

This article was downloaded by:[Gentry, William A.]
On: 13 August 2007
Access Details: [subscription number 781206261]
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



European Journal of Work and Organizational Psychology

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713684945>

A study of the discrepancy between self- and observer-ratings on managerial derailment characteristics of European managers

Online Publication Date: 01 September 2007

To cite this Article: Gentry, William A., Hannum, Kelly M., Ekelund, Bjørn Z. and de Jong, Annemarie (2007) 'A study of the discrepancy between self- and observer-ratings on managerial derailment characteristics of European managers', European Journal of Work and Organizational Psychology, 16:3, 295 - 325

To link to this article: DOI: 10.1080/13594320701394188

URL: <http://dx.doi.org/10.1080/13594320701394188>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

© Taylor and Francis 2007

A study of the discrepancy between self- and observer-ratings on managerial derailment characteristics of European managers

William A. Gentry and Kelly M. Hannum
Center for Creative Leadership, Greensboro, NC, USA

Bjørn Z. Ekelund
Human Factors AS/Agder University College, Norway

Annemarie de Jong
de Baak, Noordwijk & Driebergen, The Netherlands

Managerial derailment is costly to managers, their co-workers, and their organization. Knowing whether discrepancies (i.e., differences, dissimilarity, disagreement, incongruity) exist between self- and observer- (subordinates, peers, and bosses) ratings about derailment may help to lessen or prevent the detrimental outcomes of derailment on managers, their co-workers, and their organization. Results from 1742 European managers revealed a statistically significant difference between a manager's self-ratings and observer-ratings on the extent to which a manager displayed derailment behaviours and characteristics. The discrepancy also widened as managerial level increased, and was mostly due to inflated self-ratings. In addition, discrepancies between self–boss, self–direct report, and self–peer were examined, as well as differences between European and American managers. A discussion of these findings and implications for practice conclude this article.

Multisource (i.e., multirater, 360-degree) instruments remain a useful device for researchers and practitioners. Researchers have used multisource instruments to advance knowledge in many areas (e.g., Atwater, Ostroff, Yammarino, & Fleenor, 1998; Atwater, Roush, & Fischthal, 1995;

Correspondence should be addressed to William A. Gentry or Kelly M. Hannum, Center for Creative Leadership, 1 Leadership Place, PO Box 26300, Greensboro, NC 27438-6300, USA. E-mail: gentryb@leaders.ccl.org; hannumk@leaders.ccl.org

The first two authors contributed equally to this manuscript. The authors would like to thank Vicente González-Romá and Adam Meade for their help.

Church, 1997). Also, many Fortune 1000 and 500 companies have utilized multisource instruments for employee development, feedback, and other purposes (Atwater & Waldman, 1998; Conway, Lombardo, & Sanders, 2001; Edwards & Ewen, 1998).

Ratings from multiple “observer” perspectives (e.g., peer, subordinate, supervisor, client, customer) gathered as part of multisource instruments are frequently part of developmental feedback to managers. The ratings can signal individual and organizational outcomes (Dalessio, 1998; London & Smither, 1995; Morgeson, Mumford, & Campion, 2005; Tornow, 1993; Yammarino, 2003) and are related to performance measures and outcomes (Church, 2000; Conway et al., 2001; Smither & Walker, 2001). However, a manager’s self-rating frequently diverges from ratings provided by “observer” perspectives. Previous research has indicated that a discrepancy exists (i.e., a difference, dissimilarity, disagreement, or incongruity) between a manager’s self-rating on the constructs measured by the multisource instrument and observers’ rating of that same manager on those same constructs (Brutus, Fleenor, & McCauley, 1999; Morgeson et al., 2005; Mount, Judge, Scullen, Sytsma, & Hezlett, 1998; Sala, 2003). Consequently, further research is needed to understand why such discrepancies occur since, as Fletcher (1997) argued, it is extremely difficult for managers to handle work relationships, to contribute as a team member, and to adapt behaviours to different work circumstances when self perceptions are discrepant or differ from observer perceptions.

This study builds upon Ostroff, Atwater, and Feinberg’s (2004) and Sala’s (2003) research on self–observer discrepancies and explores whether managerial level affects self–observer discrepancies. Unlike previous studies, our research focuses specifically on self- and observer-ratings concerning managerial derailment, which is imperative to study given that managerial derailment can be very costly to managers and organizations (as described later). In addition, as multisource instruments gain popularity around the world, studies using multisource ratings are needed focusing on raters across and within different countries (Atwater, Waldman, Ostroff, Robie, & Johnson, 2005; Brutus, Leslie, & McDonald-Mann, 2001). Another way in which this study differs from previous research is that we examine managers from European countries in order to further knowledge about multisource instruments and self–observer rating discrepancies. Moreover, in an exploratory manner, this study will also compare European managers to managers in the United States.

MANAGERIAL DERAILMENT

Lombardo and McCauley (1988) indicated that derailment “occurs when a manager who was expected to go higher in the organization and who was

judged to have the ability to do so is fired, demoted, or plateaued below expected levels of achievement" (p. 1). Derailed managers leave the organization or plateau due to lack of fit between personal characteristics and job demands, or their own skills and the demands of their present job (Leslie & Van Velsor, 1996). Not only do derailed managers continue ineffective job habits because of their preference to do things their own way (Kovach, 1986), but they also do not adjust their behaviour despite differing changes and demands in their present managerial role (Lombardo & Eichinger, 1989/2005).

McCall and Lombardo (1983) interviewed executives about derailment. Those preliminary interviews and subsequent research (e.g., Leslie & Van Velsor, 1996; Lombardo & McCauley, 1988; Morrison, White, & Van Velsor, 1987) helped determine behaviours and characteristics of managerial derailment. These behaviours and characteristics are arranged into five clusters as measured by BENCHMARKS[®],¹ a multisource instrument. Though the original managerial derailment research was based on North American managers, the same clusters are found in European managers as well (Leslie & Van Velsor, 1996). The first cluster is "problems with interpersonal relationships" which describes managers as isolated from co-workers and as authoritarian, cold, aloof, arrogant, and insensitive. The second cluster, "difficulty leading a team" consists of failing to staff effectively, the inability to form and lead teams, and the inability to handle conflict. Third, "difficulty changing or adapting" involves a manager's inability to adapt to a boss with a different managerial or interpersonal style and a manager's inability to grow, learn, develop, and think strategically. Fourth, "failure to meet business objectives" examines whether managers are poor performers, are overly ambitious, or lack follow-through. Finally, "too narrow functional orientation" considers a manager's ill-preparedness for promotion and a manager's inability to supervise outside of the manager's current function. Several researchers have studied the dynamics of managerial derailment and assessed how derailment occurs (Hogan & Hogan, 2001; McCall & Lombardo, 1983) as well as derailment's costly outcomes (Finkelstein, 2004; Smart, 1999; Wells, 2005).

The direct and indirect costs associated with derailed managers and executives can be very costly to an organization, sometimes several times more than an executive's salary (Bunker, Kram, & Ting, 2002; Finkelstein, 2004; Lombardo & McCauley, 1988; Smart, 1999; Wells, 2005). Discrepancies between the extent to which a manager believes he or she displays behaviours and characteristics that lead to derailment and the extent to which observers of the manager believe that same manager displays those

¹BENCHMARKS[®] is a registered trademark of the Center for Creative Leadership.

same derailment behaviours and characteristics may signal an important and potentially expensive “blind spot” that must be addressed.

SELF–OBSERVER DISCREPANCY

Often, multisource instruments identify warning signals of managerial derailment so managers can react before real trouble occurs. However, self-ratings and observer-ratings do not always agree (Harris & Schaubroeck, 1988; Tsui & Ohlott, 1988), which may cause confusion for managers in their attempt to develop and better themselves.

In their model of self–other agreement, Atwater and Yammarino (1997) proposed that certain individual characteristics would impact the size of the discrepancy between self- and observer-ratings. Furthermore, according to Sala (2003), the size of the self–observer discrepancy may be “bigger/wider” for some managers than others, based on job level. Sala’s research indicates that higher level managers (e.g., those in “top” or “executive” level jobs) may have “bigger/wider” self–observer rating discrepancies than lower level managers (e.g., middle-level managers). One possible reason for this trend is that a “disconnect” may exist between what a higher level manager thinks of himself or herself, and what observers (i.e., bosses, peers, and direct reports) think of the higher level manager, such that higher level managers may be more “out of touch” with how they are perceived by observers than lower level managers (Goleman, Boyatzis, & McKee, 2001). Also, higher level managers may be more confident or more arrogant, may not want to perform self-assessments, may be less agreeable to accept input and truthful feedback from others or may have others around them who are reluctant to offer straightforward assessments and feedback (Conger & Nadler, 2004; Dotlich & Cairo, 2003; Kaplan, Drath, & Kofodimos, 1987; Kovach, 1986; Levinson, 1994; “Recognizing the symptoms”, 2003).

