

McKinsey estimated that companies using social technologies such as those involved in social PM to increase communications about goals and collaboration in the achievement of goals could see a increase in productivity by knowledge workers of 20% to 25%.

Between the changing expectations of the workforce and the fact that organizations are more matrixed and more collaborative than ever before, organizations simply need new processes to reflect today's business realities. True, technology is only a tool. Used correctly, this technology helps facilitate the many changes in PM that are necessary for the process to be more relevant to the employee today and to help drive organizational performance in a dynamic business environment that requires to-the-moment responsiveness and frequent shifting of priorities.

References

- Chui, M., Manyika, J., Bughin, J., Dobbs, R., Roxburgh, C., Sarrazin, H., . . . Westergren, M. (2012). *The social economy: Unlocking value and productivity through social technologies*. London, UK: McKinsey Global Institute.
- Pulakos, E. D., Mueller Hanson, R., Arad, S., & Moye, N. (2015). Performance management can be fixed: An on-the-job experiential learning approach for complex behavior change. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 8, 51–76.

Seeing the Forest but Missing the Trees: The Role of Judgments in Performance Management

John P. Meriac

University of Missouri—St. Louis

C. Allen Gorman

East Tennessee State University

Therese Macan

University of Missouri—St. Louis

Various solutions have been proposed to “fix” performance management (PM) over the last several decades. Pulakos, Mueller Hanson, Arad, and

John P. Meriac, Department of Psychological Sciences, University of Missouri—St. Louis; C. Allen Gorman, Department of Management, East Tennessee State University; Therese Macan, Department of Psychological Sciences, University of Missouri—St. Louis.

Correspondence concerning this article should be addressed to John P. Meriac, Department of Psychological Sciences, University of Missouri—St. Louis, 425 Stadler Hall, One University Boulevard, St. Louis, MO 63121-4499. E-mail: meriac@umsl.edu

Moye (2015) have presented a holistic approach to improving PM in organizations. Although this approach addresses several key elements related to the social context of PM, namely the buy-in of organizational stakeholders, timely and regular feedback, and future-directed feedback, we believe that several robust findings from the PM research literature could further improve this process. Are Pulakos et al. looking at the forest but missing the trees? In the following commentary, we offer several reasons that performance judgments and perhaps even informal ratings are still operating and occurring in the proposed holistic system. Therefore, advancements in other areas of PM research may offer additional ways to fix PM.

The Case for Facilitating Performance Judgments

We agree with Pulakos et al. (2015) that the typical rating process conducted by managers can be burdensome and time consuming, but the alternatives to not making ratings are unclear. Several unanswered questions remain about the proposed holistic system and how it could operate. Do managers still make performance judgments even if formal ratings are not assigned? Many of the authors' suggestions and proposed solutions hinge on performance judgments, and the extent that these are not ratings in the broader sense is unclear.

Pulakos et al. (2015) have emphasized the use of “meaningful performance standards” (p. 16). However, these recommended performance standards appear to be nothing more than objective measurements, which have their own well-known problems specifically related to contamination and deficiency (Murphy & Cleveland, 1995). Despite the apparent purity of objective indices, environmental and situational factors outside the employee's control have the capacity to influence outcomes. Pulakos et al. (2015, p. 16) have gone on to suggest that “performance effectiveness indicators” should be developed for situations in which objective data are not available—but there is no recommendation on how this would be conducted or how the work would be evaluated. Peer reviews and customer feedback have been suggested by the authors as potential examples, but are these likely to be superior to supervisor ratings? Will these still then be subjective performance judgments and simply shift the same, well-documented problems with supervisor ratings to another source? Thus, even if we do not call the performance judgments ratings, it is unclear how the proposed solutions in the focal article would serve to improve on the performance judgments that would be “abandoned” in the proposed system.

