speech were video recorded. All materials were sent to the Iliad Assessment Center for scoring based on the presence and effectiveness of specific behaviors. The raters were blind to this study and to the identity of the students. Raters had an average of 1 year of rating experience and were either present or former students in a master's of industrial psychology program. Two independent ratings were completed, and then conflicts between raters were settled by collaboratively reviewing the recordings and coming to a consensus. Scores are provided for the following: Selection (effective evaluation of CEO candidates in a team setting), Customer Service (effective evaluation of different customer service initiatives in a team setting), In-Basket (materials participants must act upon during the assessment by sending emails, memos, etc.), Speech (verbal, non-verbal, and content-related aspects of a 3 minute persuasive speech),

and Total Score (an equal-weighting combination of Selection, Customer Service, In-Basket, and Speech scores).

2.2 Over- and Underrating

Over- / underrating was defined as the difference score between the self-ratings of participants and the actual scores assigned by trained raters. Prior to starting the Iliad Assessment, participants were asked to complete a self-assessment of their skills in comparison to their peers; this was operationalized as a percentage of the people they would outperform in Leadership Initiative, Decision Making, Teamwork, Communication, and Organizing. The actual score on each dimension was provided by external Iliad raters. The difference between actual scores and the self-ratings (described below as the 'GAP' variables) was either positive indicating underrating (self-assessment was lower than actual score) or negative indicating overrating (self-assessment was higher than actual score).

3. Results

3.1 Data Cleaning

All variables of interest were evaluated for outliers and violations of normality. Z score analyses of all subjects on all variables showed that no significant outliers existed. Tests for Skewness and Kurtosis revealed that the data in general meet the assumption of normality. The same litmus test for significance was used for these analyses: values of Z, Skewness, or Kurtosis greater than or equal to ± 3.48 (.001 two tailed) were considered outliers or a violation of the assumption of normality, respectively.

Analyses revealed that only two variables were significantly Skewed: Decision Making self and Teamwork self. None of the variables were significantly Kurtotic, however. Based on the overall findings it was deemed that the assumption of normality was met.

3.2 Data Analyses

The mean values for each of the 5 GAP variables (Objectives scores given by the raters for Leadership Initiative, Decision Making, Organizing, Communicating, and Teamwork as compared to the self-ratings provided by the participants) were negative (see Table 1). This is indicative of a pattern of overrating across the board.

Table 1: Means and Standard Deviations

Variable	Mean	Standard Deviation
Gender	1.38	.49
Age	27.24	6.03
GMAT	478.43	84.30
Graduate School GPA	3.59	.23
Selection	50.20	29.49
Customer Service	45.09	28.05
Speech	40.54	24.71
In-Basket	42.29	28.27
Total Score	39.98	28.29
Leadership Initiative	52.47	27.54
Decision Making	39.37	27.32
Organizing	40.50	27.61
Communication	44.24	29.17
Teamwork	43.28	26.29
Leadership Initiative Self	70.33	17.66
Decision Making Self	69.20	19.03
Organizing Self	67.29	20.19
Communication Self	63.84	22.09
Teamwork Self	75.32	17.50
Leadership Initiative GAP	-17.86	30.41
Decision Making GAP	-29.84	33.79
Organizing GAP	-26.78	36.35
Communication GAP	-19.60	36.39
Teamwork GAP	-32.04	30.25

N = 115

It is interesting to note that very few (only three out of 25 coefficients) self-rating scores on the skills (Leadership Initiative, Communication, etc.) were related to the five objective outcomes (Selection, Customer Service, Speech, In-Basket, & Total Score) on the Iliad Assessment. This is consistent with findings from previous research (Zenger, 2014). However, over- and underrating (GAP variables) were related to *all* of the objective outcome variables (23 out of 25 coefficients were significant; see Tables 2 and 3).