Previous research examined self–observer discrepancies by using an index of difference scores or categories of agreement. However, these types of studies may have methodological flaws since this “difference score index” does not give a complete understanding of the underlying reason for the discrepancy (Edwards, 1995). By applying multivariate regression procedures² (see Edwards, 1995, or the Analytic Approach section of this article for details), the present research will determine whether a self–observer discrepancy exists in rating the extent to which European managers display the behaviours and characteristics of managerial derailment, whether the self–observer discrepancy is “bigger/wider” for higher managerial levels,

²We thank two anonymous reviewers for this issue and guidance for using this as our data analytic technique.

and determine the source of the discrepancy (higher self-ratings, lower observer-ratings, or both). Previous research indicated that the self-observer discrepancy widens as managerial level increases (Sala, 2003), and that managers at higher levels may have inflated self-ratings (Conger & Nadler, 2004; Goleman et al., 2001; Kaplan et al., 1987; Van Velsor, Ruderman, & Philips, 1991). As a result, we predict that:

Hypothesis 1. Managerial level will be a significant predictor of the self-observer discrepancy in rating the extent to which managers display the behaviours and characteristics that lead to derailment. The self-observer discrepancy will be “bigger/wider” for managers at higher managerial levels than for managers in lower levels due to inflated self-ratings.

Combining ratings may mask important findings based on the specific rating relationships between self-boss, self-peer, and self-direct report and therefore, it may be important to analyse observers separately. We therefore offer the following hypotheses related to each specific manager-ratee relationship:

Hypothesis 2. Managerial level will be a significant predictor of the self-boss discrepancy in rating the extent to which managers display the behaviours and characteristics that lead to derailment. The self-boss discrepancy will be “bigger/wider” for managers at higher managerial levels than for managers in lower levels due to inflated self-ratings.

Hypothesis 3. Managerial level will be a significant predictor of the self-direct report discrepancy in rating the extent to which managers display the behaviours and characteristics that lead to derailment. The self-direct report discrepancy will be “bigger/wider” for managers at higher managerial levels than for managers in lower levels due to inflated self-ratings.

Hypothesis 4. Managerial level will be a significant predictor of the self-peer discrepancy in rating the extent to which managers display the behaviours and characteristics that lead to derailment. The self-peer discrepancy will be “bigger/wider” for managers at higher managerial levels than for managers in lower levels due to inflated self-ratings.

EXPLORATORY ANALYSES

The previous hypotheses used both self- and observer-ratings as “dependent” or “outcome” variables (see Edwards, 1995), examining whether managerial level predicts the size of the self-observer discrepancy.

One can also use self- and observer-ratings as separate “independent” or “predictor” variables, to predict outcomes (see Edwards, 1994; Edwards & Parry, 1993). As briefly mentioned earlier and also later in the Analytic Approach section, methodological weaknesses and problems exist when combining self- and observer-ratings into a difference score “index” or composite. We therefore will use a squared difference model (Edwards & Parry, 1993). By using self- and observer-ratings as two separate and distinct variables, our first set of exploratory analyses will examine whether the self–observer discrepancy in rating the extent to which managers display the behaviours and characteristics that lead to derailment predicts whether or not managers are seen as likely to derail in the near future from the perspective of the target manager’s boss.

The second set of exploratory analyses will use self- and observer-scores separately as outcome variables, and determine whether origin predicts self–observer discrepancy ratings, specifically, whether the size of the self–observer discrepancy differs between managers from Europe and managers from the United States. This aspect of the study is intended to determine if the pattern of self–observer rating differences is influenced by origin. While our approach (combining European countries) is not nuanced enough to address cultural differences, it is a first step to determining if there are similar rating discrepancies based on origin.

METHOD

Participants and procedures

Data were obtained via product sales to and program use with a diverse sample of European managers. Data analysis included managers from European countries that contained at least 10 managers who were native to and currently working in the same European country. As a result, for the preliminary overall European analysis, the present research used data from 1742 European managers from over 320 different companies from Austria, Belgium, Denmark, France, Germany, Ireland, Italy, The Netherlands, Norway, Poland, Spain, Sweden, Switzerland, Turkey, and the United Kingdom. The managers’ mean age was approximately 39 years old (range = 25 – 62). Also, approximately 80% of the managers were male, 85% had at a minimum a bachelor’s degree, and 93% were from the private sector. Finally, 23% were “middle-level” managers, 50% were “upper-middle-level” managers, and 27% were “top-level” managers from the “top” or “executive” ranks.

Managers who took part in a leadership development process between October 2000 and May 2006 provided self-ratings using BENCHMARKS, a multisource instrument. The managers selected and asked observers

(including direct reports, peers, and bosses) to provide ratings by completing the observer form of BENCHMARKS.

For the exploratory analysis between European and American managers, we used a comparable sample of 1928 randomly selected managers from a database who were native to and currently working in the United States. These US managers participated in a leadership development processes during the same time period as the European managers and also took BENCHMARKS. The demographics of these managers were similar to those of the European sample (mean age 42, range 26–68 years of age, 80% male, 89% minimum of a bachelor's degree, 84% from the private sector, 20% middle-level managers, 48% upper-middle-level managers, and 32% top-level managers).

Measures

BENCHMARKS is a multisource feedback instrument, accumulating ratings from the self, direct report, peer, and boss perspective (Lombardo & McCauley, 1994). McCauley and Lombardo (1990) described the original development of BENCHMARKS, and Lombardo, McCauley, McDonald-Mann, and Leslie (1999) described the most recent updates with appropriate psychometric information. Only data from Section 2 was used, which focused on the extent to which raters believed that managers displayed the behaviours and characteristics of derailment. BENCHMARKS is based on research about how successful managers learn, grow, and change (Lindsey, Homes, & McCall, 1987; Morrison et al., 1987), and about how managers derail (Leslie & Van Velsor, 1996; Lombardo & McCauley, 1988). BENCHMARKS is a well-validated multisource feedback instrument (Center for Creative Leadership, 2002; Douglas, 2003; Lombardo & McCauley, 1994; McCauley, Lombardo, & Usher, 1989; Zedeck, 1995) and has also been used in prior research (e.g., Atwater et al., 1998; Brutus, Fleenor, & London, 1998; Brutus et al., 1999; Conway, 2000; Fleenor, McCauley, & Brutus, 1996; Gentry, Mondore, & Cox, in press).

Section 2 of BENCHMARKS contains 40 items that measure the five derailment clusters described earlier. Raters used a 5-point Likert-type scale, with 1 = *strongly disagree* (that the manager displays the following characteristic) to 5 = *strongly agree* (that the manager displays the following characteristic). Lower scores (scores closer to "1" in magnitude) implied that managers were less likely to display the behaviours and characteristics that lead to derailment. The 40 questions asked were the same for the manager and his/her observers, only the frame of reference was different (rate yourself vs. rate the person).

We also asked each boss of the target manager "What is the likelihood that this person will derail (i.e., plateau, be demoted, or fired) in the next

5 years as a result of his/her actions or behaviours as a manager?" Scores ranged from 1 = *not at all likely* to 5 = *almost certain*. This question is used for the exploratory analyses previously discussed.

Discrepancy score calculations

An average of 8.85 observers rated each European manager. Observers consisted of 1 boss, an average of 3.85 peers (range 3–10), and an average of 4 direct reports (range 3–12). To justify aggregation of ratings across observers, intraclass correlations, [or ICC(2)] were computed, reflecting the reliability of aggregated scores of each derailment cluster for the observers across each manager. ICC(2) values for observers, for each derailment cluster were above .67. The same is true for each derailment cluster for direct reports, and for peers. Past research using similar multisource data found comparable ICC(2) values (see Atwater et al., 1998; Fleenor et al., 1996; Ostroff et al., 2004). In addition, r_{wg} values were computed for the observer ratings for each manager, on each derailment cluster. The mean r_{wg} values for the observers for each manager on each derailment cluster were above .85. The same is true for peers and direct reports for each manager on each derailment cluster. For the American managers, an average of 8.89 observers (1 boss, an average of 4 direct reports ranging from 3 to 11, and an average of 3.89 peers ranging from 3 to 16) rated each manager. All ICC(2) values were above .67 and mean r_{wg} values were above .85 for the American managers. Given these findings, we aggregated scores across observers, peers, and direct reports. Self- and observer-ratings were centred based on the midpoint of their shared scale (Edwards, 1994). Therefore, scores ranged from -2 to $+2$, with ratings closer to -2 in magnitude meaning that managers were less likely to display the behaviours and characteristics that lead to derailment.

Analytic approach

Multivariate regression procedures may be a more suitable approach to analyse self–observer discrepancy (difference) scores both as predictor variables (see Edwards, 1994; Edwards & Parry, 1993) and outcome variables (see Edwards, 1995). In the past, as Edwards (1994, 1995) and Edwards and Perry (1993) pointed out, countless studies have used different approaches to analysing self–observer discrepancies or “difference scores” by combining them into a single index using algebraic, absolute, squared difference, or the sum of absolute or squared difference between self- and observer-ratings. However, these approaches are flawed (Cronbach, 1958; Cronbach & Furby, 1970; Edwards, 1994, 1995; Edwards & Cooper, 1990; Edwards & Parry, 1993). For instance, combining two distinct component

measures (such as a self-rating and observer-rating) into one variable or index cannot be unambiguously interpreted, leads to decreased reliability, confounds the effects of components by masking the true relative contribution of the components to variance, puts what is supposed to be measured with a multivariate model into a univariate framework, and consequently, provides misleading results (Edwards, 1994, 1995).

Collapsing self- and observer-ratings into a single index by subtracting the observer-rating from the self-rating for each derailment cluster does not allow researchers to investigate the independent effects from each rating source, and does not offer knowledge of the underlying explanation for the disagreement (Edwards, 1995; Ostroff et al., 2004). Instead of collapsing scores into a single index, both the self-rating and observer-rating are retained separately and tested jointly for each derailment cluster.