Researchers have long recognized that it is difficult to justify the use of performance judgments if those assessments have little to no relationship to the characteristics of the people being appraised (Ilgen, Barnes-Farrell, & McKellin, 1993). Unfortunately, however, the research literature clearly

indicates that performance judgments are rarely sufficiently reliable and valid indicators of employee performance. Raters in PM systems are routinely subjected to a multitude of contextual influences, including cognitive biases; differences in rating goals, purposes, and motivation; and political and organizational influences; these can influence the quality of performance ratings (Levy & Williams, 2004; Murphy & Cleveland, 1995). The proposed holistic system offers no solution to this key issue, although performance judgments are still being made. Fortunately, PM researchers have designed several successful interventions for improving the relevance and quality of performance judgments. Below, we summarize a few of these advancements that can be integrated into a successful PM system.

Rater Training and Improvements in Judgments

Rater training represents one of the most robust interventions that can influence the quality of a PM system. Although a broader focus on all stakeholders is critical to consider, the raters (typically managers) still represent important (if not the most important) sources of information in the system. The importance of the rater is central to all PM systems (Murphy & Cleveland, 1995; Pulakos & O’Leary, 2011). However, performance judgments made by these individuals, whether these are formal ratings or some other type of evaluation, hinge on the proper observation, categorization, and subsequent scaling of behaviors. In the absence of strategies to align managers in this process, performance judgments could be more problematic than the formal performance ratings that would be abandoned. Pulakos et al. (2015) have touched on this point by emphasizing a “shared mindset about what effective PM is” (p. 17). However, the solutions presented in creating this shared mindset ignore key developments in this regard from the PM literature.

In those cases in which Pulakos et al.’s (2015) alternative framework is adopted or has been adopted (e.g., Cargill), rater training may serve as a viable means of facilitating the shared mindset they have suggested. Specifically, frame-of-reference training (Bernardin & Buckley, 1981) has generally been demonstrated as the most effective approach for improving rater accuracy (Roch, Woehr, Mishra, & Kieszczynska, 2012; Woehr & Huffcutt, 1994). The goal of this rating approach is to train raters to share a common schema or to use a common cognitive knowledge structure regarding performance-related behaviors and their subsequent scaling with respect to effective and ineffective levels of performance (Athey & McIntyre, 1987; Bernardin & Buckley, 1981; Gorman & Rentsch, 2009; Noonan & Sulsky, 2001; Schleicher & Day, 1998). In other words, frame-of-reference training has been empirically demonstrated as an effective strategy for fostering a shared mindset for making performance judgments. Although we agree that frequent communication is beneficial, the authors of the focal article have

suggested that only informal discussions should take place. A formal rater training process could perhaps better facilitate this process.

We agree with Pulakos et al. (2015) that focusing on behaviors that matter is critical in a PM system. Toward that end, rater training from other avenues can also serve to improve the quality of ratings. Behavioral specificity (as opposed to making inferences) has been regarded as one of the key benefits in the assessment center (AC) method, which is an integral component of rater training (Thornton & Rupp, 2006). This focus on observing behaviors helps facilitate the use of ACs for developmental purposes in general and for valuable diagnostic feedback in particular. Research has demonstrated that rater training in one context can transfer to another. For example, Macan et al. (2011) found that managers who served as raters in an AC context (i.e., assessors) provided more specific behaviors in performance appraisals compared with managers who did not serve as assessors. The difference was attributable to the extensive training that was a component of the AC process—in this case, frame-of-reference training. Moreover, a focus on behavior will improve the reliability of performance ratings. Lievens (2001), for example, found that a frame-of-reference training condition produced interrater reliability estimates of at least .80 or greater in a sample of both students and managers across three dimensions of performance. These estimates were all larger than were those from control training and data-driven training conditions. Thus, rater training appears to be one means to foster the shared mindset that the authors of the focal article have suggested.

Streamlining Rating Processes

Rater training also helps address a point that Pulakos et al. (2015) have raised—specifically the downsides of streamlining ratings. Frame-of-reference training specifically aims to form or change the rater's schema to facilitate the rating or judgment process. Research from the training and development literature indicates that as trainees move from novices to experts, they become better able to recognize and evaluate new performance information because their schemas become more pattern oriented and more highly integrated, and information is stored in larger chunks (Cannon-Bowers, Tannenbaum, Salas, & Converse, 1991). To the extent that raters develop a shared mindset, this not only has the potential to minimize idiosyncratic ratings but also has the potential to communicate information more consistently across managers and over time.