Table 2: Correlations

	Age	GMAT	Grad. GPA	Selection	Customer	Speech	In-basket	Total Score	Leader Initiative	Decision	Org.
Age											
GMAT	.044										
Grad GPA	.051	.378***									
Selection	-.020	.203*	.211*								
Customer	-.068	.049	.011	.409***							
Speech	-.110	.213*	.182*	.329***	.212*						
In-basket	.013	.165	.351***	.325***	.239*	.283**					
Total Score	-.027	.226*	.333***	.634***	.534***	.545***	.851***				
Leader Initiative	-.111	.175 ^A	.247**	.462***	.430***	.197*	.766***	.762***			
Decision	-.002	.294**	.303**	.463***	.391***	.555***	.702***	.825***	.562***		
Org.	-.052	.283**	.312**	.429***	.253**	.606***	.666***	.740***	.523***	.561***	

Note: Customer = Customer Service; Decision = Decision Making; Org. = Organizing;

*p < .05; **p < .01; ***p < .001;

Table 2: Correlations (continued)

	Age	GMAT	Grad. GPA	Selection	Customer	Speech	In-basket	Total Score	Leader Initiative	Decision	Org.
Comm.	-.031	.112	.300**	.478***	.418***	.436***	.823***	.855***	.646***	.576***	.573***
Teamwork	-.039	.043	.006	.622***	.628***	.187*	.150	.463***	.236*	.282**	.209*
LI-Self	-.093	-.029	-.026	.012	.214*	-.011	-.001	.065	.149	.069	-.031
DM-Self	-.026	.004	-.033	-.081	.128	-.085	-.050	-.038	.063	-.032	-.114
O-Self	-.082	-.006	-.013	-.132	.025	-.214*	-.044	-.121	-.019	-.175 ^D	-.136
C-Self	.002	.070	-.052	.069	.225*	.036	-.005	.080	.127	.082	-.015
T-Self	-.083	-.068	-.061	.016	.137	-.151	-.027	-.024	.040	-.037	-.129
LI-GAP	-.046	.176 ^B	.238*	.411***	.265**	.185*	.694***	.652***	.819***	.469***	.492***
DM-GAP	.013	.236*	.264**	.420***	.244**	.497***	.596***	.688***	.418***	.826***	.518***
O-GAP	.006	.218*	.244**	.399***	.178 ^C	.579***	.530***	.629***	.408***	.524***	.835***
C-GAP	-.026	.048	.272**	.341***	.199*	.328***	.663***	.637***	.441***	.412***	.469***
T-GAP	.014	.077	.041	.532***	.466***	.250**	.146	.416***	.182*	.266**	.256**

Note: Comm. = Communication; Customer = Customer Service; Decision = Decision Making; Org. = Organizing; LI-Self = Leadership Initiative Self-Rating; DM-Self = Decision Making Self-Rating; O-Self = Organizing Self-Rating; C-Self = Communication Self-Rating; T-Self = Teamwork Self-Rating; LI-GAP = Leadership Initiative GAP; DM-GAP = Decision Making GAP; O-GAP = Organizing GAP; C-GAP = Communication GAP; T-GAP = Teamwork GAP

*p < .05; **p < .01; ***p < .001;

^A p < .061; ^B p < .060; ^C p < .056; ^D p < .061

Table 2: Correlations (continued)

	Comm.	Team work	LI-Self	DM-Self	O-Self	C-Self	T-Self	LI-GAP	DM-GAP
Comm.									
Teamwork	.277**								
LI-Self	.023	.120							
DM-Self	-.074	.097	.754***						
O-Self	-.066	.100	.562***	.668***					
C-Self	.011	.192*	.734***	.631***	.581***				
T-Self	-.015	.090	.764***	.679***	.597***	.587***			
LI-GAP	.572***	.144	-.445***	-.380***	-.344***	-.311**	-.408***		
DM-GAP	.507***	.173 ^A	-.369***	-.589***	-.518***	-.290**	-.413***	.593***	
O-GAP	.472***	.103	-.336***	-.457***	-.659***	-.334***	-.430***	.565***	.681***
C-GAP	.795***	.106	-.427***	-.442***	-.405***	-.598***	-.368***	.647***	.582***
T-GAP	.249**	.817***	-.338***	-.308**	-.259**	-.173 ^A	-.500***	.361***	.389***

Note: Comm. = Communication; LI-Self = Leadership Initiative Self-Rating; DM-Self = Decision Making Self-Rating; O-Self = Organizing Self-Rating; C-Self = Communication Self-Rating; T-Self = Teamwork Self-Rating; LI-GAP = Leadership Initiative GAP; DM-GAP = Decision Making GAP; O-GAP = Organizing GAP; C-GAP = Communication GAP; T-GAP = Teamwork GAP