Analyses for hypotheses. The four a priori hypotheses tested in this article used a multivariate framework with self- and observer-ratings considered jointly as outcome variables, and determined whether managerial level predicted the discrepancy between self- and observer-ratings of the extent to which a target manager displays the characteristics and behaviours that lead to derailment. For each set of regressions, we first examined whether the relationship between managerial level and self-observer ratings for each derailment cluster was significant overall (i.e., an omnibus multivariate test based on Wilks's Λ , a test for self-ratings and for observer-ratings jointly). This omnibus test reveals whether or not managerial level is significantly related to self- and observer-ratings considered jointly. If the Wilks's Λ is statistically significant, then managerial level is related to self- and observer-ratings considered jointly. Next, we inspect the source of the discrepancy by treating self- and observer-ratings as distinct (separate) outcome variables. This information determines whether: managerial level is related to self-ratings, producing the discrepancy due to inflated self-ratings; managerial level is related to observer-ratings, producing discrepancy due to inflated observer ratings; or both.

Analyses for exploratory investigation. In the first set of exploratory analyses, we used self- and observer-ratings as separate predictor variables, and determined whether the discrepancy predicted how likely the target manager's boss thinks the target manager will derail in the next 5 years (see Edwards, 1994; Edwards & Parry, 1993, for more details about this type of analysis). To avoid same-source bias concerning the target-manager's boss, we only examined self-direct report, and self-peer ratings as predictor variables, and the boss-rating as the outcome variable. For the first set of analyses, the boss-ratings of future derailment was regressed on self-ratings of each derailment cluster, direct-report ratings of each derailment cluster,

the product of self \times direct report ratings, the square of self-ratings, and the square of direct-report ratings. We conducted the same analyses using peer-ratings as well.

In the second set of exploratory analyses, we used the self–observer discrepancy as outcome variables (similar to Hypotheses 1–4 above) to examine whether origin (Europe or United States) predicted the discrepancy between self- and observer-ratings of the extent to which a target manager displays the characteristics and behaviours that lead to derailment. Omnibus multivariate analyses were first conducted. If statistically significant, we examined the source of the discrepancy by examining regression coefficients with self- and observer-ratings treated as separate outcome variables.

RESULTS

Self–observer analysis

The means, standard deviations, and alpha-reliabilities for each of the derailment clusters from self-, observer-, peer-, direct report-, and boss-ratings for all 1742 European managers regardless of level are found in Table 1. The intercorrelations among the variables can be found in Table 2.

The first set of analyses focused on the discrepancy between self- and observer-ratings (boss, peer, and direct report aggregated). The results from the multivariate test for self- and observer-ratings are found in Table 3, and each show that managerial level was significantly related to the set of self- and observer-ratings considered jointly for each of the derailment clusters. As a result, regression analyses were used treating self- and observer-ratings separately to examine the nature of the relationship between managerial level and self–observer ratings on each of the derailment clusters.

Table 4 provides results using self-ratings and observer ratings as separate variables to examine with regards to managerial level. The regression coefficient of self-ratings for “problems with interpersonal relationships” was significant; managers at different managerial levels rated themselves differently, such that managers at higher managerial levels rated themselves as less likely to display the behaviours and characteristics of derailment on “problems with interpersonal relationships” than managers at lower levels (scores were more negative in magnitude, or closer to “–2”). The same is true for each of the other four derailment clusters.

The regression coefficient of observer-ratings for “problems with interpersonal relationships” was also significant; observers rated managers at different managerial levels differently, such that observers of managers at higher managerial levels rated managers as more likely to display the behaviours and characteristics of derailment on “problems with interpersonal relationships” than observers of managers at lower levels (scores were

TABLE 1
Means, standard deviations, and alpha reliabilities of derailment clusters variables

<i>Variables</i>	<i>M</i>	<i>SD</i>	α
Level	2.04	0.71	
Self-PIR	-1.36	0.50	.86
Self-DLAT	-1.30	0.50	.87
Self-DCA	-1.39	0.45	.82
Self-FMBO	-1.41	0.47	.79
Self-TNFO	-1.30	0.57	.74
Obs-PIR	-1.18	0.41	.91
Obs-DLAT	-1.08	0.36	.91
Obs-DCA	-1.26	0.31	.88
Obs-FMBO	-1.26	0.34	.87
Obs-TNFO	-1.12	0.41	.83
Boss-PIR	-1.36	0.56	.89
Boss-DLAT	-1.26	0.57	.89
Boss-DCA	-1.37	0.50	.86
Boss-FMBO	-1.45	0.52	.84
Boss-TNFO	-1.16	0.68	.81
DR-PIR	-1.18	0.50	.91
DR-DLAT	-1.09	0.46	.90
DR-DCA	-1.29	0.39	.88
DR-FMBO	-1.26	0.44	.88
DR-TNFO	-1.25	0.47	.83
Peer-PIR	-1.14	0.49	.92
Peer-DLAT	-1.02	0.44	.91
Peer-DCA	-1.19	0.38	.87
Peer-FMBO	-1.21	0.41	.86
Peer-TNFO	-0.99	0.53	.83

PIR=problems with interpersonal relationships; DLAT=difficulty leading a team; DCA=difficulty changing or adapting; FMBO=failure to meet business objectives; TNFO=too narrow functional orientation. Obs=all observers. DR=direct reports.

Level 1 = middle-level managers, 2 = upper-middle-level managers, 3 = high-level managers.

less negative in magnitude, or farther away from “-2”). The regression coefficient of observer-ratings for “difficulty leading a team” was not significant; observers rated managers at different managerial levels similarly on “difficulty leading a team”. The results concerning “difficulty changing and adapting”, “failure to meet business objectives”, and “too narrow functional orientation” all concluded that the regression coefficient for self-ratings was significant for each derailment cluster, but the regression coefficient for observer-ratings was not significant for the same derailment cluster. Taken together, these results support Hypothesis 1. A significant self-observer discrepancy exists, and widens as managerial level increases for each derailment cluster. Moreover, while observer-ratings was one of the

TABLE 2
Correlations among derailment clusters variables

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26			
1. Level																													
2. Self-PIR	-.05																												
3. Self-DLAT	-.12	.66																											
4. Self-DCA	-.14	.68	.73																										
5. Self-FMBO	-.09	.57	.68	.68																									
6. Self-TNFO	-.12	.36	.55	.56	.62																								
7. Obs-PIR	.06	.31	.08	.09	-.01	-.04																							
8. Obs-DLAT	.01	.18	.21	.12	.06	.07	.73																						
9. Obs-DCA	-.02	.19	.14	.19	.06	.08	.79	.82																					
10. Obs-FMBO	.01	.10	.08	.08	.12	.04	.64	.76	.78																				
11. Obs-TNFO	.01	.08	.12	.11	.10	.23	.53	.72	.77	.78																			
12. Boss-PIR	.05	.20	.06	.12	.01	-.02	.56	.40	.51	.33	.28																		
13. Boss-DLAT	-.02	.10	.15	.14	.07	.09	.32	.51	.46	.37	.38	.70																	
14. Boss-DCA	-.04	.10	.08	.18	.04	.07	.35	.38	.55	.36	.38	.76	.76																
15. Boss-FMBO	.00	.04	.04	.11	.09	.04	.23	.31	.40	.46	.38	.58	.67	.74															
16. Boss-TNFO	.00	.03	.07	.14	.07	.18	.22	.35	.45	.37	.56	.49	.61	.69	.68														
17. DR-PIR	.06	.26	.11	.07	.01	-.03	.85	.66	.66	.56	.46	.34	.20	.18	.10	.14													
18. DR-DLAT	.03	.14	.20	.06	.05	.05	.58	.81	.62	.60	.54	.19	.25	.16	.12	.17	.76												

(continued overleaf)

TABLE 2 (continued)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
19. DR-DCA	.00	.16	.15	.13	.06	.05	.67	.72	.80	.67	.62	.30	.27	.27	.19	.25	.81	.82									
20. DR-FMBO	.00	.10	.11	.04	.09	.03	.56	.66	.64	.81	.61	.18	.21	.16	.19	.20	.70	.77	.81								
21. DR-TNFO	.00	.07	.12	.06	.08	.16	.45	.63	.62	.68	.79	.14	.23	.19	.19	.29	.59	.71	.78	.81							
22. Peer-PIR	.04	.26	.03	.07	-.03	-.05	.85	.58	.66	.54	.43	.42	.19	.25	.15	.14	.49	.27	.36	.28	.21						
23. Peer-DLAT	-.01	.15	.13	.10	.03	.05	.61	.80	.68	.62	.61	.30	.32	.27	.21	.25	.34	.35	.35	.32	.32	.73					
24. Peer-DCA	-.03	.16	.07	.17	.03	.07	.64	.64	.82	.62	.63	.39	.31	.39	.29	.33	.32	.25	.37	.28	.29	.78	.81				
25. Peer-FMBO	.01	.08	.02	.06	.08	.04	.51	.57	.62	.80	.63	.22	.22	.23	.30	.24	.25	.23	.29	.36	.31	.65	.75	.75			
26. Peer-TNFO	.02	.07	.08	.09	.07	.20	.44	.57	.62	.62	.84	.21	.26	.27	.25	.37	.24	.25	.31	.28	.39	.53	.70	.74	.74		

PIR = problems with interpersonal relationships; DLAT = difficulty leading a team; DCA = difficulty changing or adapting; FMBO = failure to meet business objectives; TNFO = too narrow functional orientation. Obs = all observers. DR = direct reports.

