At the crux of dyadic leadership: Self–other agreement of leaders and direct reports — Analyzing 360-degree feedback ☆

Steven E. Markham [a,*], Ina S. Markham [b,1], Janice Witt Smith [c,2]

[a] Department of Management, Pamplin College of Business, Virginia Tech, Blacksburg, VA 24061, United States
[b] Department of Computer Information Systems & Business Analytics, College of Business, MSC, 0202, James Madison University, Harrisonburg, VA 22807, United States
[c] Department of Management and Marketing, College of Arts, Science, Business, & Education, Winston-Salem State University, Winston-Salem, NC 27110, United States

A R T I C L E I N F O

A B S T R A C T

The purpose of this research is to investigate the use of dyads instead of whole rating groups for analyzing 360° feedback. A Center for Creative Leadership sample of direct subordinates was matched with their focal leaders (J = 4810 dyads). Two CCL Benchmark scales (Building & Mending Relationships and Problems with Interpersonal Relationships) were used to determine if this dyadic pairing produced high self–other agreement (SOA). Using Within and Between Analysis (WABA), results suggest the viability of detecting unitary, whole dyadic effects only under specific boundary conditions. Other conditions produced highly dispersed dyadic results. Thus, dyadic convergence is not a universal phenomenon in these data. Tying the dyadic results back to the original rater group's consistency as reported in Markham, Smith et al. (2014), those rater groups with high internal agreement form clearer dyadic pairings when compared to rater groups that were moderate or low in agreement. Implications and future directions are discussed.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Multi-source feedback programs, be they 360, 270 or 180° in nature, are a very popular and essential tool for organizations world-wide that are engaged in formal leadership development programs. The juxtaposition of one's own perceptions compared to relevant others' perceptions on key leadership dimensions can be an eye-opening experience leading to personal change and transformation. These programs, however, beg the question of how to report such data in order to highlight relevant agreement or disagreement between the focal person and their raters. Should these data be organized and reported along the lines of all raters of the same type (peers or subordinates or others), thereby implying an average leadership style? Alternatively, should the data be examined by pairing a self-report with an individual rater report, thereby implying an individualized, vertical dyadic style? Given the large numbers of participants who experience a 360° multi-source feedback program each year, this question is timely and significant.

The purpose of this research is to expand our knowledge of fundamental data configuration issues confronting multi-source leadership feedback programs (Atwater & Waldman, 1998) by operationalizing the dyadic level of analysis in this context

---

* Corresponding author at: Department of Management (0233), Pamplin Hall, Room 2007 880 West Campus Drive, Blacksburg, VA 24061, United States. Tel.: +1 540 362 8776.
E-mail addresses: markhami@vt.edu (S.E. Markham), markhais@jmu.edu (I.S. Markham), drjwsmith@jwsmithconsulting.com (J.W. Smith).
[1] Tel.: +1 540 568 3056.
[2] Tel.: +1 919 649 4055.

(Yammarino, 2003). Thus, we will be comparing a focal respondent with each of his or her direct reports one at a time. This dyadic comparison could also apply to reports from peers, bosses, or other sources. Traditionally, multi-source feedback programs compare the responses of a focal role incumbent on various types of leadership behaviors with the average response of this focal person's direct subordinates. Similarly, average responses from peers, other organizational members, or outsiders (such as customers) (Lepsinger & Lucia, 2009) could be used. Again, each comparison of the target individual with the average group response would be done one at a time. These programs have been and continue to be very popular. For example, since 1987 the Center for Creative Leadership (CCL), one of the premier organizations specializing in this type of leadership development technique, has used their Benchmarks 360° instrument with more than 200,000 leaders embedded within 16,000 organizations (Van Velsor, McCauley, & Ruderman, 2010).