* $p < .05$; ** $p < .01$; *** $p < .001$

^A $p < .065$

Table 2: Correlations (continued)

	O-GAP	C-GAP	T-GAP
O-GAP			
C-GAP	.581***		
T-GAP	.338***	.305**	

Note: O-GAP = Organizing GAP; C-GAP = Communication GAP; T-GAP = Teamwork GAP

** $p < .01$; *** $p < .001$

Table 3: Summary of Significant Relationships between GAP Scores and Outcome Variables

Objective Outcome Variables	GAP Score Correlations
Selection Percentile	Leadership Initiative $r = .411, p < .001$ Decision Making $r = .420, p < .001$ Organizing $r = .399, p < .001$ Communication $r = .341, p < .001$ Teamwork $r = .532, p < .001$
Customer Service Percentile	Leadership Initiative $r = .265, p < .01$ Decision Making $r = .244, p < .01$ Organizing $r = .178, p < .056$ Communication $r = .199, p < .05$ Teamwork $r = .466, p < .001$
Speech Percentile	Leadership Initiative $r = .185, p < .05$ Decision Making $r = .497, p < .001$ Organizing $r = .579, p < .001$ Communication $r = .328, p < .001$ Teamwork $r = .250, p < .01$
In-basket Percentile	Leadership Initiative $r = .694, p < .001$ Decision Making $r = .596, p < .001$ Organizing $r = .530, p < .001$ Communication $r = .663, p < .001$
Total Score Percentile	Leadership Initiative $r = .652, p < .001$ Decision Making $r = .688, p < .001$ Organizing $r = .629, p < .001$ Communication $r = .637, p < .001$ Teamwork $r = .416, p < .001$

Uniformly positive correlation coefficients between the GAP variables and the Iliad outcomes indicated that as underrating increased so too did scores on the outcomes measures. Table 3 summarizes these relationships.

Multiple regression analysis was used to further examine the relationships between the GAP variables and the objective outcomes (see Table 4). The model tested entered the control variables (Gender, Age, GMAT score, and Graduate GPA) in the first step, followed by the five GAP variables in the second step. Total Score, an aggregate of the four assessment center ratings, was used as the dependent variable.

Table 4: Multiple Regression Analysis

Dependent Variable	Control & Independent Variables	r with DV	b	B
Total Score				
R ² = .632				
FΔ = 28.919	Gender	--	-3.089	-.053
(Model 2)	Age	-.027	-.098	-.021
	GMAT	.226**	.009	.027
	Grad GPA	.333***	13.631	.110
	Initiative GAP	.652***	.196*	.210
	Decision Making GAP	.688***	.236**	.281
	Organizing GAP	.629***	.102	.131
	Comm GAP	.637***	.157*	.202
	Teamwork GAP	.416***	.111 ^a	.118

Note: *p < .05; **p < .01; ***p < .001, ^a p < .075

Model one proved significant (R² = .126, FΔ (4, 110) = 3.972, p < .01), as did model two (R² = .632, FΔ (5, 105) = 28.919, p < .001). Coefficients in model two demonstrated that four of the five GAP variables had an impact on the Total Score: Leadership Initiative (b = .196; t = 2.470, p < .05), Decision Making (b = .236; t = 3.119, p < .01), Communication (b = .157; t = 2.315, p < .05), and Team, which was ‘marginally significant’ (b = .111; t = 1.797, p < .075). Thus, in all four cases underrating oneself on a given dimension led to a higher Total Score on the Iliad Assessment Center.

4. Discussion

The results of this study demonstrated that underrating of managerial self-competencies is associated with better performance on an objectively scored assessment center. We found that underraters, individuals who provided lower self-scores compared to other raters on several managerial skills, were more likely to have higher scores on the assessment center outcomes, particularly the total score. Our study improved on previous research in this area by including objective, unbiased measures of both over- and underrating and performance, lending support to the integrity of the findings. These findings are notable because the assessment center method (ACM) is a strong predictor of several important work outcomes: salary progress, job performance, training performance, and promotions (see Thornton and Gibbons, 2009 for a review). Whetzel et al. (2014) concluded the ACM is appropriate for “high stakes settings” and meta-

analysis has estimated its mean population criterion-related validity to be .36 for the prediction of job performance (see Gaugler et al, 1987, p. 502).