Decades earlier, rating forms were regularly examined as potential solutions to improve the PM process. However, given their minimal improvement on rater errors, these approaches were largely abandoned (Landy & Farr, 1980). Unfortunately, the early rating-format research paradigm was based almost entirely on a false premise: that rating errors were actually

errors. It is now widely recognized that rater “errors” are poor indicators of the quality of rating formats (Fiscaro, 1988; Murphy, 2008; Murphy & Balzer, 1989; Nathan & Tippins, 1990). Indeed, recent research has demonstrated that the form does matter and can improve the reliability, validity, and factor structure of ratings as well as rater accuracy (Borman et al., 2001; Goffin, Jelley, Powell, & Johnston, 2009; Hoffman et al., 2012; Kane & Woehr, 2006). To be clear, we do not suggest that rater training or rating forms should be emphasized over the social context of the PM system but rather suggest that the measurement tools and the processes used to create the shared mindset among managers can facilitate the effective implementation of such a holistic system.

Back Into the Forest

We completely agree with several of Pulakos et al.’s (2015) key points related to regular feedback and that more frequent attempts to improve performance are beneficial for employees. However, meaningful feedback that has the potential to improve employee performance relies on reliable and valid performance judgments. The proposed holistic system hinges on the assumption that the core problems in traditional PM systems are attributable to the task of making ratings and evaluating performance on an annual basis, not information processing. To the extent that we aim to fix PM, let us not ignore the research advancements related to the facilitation of higher quality performance ratings.

References

- Athey, T. R., & McIntyre, R. M. (1987). Effect of rater training on rater accuracy: Level of processing theory and social facilitation theory perspectives. *Journal of Applied Psychology*, *72*, 567–572. doi:10.1037/0021-9010.72.4.567
- Bernardin, H. J., & Buckley, M. R. (1981). Strategies in rater training. *Academy of Management Review*, *6*, 205–212. doi:10.5465/AMR.1981.4287782
- Borman, W. C., Buck, D. E., Hanson, M. A., Motowidlo, S. J., Stark, S., & Drasgow, F. (2001). An examination of the comparative reliability, validity, and accuracy of performance ratings made using computerized adaptive rating scales. *Journal of Applied Psychology*, *86*, 965–973. doi:10.1037/0021-9010.86.5.965
- Cannon-Bowers, J. A., Tannenbaum, S. L., Salas, E., & Converse, S. A. (1991). Toward an integration of training theory and technique. *Human Factors*, *33*, 281–292.
- Fiscaro, S. A. (1988). A reexamination of the relation between halo error and accuracy. *Journal of Applied Psychology*, *73*, 239–244. doi:10.1037/0021-9010.73.2.239
- Goffin, R. D., Jelley, R. B., Powell, D. M., & Johnston, N. G. (2009). Taking advantage of social comparisons in performance appraisal: The relative percentile method. *Human Resource Management*, *48*, 251–268.
- Gorman, C. A., & Rentsch, J. R. (2009). Evaluating frame-of-reference rater training effectiveness using performance schema accuracy. *Journal of Applied Psychology*, *94*, 1336–1344. doi:10.1037/a0016476