Level 1 = middle-level managers, 2 = upper-middle-level managers, 3 = high-level managers.

Correlations at or above .05 in absolute value, significant at $p < .05$.

TABLE 3
Multivariate analysis results with managerial level predicting self-other ratings

Variable	PIR		DLAT		DCA		FMBO		TNFO	
	Wilks's Λ	F	Wilks's Λ	F	Wilks's Λ	F	Wilks's Λ	F	Wilks's Λ	F
Self-Observer	.99	5.80**	.98	8.76**	.98	10.49**	.99	5.15**	.98	11.07**
Self-Boss	.99	4.81*	.98	8.23**	.98	11.06**	.99	4.46*	.98	10.59**
Self-Direct report	.99	5.85**	.98	8.99**	.98	10.41**	.99	4.06*	.98	9.64**
Self-Peer	.99	4.20*	.98	9.21**	.98	10.67**	.99	5.94**	.97	11.43**

PIR = problems with interpersonal relationships; DLAT = difficulty leading a team; DCA = difficulty changing or adapting; FMBO = failure to meet business objectives; TNFO = too narrow functional orientation. * $p < .01$, ** $p < .001$.

TABLE 4
Regression analysis of self and other ratings on managerial level

Variable	PIR		DLAT		DCA		FMBO		TNFO	
	<i>b</i>	<i>R</i> ²	<i>b</i>	<i>R</i> ²	<i>b</i>	<i>R</i> ²	<i>b</i>	<i>R</i> ²	<i>b</i>	<i>R</i> ²
Self	-0.04*	.002	-0.09***	.015	-0.09***	.019	-0.06***	.007	-0.10***	.015
Observer	0.04*	.004	0.01	.000	-0.01	.000	0.00	.000	0.01	.000
Boss	0.03	.002	-0.01	.000	-0.03	.002	0.00	.000	0.00	.000
Direct report	0.04*	.004	0.02	.001	0.00	.000	0.00	.000	0.00	.000
Peer	0.03	.002	-0.01	.000	-0.02	.001	0.01	.000	0.02	.000

PIR = problems with interpersonal relationships; DLAT = difficulty leading a team; DCA = difficulty changing or adapting; FMBO = failure to meet business objectives; TNFO = too narrow functional orientation.

* $p < .05$, *** $p < .001$.

causes of the self–observer discrepancy only in ratings concerning “problems with interpersonal relationships”, those at higher managerial levels appear as “overestimators” because the inflated self-ratings were the cause of the discrepancy between self- and observer-ratings of the extent to which managers display the behaviours and characteristics of all derailment clusters. Figure 1 shows this relationship using “difficulty leading a team” as an example.

Self–boss analysis

The second set of analyses focused on the discrepancy between self- and boss-ratings. The results from the multivariate test for self- and boss-ratings are found in Table 3. Managerial level was significantly related to the set of self- and observer-ratings considered jointly for each of the derailment clusters. Next, we conducted regression analyses treating self- and boss-ratings separately to examine the nature of the relationship between managerial level and self-boss ratings on each of the derailment clusters, as is shown in Table 4.

The regression coefficient and findings for self-ratings are the same as those previously discussed in the self–observer analysis. The regression coefficient of boss-ratings for “problems with interpersonal relationships” was not significant; bosses rated managers at different managerial levels similarly. Figure 2a reveals this relationship. The regression coefficient of “difficulty leading a team” (see Figure 2b), “difficulty changing and adapting” (see Figure 2c), “failure to meet business objectives” (see Figure 2d), and “too narrow functional orientation” (see Figure 2e) was also similar in that (a) self-ratings were significant, meaning that managers

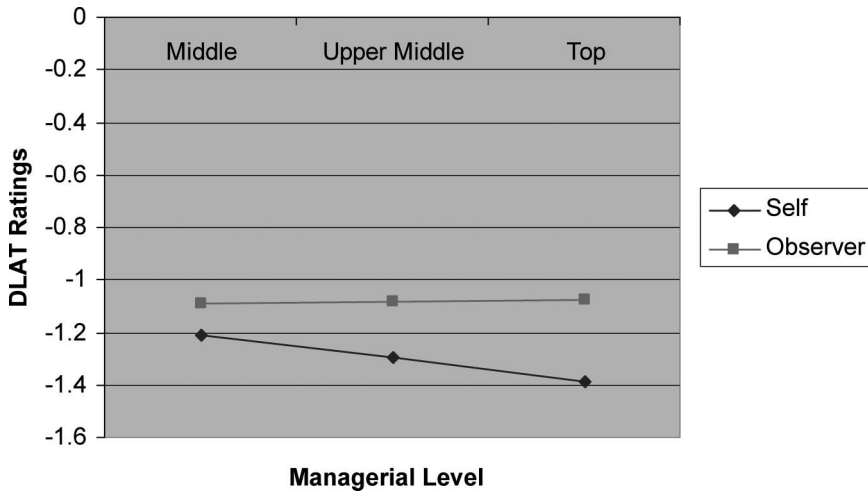


Figure 1. Self-observer discrepancies as a function of organizational level (“difficulty leading a team” as an example).

at higher managerial levels rated themselves as less likely to display the behaviours and characteristics of a particular derailment cluster than managers at lower levels (scores were more negative in magnitude, or closer to “-2”), but (b) the regression coefficient for boss-ratings was not statistically significant, meaning that bosses rated managers at different managerial levels similarly on that same derailment cluster. From the figures, one can see that middle-level managers rated themselves as more likely to display the behaviours and characteristics that lead to derailment than their own boss (self-ratings tend to be above boss-ratings in the figures and further away from “-2” in magnitude). However, managers at higher levels (e.g., “top” levels) see themselves as less likely to display the behaviours and characteristics that lead to derailment than their bosses (self-ratings are below boss-ratings in the figures and closer to “-2” in magnitude). Taken together, these results partially support Hypothesis 2. For the most part, managers at lower managerial levels rated themselves as more likely to display those behaviours and characteristics (scores are less negative in magnitude or farther away from “-2” in magnitude) than their own boss. However, at higher managerial levels, self-boss ratings are less congruent, the discrepancy becomes “bigger/wider”, and managers believe they are less likely to display the behaviours and characteristics that lead to certain derailment clusters than their boss (scores are more negative in magnitude or closer to “-2” in magnitude). Moreover, the source of the discrepancy tends to be due to inflated self-ratings.

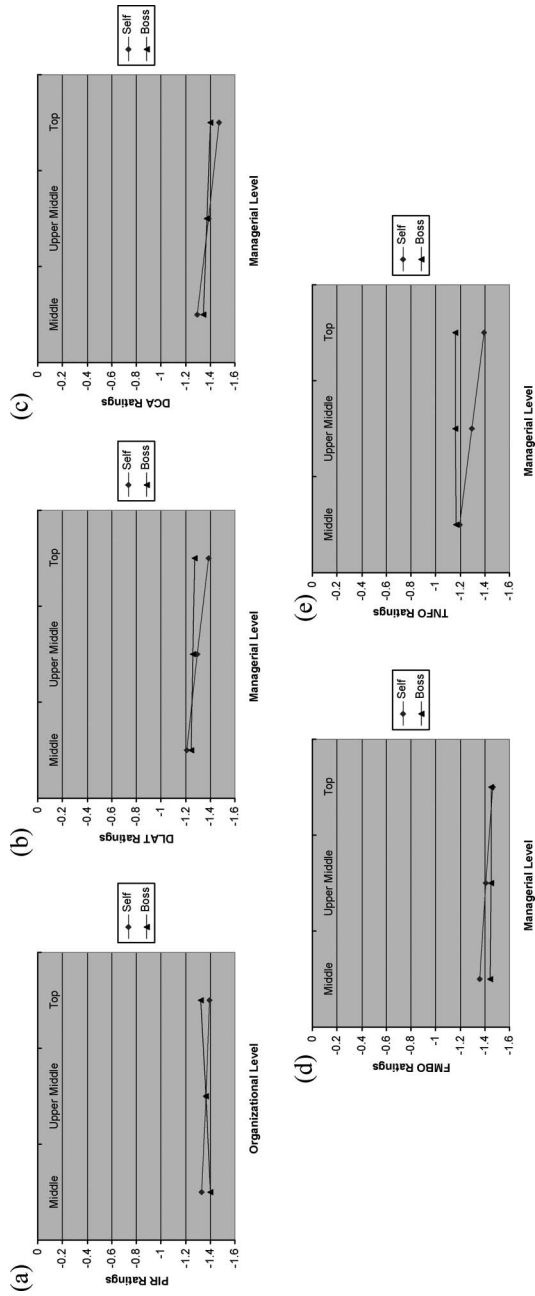


Figure 2. Self-boss discrepancies as a function of organizational level.

Self-direct report analysis

The third set of analyses focused on the discrepancy between self- and direct report-ratings. The results from the multivariate test for self- and direct report-ratings are found in Table 3. Managerial level was significantly related to the set of self- and direct report-ratings considered jointly for each of the derailment clusters. Next, we conducted regression analyses treating self- and direct report-ratings separately to examine the nature of the relationship between managerial level and self-direct report ratings on each of the derailment clusters as is shown in Table 4.