An inherent conundrum in all multi-source feedback techniques within leadership development programs is the Achilles' heel issue of aggregation (Markham, Smith, Markham, & Braekkan, 2014). In other words, should each rater group (composed of subordinates, peers, or others) for each focal participant have an average score calculated and reported? On one hand, the organizational reality is that feedback data must somehow be summarized and raters' identities protected. On the other hand, it is now clear that such aggregation can be problematic when rater groups have high internal variance resulting in inaccurate (or incomplete) information being returned to the focal respondent. According to Markham, Smith et al. (2014), depending upon the decision criteria utilized, up to 25% of rating groups should not have their scores averaged or reported. The purpose of this research is to demonstrate that for analyzing multi-source feedback reports an alternative level of analysis, the dyad, can be used either in conjunction with and/or instead of whole rater groups.

The opportunity to evaluate actual field data to investigate alternative solutions for this potential problem is worthwhile; especially when addressing the general question, do dyadic self–other pairings capture and describe this type of multi-source feedback data better than whole groups of raters? The examination of leadership processes that focus on dyads instead of whole groups has a long history (Dansereau, 1995; Dansereau et al., 1995). Similarly, the general analysis of dyadic data also has been recently recommended (Kenny, Kashy, & Cook, 2006) as has the notion that multi-level methods can be applied to multi-source feedback (Yammarino, 2003). However, dyadic studies using complete entity-based multi-level methods (Markham, Yammarino, Murry, & Palanski, 2010) are scarce (Gooty & Yammarino, 2011), and, to date, they have not been specifically applied to multi-source feedback data. A detailed explanation of the history, administration, and efficacy of these multi-source leadership feedback and development programs can be found in Markham, Smith et al. (2014).

Leadership derailment

We will be using the same model and subset of variables that were analyzed by Markham, Smith et al. (2014). These data were provided by CCL, and they are part of a larger set contained within the Benchmarks 360° Instrument that focuses on preventing managerial derailment (Leslie & Braddy, 2013). A derailment factor is any item or event that has been shown to side-track a manager's career such that he or she is no longer considered "hi-po" (high potential) by the organization (Zenger & Folkman, 2009). While such a manager is not often fired, they may be removed from an internal pool of candidates for promotion (Herbst, 2014; Thuraisingham, 2010).

Leadership derailment factors have become the focus of a dedicated literature (Braddy, Gooty, Fleenor, & Yammarino, 2014; Leslie & Braddy, 2013). This research stream attempts to identify the kinds of developmental activities and training necessary for individuals to attain senior-level positions. At CCL the Executive Derailment research project captured data from individuals who were successful and then compared this information to data from executives who had been only somewhat successful; in other words, they had been derailed. This comparison group had not been terminated, but they had not reached the organizational level to which they and their organization had expected them to reach. The Executive Derailment and follow-up studies provided qualitative data that were content analyzed resulting in sixteen categories of key events (Leslie & Braddy, 2013). Further studies have supported the investigation of derailment and have found that ratings from peers and subordinates, aside from direct reports, can be important in predicting the risk of derailment potential (Braddy et al., 2014).

We will be using two illustrative variables: (1) Building and Mending Relationships and (2) Problems with Interpersonal Relationships. The scale, Building and Mending Relationships, refers to the amount of time and energy spent by the subject in reaching out to and investing in his or her social support network. The second refers to the severity and difficulty of interpersonal problems faced by the focal respondent. Both variables are highly salient to the leader–direct report vertical relationship.

We have one potential control measure, age, which was also collected by CCL from both subjects and raters. Age attempts to account for any maturity effects such that older raters might have more experience in accurately filling out these types of instruments. Vecchio and Anderson (2009) found that gender and age were associated with self- and others' ratings as it relates to leadership effectiveness. Males and older managers were more likely to overestimate their own effectiveness compared to the ratings of others.

Effective leadership and self–other agreement (SOA)

In reviewing the history of self–other agreement (SOA), we note that researchers historically have examined the differences between self- and other raters as error variance that should either be reduced or eliminated. More recently, in this approach to dealing with SOA, we are faced with methodological, theoretical and practical challenges in interpretation when there is a difference within rater groups or dyads. Tornow (1993) suggested that these rater differences can provide both meaningful and useful