Group Based Performance Talent Reviews:
Advantages, challenges and potential improvements

Steven T. Hunt, Ph.D. (with Ariel Roberts, Doug Haaland, Paul Karavis, Paul Levy, & Sylvia Roche)
Click to follow my posts: www.linkedin.com/in/steventhunt

All employees do not contribute equally to company performance (Boudreau & Ramstad, 2005, Viswesvaran & Ones, 2000). The contributions of high performing employees can be several orders of magnitude greater than contributions of "average" employees (O'Boyle & Aguinis, 2012). And companies that treat employees differently considering their relative contributions significantly outperform companies that treat employees as though they all provide equal value (Bloom & Van Reenan, 2007). Recognizing that some employees contribute more than others is also good for workers. Use of consistent, transparent methods to assess and reward employee contributions is a key factor influencing people's perceptions of justice, fairness, and equity (Colquitt, Conlon, Wesson, Porter, & Yee,2001).

Managing differences in employee contributions is critical to maximizing company performance and ensuring employee equity. But rating employee performance contributions is one of the most sensitive and difficult areas of human capital management (Aguinis et al., 2011; Austin & Villanova, 1992). The traditional approach of having individual managers rate employee performance has significant problems due to implicit biases associated with gender, race and other non-job relevant characteristics (Heilman & Chen, 2005; Koch, D'Mello, & Sacket, 2015; McKay & McDaniels, 2006; Stauffer & Buckley, 2005). Manager ratings are also constrained by the limited perspective managers have into the contributions of employees (Lance, Hoffman, Gentry, & Baranik, 2008; Scullen et al., 2000).

Concerns with individual manager ratings have spurred increased use of group-based methods where organizational stakeholders meet to collectively discuss, evaluate, and compare employee contributions. Frequently referred to as "talent reviews" or "calibration sessions", group-based methods have been widely adopted by organizations to rate employee performance and potential (Hasting, 2012). They are typically used to identify critical talent, guide staffing and development actions, and inform compensation decisions.   It is important to note that group-based talent reviews do not require nor necessarily involve the use of numeric ratings or forced distributions. Their only requirement is that stakeholder meet to discuss and compare the relative contributions of different employees.   Although most talent review processes do generate some form of employee categorization or rating.

Despite the widespread use of group-based reviews, very few academic research studies have examined group-based performance evaluations. A review of performance management research conducted between 1980 and 2017 did not identify any studies that looked at group-based talent reviews as they are performed in many organizations (Schleicher et al., 2018). To bridge this gap between science and practice, we held a session at the 2019 SIOP Conference that brought together practitioners and academics to discuss advantages, challenges and possible improvement to group-based talent reviews. The following summarizes observations and insights that came out of this panel

Advantages of Group Based Talent Reviews
There are several reasons why companies use talent reviews instead of individual manager ratings. Rather than relying on the opinions and perceptions of a single manager, talent reviews incorporate the opinions of multiple people with different perspectives to ensure that employees are assessed based on their actual behaviors, skills and accomplishments. Most talent reviews are used to facilitate dialogue

©Steven T. Hunt, www.steventhunt.com

that challenges and helps managers to differentiate between outstanding, solid, and below average contributors. This dialogue promotes consistency across managers in terms of how they define effective vs. ineffective performance. The dialogue that occurs during group talent reviews also provides organizational leaders with insight into workforce strengths, accomplishments and capabilities.

Academic research investigating individual versus group decision making has shown that on intellectual tasks (tasks with correct answers), groups perform better than individuals (e.g., Hastie, 1986; Kerr, MacCoun, & Kramer, 1996; McGrath, 1984; Vollrath, Sheppard, Hinsz, & Davis, 1989). Joshi et al (2015) also suggest that group-based calibration meetings could reduce gender-based biases in performance and compensation decisions. A few lab studies have also specifically investigated individual versus group ratings in a performance evaluation context (Martell & Borg, 1993; Martell & Leavitt, 2002; Roch, 2006, 2007). These studies found that groups of raters have better behavioral accuracy, rating accuracy, more motivation and are less influenced by feedback cues than individual raters. Although these studies were conducted with undergraduate raters who watched taped performance, they align with the experiences practitioners on this panel have had with high stakes talent reviews used to assess contributions of employees.

Challenge and Potential Improvements to Group Based Talent Reviews
The practitioners on the panel identified five specific challenges frequently encountered when overseeing group-based talent reviews.   Each is discussed below, along with suggestion made during the panel discussion on how they might be addressed.