The regression coefficient and findings for self-ratings is the same as those previously discussed in the self-observer analysis. The regression coefficient of direct report-ratings for “problems with interpersonal relationships” was also statistically significant; direct reports rated managers at different managerial levels differently, such that direct reports of managers at higher managerial levels rated managers as more likely to display the behaviours and characteristics of derailment on “problems with interpersonal relationships” than observers of managers at lower levels (scores were less negative in magnitude, or farther away from “-2” on magnitude). Regarding the results obtained for “difficulty leading a team”, the regression coefficient for direct report-ratings was not significant; direct reports rated managers at different managerial levels similarly on “difficulty leading a team”. The same findings for self- and direct report-ratings are true for the other derailment clusters. The graphical display of the relationship is similar to the self-observer display of Figure 1. Taken together, these results support Hypothesis 3. A significant self-direct report discrepancy exists, and widens as managerial level increases for each derailment cluster. Moreover, while direct report-ratings was one of the causes of the self-observer discrepancy only in “problems with interpersonal relationships”, those at higher managerial levels appear as “overestimators” because the inflated self-ratings were the cause of the discrepancy between self- and observer-ratings of the extent to which managers display the behaviours and characteristics of all derailment clusters.

Self-peer analysis

The fourth set of analyses focused on the discrepancy between self- and peer-ratings. The results from the multivariate test for self- and peer-ratings are found in Table 3. Analyses were conducted exactly as the previous analyses. Managerial level was significantly related to the set of self- and peer-ratings considered jointly for each of the derailment clusters. Findings from the separate regression analyses treating self- and peer-ratings separately are shown in Table 4.

Results remained the same for self-ratings. The regression coefficient for peer-ratings for each derailment cluster was not significant, supporting Hypothesis 4. A significant self–peer discrepancy exists, and widens as managerial level increases for each derailment cluster. Moreover, those at higher managerial levels appear as “overestimators” because the inflated self-ratings were the cause of the discrepancy between self- and peer-ratings of the extent to which managers display the behaviours and characteristics of all derailment clusters. The graphical display is very similar to the self–observer ratings of Figure 1.

Exploratory analyses

In addition to the analyses described above, we conducted two sets of exploratory analyses. The first used the self- and direct report-ratings of each derailment cluster as predictor variables to examine whether these ratings could predict boss-ratings of likelihood of derailing in 5 years. We conducted the same analyses using the self and peer ratings. The second exploratory analysis used the self–observer discrepancy ratings as an outcome variable in determining whether origin (Europe or United States) would predict the ratings.

Likelihood of derailing. Each boss of the target manager was asked the likelihood of the target manager to derail in the next 5 years based on his/her actions. We used this response as an outcome variable, to determine first if the self–direct report discrepancy and then the self–peer discrepancy could predict answers to this question. Of the original 1742 managers from Europe, 1715 had bosses answer the question.

In examining self- and direct report-ratings, the “likelihood to derail” question was regressed on self-ratings (the regression coefficient b_1), direct report-ratings (the regression coefficient b_2), the square of self-ratings (the regression coefficient b_3), the product of self- and direct report-ratings (the regression coefficient b_4), and the square of direct-report ratings (the regression coefficient b_5) for each derailment cluster. Five regressions were computed, one for each derailment cluster. Results can be found in Table 5.

From Table 5 one can determine $a_1 = (b_1 + b_2)$ to calculate the slope of the line of perfect agreement, where self-ratings equal direct observer ratings ($S = DR$). For “problems with interpersonal relationships” and “difficulty changing and adapting”, a_1 is statistically significant, meaning the hypothesis that the surface is flat along the $S = DR$ line is rejected. Results for self–direct report agreement with “problems with interpersonal relationships” and “difficulty changing and adapting” are presented graphically in Figures 3a and 3b respectively. As one can see by the figures, along the line of perfect agreement ($S = DR$), agreement at higher levels

TABLE 5
Self-direct report discrepancy and boss-rated likelihood to derail in five years

<i>Variable</i>	<i>PIR</i> <i>b (SE)</i>	<i>DLAT</i> <i>b (SE)</i>	<i>DCA</i> <i>b (SE)</i>	<i>FMBO</i> <i>b (SE)</i>	<i>TNFO</i> <i>b (SE)</i>
Constant	2.20 (0.12)**	2.14 (0.13)**	3.15 (0.21)**	2.12 (0.20)**	2.05 (0.13)**
Self	0.20 (0.12)	0.20 (0.13)	0.64 (0.18)**	0.03 (0.17)	0.02 (0.12)
Direct report	0.36 (0.11)*	0.15 (0.15)	10.05 (0.22)**	0.28 (0.19)	0.13 (0.15)
Self squared	0.00 (0.04)	-0.01 (0.05)	-0.02 (0.04)	-0.02 (0.05)	-0.01 (0.04)
Self × Direct report	0.15 (0.08)	0.17 (0.09)	0.35 (0.11)*	0.04 (0.10)	-0.01 (0.08)
Direct report squared	0.02 (0.05)	-0.13 (0.07)	0.14 (0.08)	0.00 (0.07)	-0.03 (0.06)
R^2	0.01**	0.01**	0.04**	0.01**	0.02**
Surface tests					
a ₁	0.56*	0.35	10.69**	0.31	0.15
a ₂	0.17	0.03	0.47*	0.02	-0.05
a ₃	0.16	-0.05	0.41	0.25	0.11
a ₄	-0.13	-0.31	-0.23	-0.06	-0.03

PIR = problems with interpersonal relationships; DLAT = difficulty leading a team; DCA = difficulty changing or adapting; FMBO = failure to meet business objectives; TNFO = too narrow functional orientation.
* $p < .05$, ** $p < .01$.

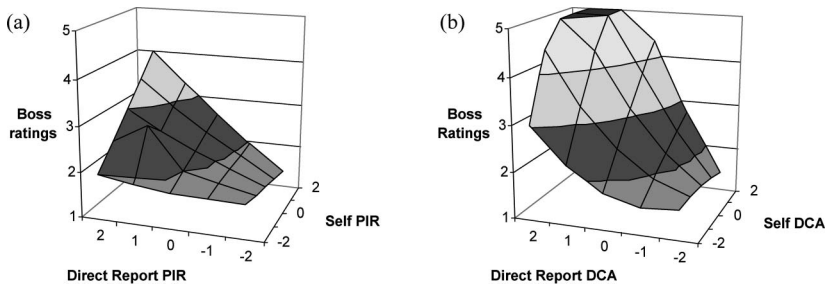


Figure 3. Relationships between self–direct report agreement (X – Y axis) and boss-ratings of derailment outcome (Z axis).

(both self and direct reports agree that the manager displays “problems with interpersonal relationships” and “difficulty changing and adapting” derailment characteristics) is related to higher boss-ratings of likelihood of derailment, than agreement at lower levels, where both the manager and direct reports believe that the manager does not display “problems with interpersonal relationships” and “difficulty changing and adapting” derailment characteristics.

The curvature along the $S = DR$ line is given by $a_2 = (b_3 + b_4 + b_5)$. For all clusters, a_2 is positive, indicating a convex surface, except “too narrow functional orientation” where the result is negative, indicating a concave surface. Looking at Table 5, a_2 is statistically significant only for “difficulty changing and adapting”, which means that the surface along the $S = DR$ line is not flat, and is a convex surface. For “difficulty changing and adapting”, the significant a_2 value leads to the conclusion that boss-ratings of likelihood of derailment decreased as self- and direct report-ratings of displaying “difficulty changing and adapting” derailment characteristics became less likely. This also suggests that as self- and direct report-ratings of displaying “difficulty changing and adapting” derailment characteristics both increased, boss-ratings of likelihood of derailment increased at an increasing rate. One can see this by the upward curvature along the $S = DR$ line in Figure 3b.

When inspecting the term $a_3 = (b_2 - b_1)$, if a_3 differs significantly from zero, there is a linear slope along the $S = -DR$ line. For a_3 none of the results were statistically significant. Next, when inspecting the term $a_4 = (b_3 - b_4 + b_5)$, one can determine whether a wider discrepancy predicts likelihood of derailment in 5 years. For a_4 none of the results were statistically significant.

It is also important to note from Table 5 that the regression coefficients for direct report-ratings for each derailment cluster are larger than self-ratings (except “difficulty leading a team”), and the regression coefficients

for direct report-ratings are statistically significant for “problems with interpersonal relationships” and “difficulty leading a team”. This leads to the conclusion that underrating (managers believing themselves to be less likely to display derailment characteristics than their direct reports) is related to higher boss-ratings of the likelihood of derailment, than overrating.