It is also interesting to note that self-ratings in the current study did not predict objective assessment center performance. Self-ratings were significantly associated with only 12% of the potential outcomes (three of the 25 coefficients). This would suggest that the emphasis placed on the accuracy of self-ratings may be unwarranted. While several authors have suggested that accurate self-ratings, agreement between self and other ratings, is important for performance (see Guenole, Cockerill, Chamorro-Premuzic, 2011), our findings indicate that underrating may be a better predictor of managerial performance.

A central question is whether underrating actually represents a lack of self-awareness for an individual. Zenger and Folkman (2015) noted that underraters were humble, had high personal standards, and a continuous drive to improve. The concept of humility does fit with the construct of underrating, and may help to explain the outcomes associated with it. Owens and Hekman (2012) indicated that humility is a “virtue” leading to growth, such that humble people see themselves as a work in progress and thus may rate themselves somewhat lower than their peers. This coupled with the fact that, “they fall short of a standard they are striving to reach” (Ou, Tsui, Kinicki, Waldman, Xiao, & Song, 2014, p. 38) may explain why individuals underrate—they recognize the constant need for growth and continually fall short of this high standard. Perhaps underraters do not lack self-awareness, rather they are aware of their skills and abilities, but understand they need to grow, evolve, and improve.

If underraters are humble, perhaps overraters are narcissistic. Self-enhancement is a central aspect of narcissism (Grijalva & Zhang, 2016), and narcissists were found to overrate their attributes and abilities, such as leadership effectiveness, intelligence, physical attractiveness, openness, and honesty. Essentially, narcissists appear to be over-raters. Interestingly, Owens, Wallace, and Waldman (2015) found that leaders with high levels of narcissism and low levels of humility had lower scores on objective and subjective measures of job performance, perceived effectiveness, and job engagement than their low narcissism/high humility peers. It appeared that humility seemed to diminish the negative effects of narcissism. The overraters in the current study therefore may be those with high levels of narcissism and low levels of humility.

5. Limitations and Future Directions

The current article had several limitations. First, the study included MBA student performance on a managerial assessment center. While this is indicative of managerial performance, we did not have actual work-related outcomes (for example, like those in Owens, Wallace, & Waldman, 2015). Future research is needed to examine over- and underrating and objective measures of actual on-the-job performance. In addition, nearly all of the participants were from the United States, and there are most likely cross-cultural differences in tendencies toward over- and underrating (e.g., cultures that are individualistic vs collectivistic). Future studies need to include measures of narcissism, humility, and related variables in order to better understand the behavioral and psychological mechanisms that underlie over- and underrating. Specifically, more clarity is needed to better understand what personality types are more prone to inflating or underrating ability. Knowledge of such attributes may contribute to understanding the value of self-awareness and leadership outcomes. Understanding why underraters tend to perform better on both objective and subjective measures of leadership has important implications for management education and management development.