- Hoffman, B. J., Gorman, C. A., Blair, C. A., Meriac, J. P., Overstreet, B., & Atchley, E. K. (2012). Evidence for the effectiveness of an alternative multi-source performance rating methodology. *Personnel Psychology*, 65, 531–563. doi:10.1111/j.1744-6570.2012.01252.x
- Ilgen, D. R., Barnes-Farrell, J. L., & McKellin, D. B. (1993). Performance appraisal process research in the 1980s: What has it contributed to appraisals in use? *Organizational Behavior and Human Decision Processes*, 54, 321–368. doi:10.1006/obhd.1993.1015
- Kane, J. S., & Woehr, D. J. (2006). Performance measurement reconsidered: An examination of frequency estimation as a basis for assessment. In W. Bennett Jr., C. E. Lance, & D. J. Woehr (Eds.), *Performance measurement: Current perspectives and future challenges* (pp. 77–110). Mahwah, NJ: Erlbaum.
- Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin*, 87, 72–107.
- Levy, P. E., & Williams, J. R. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management*, 30, 881–905. doi:10.1016/j.jm.2004.06.005
- Lievens, F. (2001). Assessor training strategies and their effects on accuracy, interrater reliability, and discriminant validity. *Journal of Applied Psychology*, 86, 255–264. doi:10.1037/0021-9010.86.2.255
- Macan, T., Mehner, K., Havill, L., Roberts, L., Heft, L., & Meriac, J. P. (2011). Two for the price of one: Assessment center training to focus on behaviors can transfer to performance appraisals. *Human Performance*, 24, 443–457. doi:10.1080/08959285.2011.614664
- Murphy, K. R. (2008). Explaining the weak relationship between job performance and ratings of job performance. *Industrial and Organizational Psychology: Perspectives on Research and Practice*, 1, 148–160. doi:10.1111/j.1754-9434.2008.00030.x
- Murphy, K. R., & Balzer, W. K. (1989). Rater errors and rating accuracy. *Journal of Applied Psychology*, 74, 619–624.
- Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social, organizational, and goal-based perspectives*. Thousand Oaks, CA: Sage.
- Nathan, B. R., & Tippins, N. (1990). The consequences of halo “error” in performance ratings: A field study of the moderating effect of halo on test validation results. *Journal of Applied Psychology*, 75, 290–296. doi:10.1037/0021-9010.75.3.290
- Noonan, L. E., & Sulsky, L. M. (2001). Impact of frame-of-reference and behavioral observation training on alternative training effectiveness criteria in a Canadian military sample. *Human Performance*, 14, 3–26. doi:10.1207/S15327043HUP1401_02
- Pulakos, E. D., Mueller-Hanson, R., Arad, S., & Moye, N. (2015). Performance management can be fixed: An on-the-job experiential learning approach for complex behavior change. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 8, 51–76.
- Pulakos, E. D., & O’Leary, R. S. (2011). Why is performance management broken? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 4, 146–164. doi:10.1111/j.1754-9434.2011.01315.x
- Roch, S. G., Woehr, D. J., Mishra, V., & Kieszczynska, U. (2012). Rater training revisited: An updated meta-analytic review of frame-of-reference training. *Journal of Occupational and Organizational Psychology*, 85, 370–395. doi:10.1111/j.2044-8325.2011.02045.x
- Schleicher, D. J., & Day, D. V. (1998). A cognitive evaluation of frame-of-reference rater training: Content and process issues. *Organizational Behavior and Human Decision Processes*, 73, 76–101. doi:10.1006/obhd.1998.2751

- Thornton, G. C., III, & Rupp, D. E. (2006). *Assessment centers in human resource management*. Mahwah, NJ: Erlbaum.
- Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance appraisal: A quantitative review. *Journal of Occupational and Organizational Psychology*, 67, 189–205. doi:10.1111/j.2044-8325.1994.tb00562.x

Informal and Formal Performance Management: Both Are Needed

Robert L. Cardy

University of Texas at San Antonio

Dissatisfaction with performance management (PM) has had a long history. Managers and employees alike have frustrations with the system, and numerous calls for the elimination of performance appraisal have been made over the years (e.g., Scholtes, 1999). The dissatisfaction and calls for elimination have created pressure for change in the practice of PM, and I applaud the focus on feedback and coaching that Pulakos, Mueller Hanson, Arad, and Moye (2015) have proposed. Providing feedback and being actively involved in the PM process would seem to be a key part of the job of managers, yet many managers are uncomfortable addressing this central task. Focusing efforts in our field on improving the skills of managers and helping managers become effective coaches can play a key role in improving PM. The need for improvement in the informal process, however, does not mean that the formal process is not needed. Pulakos et al. have suggested streamlining the formal PM system as much as possible, with particular emphasis on the possibility of eliminating performance ratings. It is argued here that the formal PM system still serves important purposes. It is also argued that a balance between the informal and formal aspects of the PM system needs to be maintained. These two issues are addressed below.

The Need for Formal Performance Management

The formal PM system serves a number of purposes and has some positive effects. The intent here is to focus on only some of the major uses and effects.

Robert L. Cardy, Department of Management, University of Texas at San Antonio.

Correspondence concerning this article should be addressed to Robert L. Cardy, Department of Management, University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249–0634. E-mail: robert.cardy@utsa.edu