The same analyses were conducted examining self- and peer-ratings. Results can be found in Table 6. Results were similar to the direct report analysis. The response surface a_1 was statistically significant for “problems with interpersonal relationships”, “difficulty leading a team”, and “difficulty changing and adapting”, meaning that along the line of perfect agreement ($S = P$), agreement at higher levels (both self-ratings and peer-ratings are in agreement that the manager displays the aforementioned derailment characteristics) is related to higher ratings of likelihood of derailment from the boss, than agreement at lower levels (both self-ratings and peer-ratings agree that the manager does not display the aforementioned derailment characteristics). Also, the response surface a_2 was significant for “difficulty changing and adapting” only, leading to the same conclusion as was for direct reports. In addition, results for a_3 and a_4 were not statistically significant for each derailment cluster. Finally, the regression coefficients for peer-ratings were larger than self-ratings, and were significant for “problems

TABLE 6
Self-peer discrepancy and boss-rated likelihood to derail in five years

Variable	PIR <i>b</i> (SE)	DLAT <i>b</i> (SE)	DCA <i>b</i> (SE)	FMBO <i>b</i> (SE)	TNFO <i>b</i> (SE)
Constant	2.24 (0.11)**	2.31 (0.13)**	3.39 (0.19)**	2.30 (0.19)**	2.05 (0.09)**
Self	0.16 (0.12)	0.17 (0.13)	0.65 (0.17)**	0.05 (0.16)	-0.03 (0.10)
Peer	0.39 (0.12)*	0.42 (0.15)*	1.32 (0.21)**	0.38 (0.20)	0.20 (0.11)
Self squared	0.02 (0.04)	0.01 (0.05)	0.02 (0.04)	-0.02 (0.05)	0.00 (0.04)
Self × Peer	0.12 (0.08)	0.11 (0.09)	0.37 (0.11)*	0.06 (0.09)	-0.06 (0.06)
Peer squared	-0.02 (0.06)	-0.05 (0.07)	0.12 (0.09)	-0.04 (0.08)	-0.03 (0.05)
R^2	0.03**	0.04**	0.10**	0.04**	0.04**
Surface tests					
a_1	0.55*	0.59*	1.97**	0.43	0.17
a_2	0.12	0.07	0.51*	0.00	-0.09
a_3	0.23	0.25	0.67	0.33	0.23
a_4	-0.12	-0.15	-0.23	-0.12	0.03

PIR = problems with interpersonal relationships; DLAT = difficulty leading a team; DCA = difficulty changing or adapting; FMBO = failure to meet business objectives; TNFO = too narrow functional orientation.

* $p < .05$, ** $p < .01$.

with interpersonal relationships”, “difficulty leading a team”, and “difficulty changing and adapting”, concluding that underrating is related to higher boss-ratings of the likelihood of derailment, than overrating.

Europe and United States comparison. Our second exploratory analysis compared the European managers with managers in the United States on the self–observer discrepancy in ratings of the extent to which managers displayed the behaviours and characteristics of derailment. We used the aforementioned comparable sample of 1928 randomly selected managers.

The multivariate test for self- and observer-ratings indicated that origin (i.e., Europe or United States) was significantly related to the set of self- and observer-ratings considered jointly for “problems with interpersonal relationships”, Wilks’s $\Lambda = .99$, $F(2, 3667) = 13.09$, $p < .001$, “failure to meet business objectives”, Wilks’s $\Lambda = .99$, $F(2, 3667) = 4.02$, $p < .05$, and “too narrow functional orientation”, Wilks’s $\Lambda = .99$, $F(2, 3667) = 17.63$, $p < .001$. However, “difficulty leading a team”, Wilks’s $\Lambda = .99$, $F(2, 3667) = 1.74$, *ns*, and “difficulty changing and adapting”, Wilks’s $\Lambda = .99$, $F(2, 3667) = 2.56$, *ns*, revealed findings that were not statistically significant. As a result, regression analyses were used treating self- and observer-ratings separately to examine the nature of the relationship between origin and self–observer ratings in the first three derailment clusters mentioned. Results can be found in Table 7.

The regression coefficient of self-ratings for “problems with interpersonal relationships” was not significant; managers in Europe rated themselves similarly to managers in the United States. The regression coefficient of observer-ratings for “problems with interpersonal relationships” was significant such that observers rated managers from Europe as less likely

TABLE 7
Regression analysis of self- and observer-ratings on managers from Europe and United States

<i>Variable</i>	<i>Self b (SE)</i>	<i>Observer b (SE)</i>
PIR	−0.01 (0.02)	0.08 (0.02)**
DLAT	0.01 (0.02)	0.02 (0.01)
DCA	−0.02 (0.02)	0.02 (0.01)
FMBO	−0.04 (0.02)*	0.02 (0.01)
TNFO	−0.07 (0.02)**	0.05 (0.01)**

PIR = problems with interpersonal relationships; DLAT = difficulty leading a team; DCA = difficulty changing or adapting; FMBO = Failure to meet business objectives; TNFO = too narrow functional orientation. For coding of origin, 1 = Europe, 2 = United States.

* $p < .05$, ** $p < .01$.

to display the behaviours and characteristics of derailment on “problems with interpersonal relationships” than observers of American (scores are more negative or closer to “-2” in magnitude for observers of European managers than American managers). Figure 4a shows this relationship. From the regression analysis of Table 7, the self-observer discrepancy for “problems with interpersonal relationships” is due to observers rating American managers differently than observers of European managers, and is not due to self-ratings.

The regression coefficient of self-ratings for “failure to meet business objectives” was significant; managers in Europe rated themselves as more likely to display the behaviours and characteristics that lead to derailment on “failure to meet business objectives” than American managers (scores are less negative or farther away from “-2” in magnitude for European managers than American managers). The regression coefficient of observer-ratings for “failure to meet business objectives” was not significant, indicating that observers rated Europeans and Americans similarly (see Figure 4b). For “failure to meet business objectives”, inflated self-ratings are a cause of the self-observer discrepancy.

Finally, the regression coefficient for both self-ratings and observer-ratings was significant for “too narrow functional orientation” (see Figure 4c). European managers believe they are more likely to display the behaviours and characteristics that lead to derailment on “too narrow functional orientation” than American managers (self-ratings for Europeans

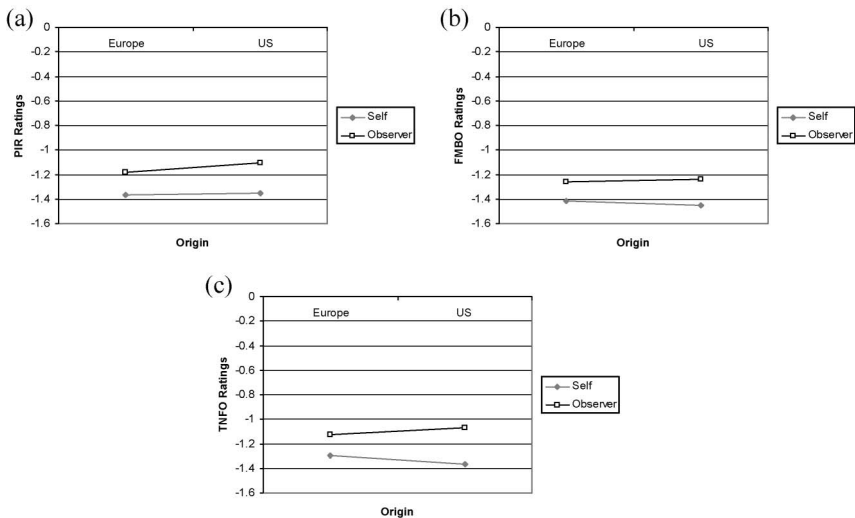


Figure 4. Self-observer discrepancies as a function of origin (Europe or United States).

are less negative or farther away from “-2” in magnitude than American managers). Also, observers rated European managers as less likely to display behaviours and characteristics that lead to derailment than observers of American managers (observer-ratings for European managers are more negative or closer to “-2” in magnitude than observer-ratings for American managers). Consequently, the self-observer discrepancy is wider for American managers than European managers, which means that self-observer ratings are more congruent for European managers than American managers, and the discrepancy is due to both self- and observer-ratings.

DISCUSSION

Managerial derailment is costly to managers, their co-workers, and the organization (Bunker et al., 2002; Finkelstein, 2004; Lombardo & McCauley, 1988; Smart, 1999; Wells, 2005). Knowing whether a discrepancy exists between what managers believe and what observers of managers believe about displaying behaviours and characteristics of managerial derailment may be an important warning signal for managers to prevent derailment. This discrepancy could be a potential and expensive “blind spot” if not dealt with in a timely manner.

In the past, derailment and failure have focused on senior leadership positions, high-ranking executives, or have been defined as an executive phenomenon (Dotlich & Cairo, 2003; Lombardo & Eichinger, 1989/2005). However, this research reveals that derailment could be a problem even for managers in lower organizational levels, reiterating Shipper and Dillard’s (2000) proposition. Furthermore, the present study’s results reaffirm and extend Ostroff et al.’s (2004) and Sala’s (2003) previous research by (a) showing that self-observer discrepancies existed and widened as managerial level increased; (b) focusing on a construct that had not been used before in self-observer discrepancy research, namely managerial derailment; and (c) utilizing a European sample, and comparing European managers to American managers.

The present study revealed that a discrepancy exists between self- and observer-ratings on the extent to which European managers display the behaviours and characteristics of managerial derailment. Moreover, the discrepancy widened as managerial levels increased and was mostly due to inflated self-ratings, supporting Hypothesis 1. The same pattern emerged across all five derailment clusters and across all observers. Knowing this pattern concerning self-observer discrepancies in derailment is important since Fletcher (1997) proposed that a manager’s contribution, adaptability to work, and relationships may be hindered when self perceptions are discrepant or differ from observer perceptions. The main trend in the

findings is that overinflated self-ratings may be a cause of the discrepancy at higher managerial levels. The impact of managerial level remains similar regardless of the perspective of the rater.

It therefore becomes important to focus on why managers' ratings are discrepant (i.e., different) from the ratings of their peers and direct report, especially at higher managerial levels. Future research needs to attempt to examine why discrepancies between self- and observer-ratings occur. Accordingly, others have discussed their thoughts as to why the self-observer discrepancy is wider for those at higher managerial levels; reasons include higher level managers not receiving honest feedback (Goleman et al., 2001) or fear of retribution (Dalton, 1997).