References

- Atwater, L & Yammarino, F 1997, "Self-other rating agreement: A review and a model," *Research in Personnel and Human Resources Management*, 15, pp. 121-174.
- Bommer, W & Bartels, L 1996, "Assessing the unassessable: Interpersonal and managerial skills," in Martell, K & Calderon, T 2005, *Assessment of student learning in business schools: Best practices each step of the way*, The Iliad Assessment Center, Bloomington.
- Brutus, S, Fleenor, J & McCauley, C 1999, "Demographic and personality predictors of congruence in multi-source ratings," *Journal of Management Development*, 18, 5, pp. 417-435.
- Fleenor, J, Smither, J, Atwater, L, Braddy, P & Sturm, R 2010, "Self-other rating agreement in leadership: A review," *The Leadership Quarterly*, 21, 6, pp.1005-1034.
- Gaugler, B, Rosenthal, D, Thornton, G & Bentson, C 1987, "Meta-analysis of assessment center validity," *Journal of Applied Psychology*, 72, pp. 493-511.
- Goffin, R & Anderson, D 2002, "Differences in self and superior ratings of performance: Personality provides clues," *Society for Industrial and Organizational Psychology, Toronto, Canada*.
- Goffin, R & Anderson, D 2007, "The self-rater's personality and self-other disagreement in multi-source performance ratings: Is disagreement healthy?" *Journal of Managerial Psychology*, 22, 3, pp. 271-289.
- Grijalva, E & Zhang, L 2016, "Narcissism and self-insight: A review and meta-analysis of narcissists' self-enhancement tendencies," *Personality and Social Psychology Bulletin*, 42, 1, pp. 3-24.
- Guenole, N, Cockerill, T, Chamorro-Premuzic, T & Smillie, L 2011, "Evidence for the validity of 360 dimensions in the presence of rater-source factors," *Consulting Psychology Journal: Practice and Research*, 63, 4, pp. 203-218.
- Harris, M & Schaubroeck, J 1988, "A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings," *Personnel Psychology*, 41, 1, pp. 43-62.
- Howard, A 1997, "A reassessment of assessment centers: Challenges for the 21st century," *Journal of Social Behavior and Personality*, 12, 5, pp. 13-52.
- Kaplan, R & Kaiser, R 2009, "Stop overdoing your strengths," *Harvard Business Review*, 87, 2, pp. 100-103.
- Mabe, P.A. and West, S.G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), pp. 280-296.
- Moshavi, D, Brown, F & Dodd, N 2003, "Leader self-awareness and its relationship to subordinate attitudes and performance," *Leadership & Organization Development Journal*, 24, 7, pp. 407-418.

- Nowack, K & Mashihi, S 2012, "Evidence-based answers to 15 questions about leveraging 360-degree feedback," *Consulting Psychology Journal: Practice and Research*, 64, 3, pp. 157-182.
- Ou, A, Tsui, A, Kinicki, A, Waldman, D, Xiao, Z & Song, L 2014, "Humble chief executive officers' connections to top management team integration and middle managers' responses," *Administrative Science Quarterly*, 59, 1, pp. 34-72.
- Owens, B & Hekman, D 2012, "Modeling how to grow: An inductive examination of humble leader behaviors, contingencies, and outcomes," *Academy of Management Journal*, 55, 4, pp. 787-818.
- Owens, B, Wallace, A & Waldman, D 2015, "Leader narcissism and follower outcomes: The counterbalancing effect of leader humility," *Journal of Applied Psychology*, 100, 4, pp. 1203-1213.
- Rode, J, Arthaud-Day, M, Mooney, C, Near, J, Baldwin, T, Bommer, W & Rubin, R 2005, "Life satisfaction and student performance," *Academy of Management Learning & Education*, 4, 4, pp. 421-433.
- Thorton, G and Gibbons, A 2009, Validity of assessment centers for personnel selection. *Human Resource Management Review*, 19, pp. 169-187.
- Vecchio, R & Anderson, R 2009, "Agreement in self-other ratings of leader effectiveness: The role of demographics and personality," *International Journal of Selection and Assessment*, 17, 2, pp. 165-179.
- Whetzel, D, Rotenberry, P & McDaniel, M 2014, "In-basket validity: A systematic review," *International Journal of Selection and Assessment*, 22, 1, pp. 62-79.
- Wilson, A, and Ross, M 2001, "From chump to champ: People's appraisals of their earlier and present selves," *Journal of Personality and Social Psychology*, 80, 4, pp. 572-584.
- Woo, S, Sims, C, Rupp, D & Gibbons, A 2008, "Development engagement within and following developmental assessment centers: Considering feedback favorability and self-assessor agreement," *Personnel Psychology*, 61, 4, pp. 727-759.
- Zenger, J 2014, "The singular secret for a leader's success: Self-awareness," *Forbes*, Available: <http://www.forbes.com/sites/jackzenger/2014/04/17/the-singular-secret-for-a-leaders-success-self-awareness/#4571114177e4>.
- Zenger, J & Folkman, J 2015, "We like leaders who underrate themselves. *Harvard Business Review*," Available: <https://hbr.org/2015/11/we-like-leaders-who-underrate-themselves>.