However, results of the present study tend to show that inflated self-ratings are the source for the self-observer discrepancy. Therefore, it may be more pertinent to focus less on observers and more on managers themselves. For instance, managers may not be realistic or accurate in their self-perception, and therefore their self-ratings, because they have never received feedback, or have never received specific negative feedback (Yammarino & Atwater, 2001). Also, those at higher managerial levels have more experience, which may lead to overconfidence about one's ability (Atwater & Yammarino, 1997), and less feedback seeking (Ashford, 1986). These reasons may be more plausible explanations for the discrepancy in self-observer ratings of derailment behaviours and characteristics, especially at higher managerial levels.

The findings were more complicated and mixed with the self-boss discrepancy (partially supporting Hypothesis 2). The self-boss discrepancy was the only rater group where the lines "crossed" such that middle-level managers believed they were *more* likely to display derailment behaviours and characteristics than the boss, but top-level managers believed they were *less* likely to display derailment behaviours and characteristics than the boss. One reason may be because bosses have a different view of managers who reside at different managerial levels. Alternatively, managers at different levels relate to their bosses in different ways or with a different level of frequency. Further research should examine why boss-ratings may be different from peer- or direct report-ratings.

The most accurate and honest feedback system will be of little use if managers disregard or deny information from multisource instruments. Recent research has demonstrated that superior (boss) and subordinate (direct report) ratings are most predictive of objective performance criteria (Sala & Dwight, 2002). If "merely a difference of perception" could influence external consequences such as rewards or promotion, more consideration may be given to self-observer discrepancies. This is especially important in light of the inconsistent findings regarding self-boss rating differences in the present study.

Finally, it is interesting to note that there appear to be differences in self–observer discrepancies between European and American managers, though results were mixed for each of the derailment clusters. For instance, there were no differences in the self–observer discrepancy between European and American managers in “difficulty leading a team” and “difficulty changing and adapting”. However, there were differences in self–observer discrepancy between European and American managers in “problems with interpersonal relationships”, “failure to meet business objectives”, and “too narrow functional orientation”. In all three cases, the self–observer discrepancy was “bigger/wider” for the American managers than the European managers. Analysing these differences further, the reasons for the discrepancies varied as well. For instance, the cause for the discrepancy in “problems with interpersonal relationships” was observer-ratings, while the cause for the discrepancy in “failure to meet business objectives” was self-ratings. Self- and observer-ratings were both influential in the self–observer discrepancy in “too narrow functional orientation”. This exploratory analysis should lead to more research regarding differences between specific European countries and American managers, particularly drawing on research regarding specific cultural nuances within Europe.

LIMITATIONS AND FUTURE RESEARCH

Like all research, this research is limited. The sample, which used practising European managers participating in a leadership development process, is one limitation. Managers selected for leadership development processes may not mirror practising managers in general. For instance, managers in a leadership development process may be a more high-performing or high-potential calibre of managers. Moreover, only managers that were native to and working in European countries were used. How to categorize a manager as “European” or not is a complex undertaking. Furthermore, managers in Europe undoubtedly vary in terms of individual characteristics as well as organizational and cultural contexts.

It would be beneficial for future research to not study all European managers as a whole, but rather, to study differences within and between European countries to bring more knowledge about self–observer discrepancies in a European context. This approach would allow for the exploration of hypotheses based on cultural categorizations identified in other research such as House, Hanges, Javidan, Dorfman, and Gupta (2004). Future research should attempt to determine the antecedents and consequences of self–observer discrepancies between European and American managers. In sum, similar studies using managers in other parts of Europe and the world are needed, especially since multisource

instruments are increasingly used in practice and in research around the world (Atwater et al., 2005; Brutus et al., 2001).

Another limitation could be the one-item measure of “likelihood of derailment”. Measures with one item may be limited as compared to longer measures, but may be easier for interpretation. The exploratory manner in which this analysis was conducted should be taken under consideration regarding this limitation. In addition, though the findings may be of statistical significance, one may question the practical significance. Finally, it should be noted that having or displaying the behaviours and characteristics of derailment does not equal actual derailment. Future research should observe managers who have actually derailed, and ascertain whether self-observer rating discrepancies in multisource instruments may give insight into why the manager derailed. Similarly, researchers could also examine managers who recovered from managerial derailment (Kovach, 1989; Shipper & Dillard, 2000).

IMPLICATIONS FOR PRACTICE

As noted previously, managerial derailment is a concern for individuals as well as organizations. This research illuminates an important dynamic related to ratings of derailment characteristics. In light of these findings, it is worth noting means by which managers can cope with or prevent derailment. This is especially important since the findings show that it is self-ratings (rather than observer-ratings) that influence self-observer discrepancies, particularly at higher managerial levels.

For instance, becoming aware of times that could prompt derailment is important (Dotlich & Cairo, 2003). In addition, becoming an active learner, concentrating more on leadership roles, being more tolerant of ambiguity, focusing on problem solving, and understanding one’s own interpersonal impact may help managers cope with or prevent derailment (Lombardo & Eichinger, 1989/2005).

Lombardo and Eichinger (1989/2005) also proposed that a manager must openly receive honest, constructive, developmental feedback. With this in mind, if managers are to receive this type of feedback, it may be best for managers to be more “learning-goal oriented” than “performance-goal oriented” and, hence, managers becoming more “learning-goal oriented” may be important in receiving and using their feedback effectively. VandeWalle and Cummings (1997) offered different self-regulation strategies to build or enhance a learning-goal orientation, as well as suggesting that managers should change attributions about ability and performance and have alternative sources of feedback other than face to face.

In addition, enhancing managerial self-awareness may be especially important given the findings of the present research, showing that

“overestimators” or managers whose ratings are discrepant from observers are those who may be more likely to derail in the future. In reality, self-awareness is a crucial feature of effective leadership (Konger & Benjamin, 1999). Much has been written on how to increase or enhance self-awareness in the applied setting. By reflecting on life-shaping moments (Van Velsor, Moxley, & Bunker, 2004), using executive coaches (Rosinski, 2003; Wales, 2003), or mentors (Couzins & Beagrie, 2005), taking personality tests (Baillie, 2004; Couzins & Beagrie, 2005; McCarthy & Garavan, 1999) and journaling (Loo, 2002; Van Velsor et al., 2004), a manager has several methods to enhance self-awareness. The relationship between use (or nonuse) of these methods and organizational level would be worth exploring in light of the current research findings. It is possible that as managers reach higher organizational levels, there are more opposing demands for their time, and their free time for reflection and engagement in practices such as those previously mentioned may be reduced.

REFERENCES

- Ashford, S. J. (1986). The role of information seeking in individual adaptation: A resource perspective. *Academy of Management Journal*, *29*, 465–487.
- Atwater, L. E., Ostroff, C., Yammarino, F. J., & Fleenor, J. W. (1998). Self–other agreement: Does it really matter? *Personnel Psychology*, *51*, 577–598.
- Atwater, L. E., Roush, P., & Fischthal, A. (1995). The influence of upward feedback on self and follower ratings of leadership. *Personnel Psychology*, *48*, 35–59.
- Atwater, L., & Waldman, D. (1998). Accountability in 360 degree feedback. *HR Magazine*, *43*, 96–104.
- Atwater, L., Waldman, D., Ostroff, C., Robie, C., & Johnson, K. M. (2005). Self–other agreement: Comparing its relationship with performance in the US and Europe. *International Journal of Selection and Assessment*, *13*, 25–40.
- Atwater, L. E., & Yammarino, F. (1997). Self–other rating agreement: A review and model. *Research in Personnel and Human Resources Management*, *15*, 141–164.
- Baillie, S. (2004). Know thyself. *Profit*, *23*, 107–108.
- Brutus, S., Fleenor, J. W., & London, M. (1998). Does 360-degree feedback work in different industries? A between-industry comparison of the reliability and validity of multi-source performance ratings. *Journal of Management Development*, *17*, 177–190.
- Brutus, S., Fleenor, J. W., & McCauley, C. D. (1999). Demographic and personality predictors of congruence in multi-source ratings. *Journal of Management Development*, *18*, 417–435.
- Brutus, S., Leslie, J. B., & McDonald-Mann, D. (2001). Cross-cultural issues in multisource feedback. In D. Bracken, C. Timmreck, & A. Church (Eds.), *Handbook of multisource feedback* (pp. 433–446). San Francisco: Jossey-Bass.
- Bunker, K. A., Kram, K. E., & Ting, S. (2002). The young and the clueless. *Harvard Business Review*, *80*, 80–87.
- Center for Creative Leadership. (2002). *BENCHMARKS[®] facilitator's manual*. Greensboro, NC: Center for Creative Leadership.
- Church, A. H. (1997). Managerial self-awareness in high-performing individuals in organizations. *Journal of Applied Psychology*, *82*, 281–292.

- Church, A. H. (2000). Do higher performing managers actually receive better ratings? A validation of multirater assessment methodology. *Consulting Psychology Journal: Practice and Research*, 52, 99–116.
- Conger, J. A., & Nadler, D. A. (2004). When CEOs step up to fail. *MIT Sloan Management Review*, 45, 50–56.
- Conway, J. M. (2000). Managerial performance development constructs and personality correlates. *Human Performance*, 13, 23–46.
- Conway, J. M., Lombardo, K., & Sanders, K. C. (2001). A meta-analysis of incremental validity and nomological networks for subordinate and peer ratings. *Human Performance*, 14, 267–303.
- Couzins, M., & Beagrie, S. (2005, May 17). How to . . . develop your self-awareness. *Personnel Today*, p. 31.
- Cronbach, L. J. (1958). Proposals leading to analytic treatment of social perceptions scores. In R. Tagiuri & L. Petrulo (Eds.), *Person perception and interpersonal behavior* (pp. 353–379). Stanford, CA: Stanford University Press.
- Cronbach, L. J., & Furby, L. (1970). How should we measure “change”—or should we? *Psychological Bulletin*, 74, 68–80.
- Dallessio, A. T. (1998). Using multi-source feedback for employee development and personnel decisions. In J. W. Smither (Ed.), *Performance appraisal: State-of-the-art in practice* (pp. 278–330). San Francisco: Jossey-Bass.
- Dalton, M. A. (1997). When the purpose of using multi-rater feedback is behavior change. In D. W. Bracken, M. A. Dalton, R. A. Jako, C. D. McCauley, & V. A. Pollman (Eds.), *Should 360-degree feedback be used only for developmental purposes?* (pp. 1–6). Greensboro, NC: Center for Creative Leadership.
- Dotlich, D. L., & Cairo, P. C. (2003). *Why CEOs fail: The 11 behaviors that can derail your climb to the top—and how to manage them*. San Francisco: Jossey-Bass.
- Douglas, C. A. (2003). *Key events and lessons for managers in a diverse workforce: A report on research and findings*. Greensboro, NC: Center for Creative Leadership.
- Edwards, J. R. (1994). The study of congruence in organizational behavior research: Critique and proposed alternative. *Organizational Behavior and Human Decision Processes*, 58, 51–100.
- Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes*, 64, 307–324.
- Edwards, J. R., & Cooper, C. L. (1990). The person–environment fit approach to stress: Recurring problems and some suggested solutions. *Journal of Organizational Behavior*, 10, 293–307.
- Edwards, J. R., & Parry, M. E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal*, 36, 1577–1613.
- Edwards, M., & Ewen, A. (1998). *Multisource assessment survey of industry practice*. Paper presented at the 360-Degree Feedback Global Users Conference, Orlando, FL.
- Finkelstein, S. (2004). *Why smart executives fail*. New York: Portfolio.
- Fleener, J. W., McCauley, C. D., & Brutus, S. (1996). Self–other rating agreement and leader effectiveness. *Leadership Quarterly*, 7, 487–506.
- Fletcher, C. (1997). Self awareness: A neglected attribute in selection and assessment. *International Journal of Selection and Assessment*, 5, 183–187.
- Gentry, W. A., Mondore, S. P., & Cox, B. D. (in press). An exploratory study of managerial derailment characteristics and personality preferences. *Journal of Management Development*.
- Goleman, D., Boyatzis, R., & McKee, A. (2001). Primal leadership: The hidden driver of great performance. *Harvard Business Review*, 79, 42–51.

- Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology, 41*, 43–61.
- Hogan, R., & Hogan, J. (2001). Assessing leadership: A view from the dark side. *International Journal of Selection and Assessment, 9*, 40–51.
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage.
- Kaplan, R. E., Drath, W. H., & Kofodimos, J. R. (1987). High hurdles: The challenge of executive self-development. *Academy of Management Executive, 1*, 195–205.
- Konger, J. A., & Benjamin, B. (1999). *Building leaders: How successful companies develop the next generation*. San Francisco: Jossey-Bass.
- Kovach, B. E. (1986). The derailment of fast-track managers. *Organizational Dynamics, 15*, 41–48.
- Kovach, B. E. (1989). Successful derailment: What fast-trackers can learn while they're off the track. *Organizational Dynamics, 18*, 33–47.
- Leslie, J. B., & Van Velsor, E. (1996). *A look at derailment today: North America and Europe*. Greensboro, NC: Center for Creative Leadership.
- Levinson, H. (1994). Beyond the selection failures. *Consulting Psychology Journal, 46*, 3–8.
- Lindsey, E. H., Homes, V., & McCall, M. W., Jr. (1987). *Key events in executives' lives*. Greensboro, NC: Center for Creative Leadership.
- Lombardo, M. M., & Eichinger, R. W. (2005). *Preventing derailment: What to do before it's too late*. Greensboro, NC: Center for Creative Leadership. (Original work published 1989).
- Lombardo, M. M., & McCauley, C. D. (1988). *The dynamics of management derailment* (Tech. Rep. No. 34). Greensboro, NC: Center for Creative Leadership.
- Lombardo, M. M., & McCauley, C. D. (1994). *BENCHMARKS[®]: A manual and trainer's guide*. Greensboro, NC: Center for Creative Leadership.
- Lombardo, M. M., McCauley, C. D., McDonald-Mann, D., & Leslie, J. B. (1999). *BENCHMARKS[®] developmental reference points*. Greensboro, NC: Center for Creative Leadership.
- London, M., & Smither, J. W. (1995). Can multi-source feedback change self-evaluations, skill development, and performance? Theory-based applications and directions for research. *Personnel Psychology, 48*, 375–390.
- Loo, R. (2002). Journaling: A learning tool for project management training and team-building. *Project Management Journal, 33*, 61–66.
- McCall, M., & Lombardo, M. (1983). *Off the track: Why and how successful executives get derailed*. Greensboro, NC: Center for Creative Leadership.
- McCarthy, A. M., & Garavan, T. N. (1999). Developing self-awareness in the managerial career development process: The value of 360-degree feedback and the MBTI. *Journal of European Industrial Training, 23*, 437–445.
- McCauley, C., & Lombardo, M. (1990). BENCHMARKS[®]: An instrument for diagnosing managerial strengths and weaknesses. In K. E. Clark & M. B. Clark (Eds.), *Measures of leadership* (pp. 535–545). West Orange, NJ: Leadership Library of America.
- McCauley, C., Lombardo, M., & Usher, C. (1989). Diagnosing management development needs: An instrument based on how managers develop. *Journal of Management, 15*, 389–403.
- Morgeson, F. P., Mumford, T. V., & Campion, C. A. (2005). Coming full circle: Using research and practice to address 27 questions about 360-degree feedback programs. *Consulting Psychology Journal: Practice and Research, 57*, 196–209.
- Morrison, A. M., White, R. P., & van Velsor, E. (1987). *Breaking the glass ceiling: Can women reach the top of America's largest corporations?* Reading, MA: Addison-Wesley.
- Mount, M. K., Judge, T. A., Scullen, T. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. *Personnel Psychology, 51*, 557–576.

- Ostroff, C., Atwater, L. E., & Feinberg, B. J. (2004). Understanding self-other agreement: A look at rater and rate characteristics, context, and outcomes. *Personnel Psychology*, *57*, 333–375.
- Recognizing the symptoms of reckless leadership. (2003). *Harvard Business Review*, *81*, 65.
- Rosinski, P. (2003). *Coaching across cultures: New tools for leveraging national, corporate, and professional differences*. London: Nicholas Brealey.
- Sala, F. (2003). Executive blind spots: Discrepancies between self- and other-ratings. *Consulting Psychology Journal*, *55*, 222–229.
- Sala, F., & Dwight, S. A. (2002). Predicting executive performance with multi-rater surveys: Whom you ask makes a difference. *Consulting Psychology Journal*, *54*, 166–172.
- Shipper, F., & Dillard, J. E., Jr. (2000). A study of impending derailment and recovery of middle managers across career stages. *Human Resource Management*, *39*, 331–345.
- Smart, B. (1999). *Topgrading: How leading companies win by hiring, coaching, and keeping the best people*. Upper Saddle River, NJ: Prentice Hall.
- Smither, J. W., & Walker, A. G. (2001). Measuring the impact of multisource feedback. In D. Bracken, C. Timmreck, & A. Church (Eds.), *Handbook of multisource feedback* (pp. 256–271). San Francisco: Jossey-Bass.
- Tornow, W. W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management*, *32*, 246–264.
- Tsui, A. S., & Ohlott, P. J. (1988). Multiple assessment of managerial effectiveness: Interrater agreement and consensus in effectiveness models. *Personnel Psychology*, *41*, 779–803.
- Van Velsor, E., Moxley, R. S., & Bunker, K. A. (2004). The leader development process. In C. McCauley & E. van Velsor (Eds.), *The Center for Creative Leadership handbook of leadership development* (2nd ed., pp. 204–233). San Francisco: Jossey-Bass/Center for Creative Leadership.
- Van Velsor, E., Ruderman, M., & Philips, A. (1991, April). *Enhancing self-objectivity and performance on the job: The developmental impact of feedback*. Paper presented at the sixth annual conference of the Society for Industrial and Organizational Psychology, St Louis, MO.
- VandeWalle, D., & Cummings, L. L. (1997). A test of the influence of goal orientation on the feedback-seeking process. *Journal of Applied Psychology*, *82*, 390–400.
- Wales, S. (2003). Why coaching? *Journal of Change Management*, *3*, 275–282.
- Wells, S. J. (2005). Diving in. *HR Magazine*, *50*, 54–59.
- Yammarino, F. (2003). Modern data analytic techniques for multisource feedback. *Organizational Research Methods*, *6*, 6–14.
- Yammarino, F., & Atwater, L. (2001). Understanding agreement in multisource feedback. In D. Bracken, C. Timmreck, & A. Church (Eds.), *Handbook of multisource feedback* (pp. 204–220). San Francisco: Jossey-Bass.
- Zedeck, S. (1995). Review of BENCHMARKS[®]. In J. Conoley & J. Impara (Eds.), *The twelfth mental measurements yearbook* (Vol. 1, pp. 128–129). Lincoln, NE: Buros Institute of Mental Measurements.

Original manuscript received May 2006

Revised manuscript received April 2007

First published online 16 July 2007