Challenge: Distinguishing between performance and impact
Talent reviews often involve comparing employees who work at the same organization but who hold very different jobs. For example, comparing the Director of Marketing to the Director of Information Technology. This can create confusion whether employees are being evaluated based on how effectively they perform their role or based on how much impact their role has on the overall performance of the organization.   This issue often arises when people start posing questions like "how critical is this person to the company?" or "what would happen if this person left the organization?".
Potential Solutions
- Develop criteria to define performance using behavioral definitions (e.g. competencies). If possible, also include some method to evaluate objective criteria (e.g. performance relative to specific, job-based key performance indicators). Use these criteria to keep the discussion focused on what people are doing within the scope of their role. Avoid making wholistic judgments about a person's overall value to the company.
- Evaluate the impact of different jobs using methods such as those used to identify pivotal roles (Boudreau & Ramstad, 2007). Share this data during the talent review so that a) raters know that impact is being considered as part of the evaluation process, and b) raters can better differentiate between ratings of impact and ratings of performance.

Challenge: Incorporating information and feedback from people outside of the session
The data about employees that is discussed during talent reviews is largely limited to whatever information exists in the heads of the people participating the review itself. Leaders in many companies often have 15 or more direct reports and may have very limited information about certain employees. Similarly, as organizational restructuring becomes more commonplace, many leaders may have only worked with employees for a very limited time.   This raises serious concerns that employees could be assessed based on very limited and incomplete data.
Potential Solutions

- Explore the use of 360 survey technology and similar tools to gather information about employees from their colleagues and customers prior to the session.   Note that this does raise questions about the willingness of employees to provide feedback and potential accuracy concerns related to using feedback for evaluation instead of development.
- Consider incorporating data from continuous performance feedback technology. This idea also raises concerns since this technology is usually designed to support coaching. Specifically, are their ways to use coaching data for evaluation without damaging its use as a tool for development?
- Designate individuals to act as "advocates" for employees during talent reviews. Ensure each employee assessed during the talent review is represented by someone who has taken time to understand their contributions and performance. This person might be the employee's manager or could be a designated employee advocate who comes from another role in the company.

Challenge: Paying for performance

A common reason for conducting talent reviews is to guide compensation decisions. There is considerable argument around the effectiveness of "pay for performance". To quote Gerhart & Fang (2015), "pay for performance when it works seems capable of producing spectacularly good results. And when it does not work, it can likewise produce spectacularly bad results". Despite these concerns, the practitioners on the panel observed that virtually all CEOs believe in the concept of pay for performance in one form or another. The challenge is figuring out how to do it well.   Specific questions in need of answers in this area include determining how much difference should there be in the pay given to average performers versus high performers? What is the optimal level of differentiation in compensation levels? Also, how can companies avoid problems were managers change their performance ratings purely to justify making pay decisions? For example, rating someone lower one year because they got a high pay increase the previous year.

Potential Solutions

- Ensure employees understand how their performance has influenced their pay. Even small differences in compensation increases (e.g. 2% vs 2.5%) can significantly influence employee motivation if employees know they are receiving more than their peers and they understand how the decision was influenced by their performance (Schaubroeck et al, 2007). Make sure managers are trained on how to effectively explain compensation decisions to their direct reports.
- Treat performance as one factor among many that influences pay. Factors such as internal equity or retention risk also influence compensation decisions. Ensure your process gives managers some leeway to account for these factors.   For example, tie performance ratings to broader pay ranges. This allows managers to adjust pay up or down based on non-performance relevant variables, but still retains a link to performance.
- "Pay dispersion" refers to differences in pay tied to performance. If pay dispersion is too low then differences in pay may be viewed as meaningless, but if it is too high then differences can hurt performance by creating anxiety and negative competition (Shaw, 2014). The challenge is what might be effective pay dispersion for one population of employees could be ineffective for another. The only way to determine the right level of pay dispersion may be to conduct research measuring how employees in specific jobs or organizations react to different levels of pay.

Challenge: Increasing accountability for accurately assessing employees

Some managers in companies fail to put adequate time and effort into the assessment process. For example, they may rate all employees as being similar, so they do not have to have difficult conversations that require explaining why some employees are perceived as more valuable than others. Or they may make quick assessments based on relatively superficial reviews of employee data with little

apparent concern about making sure their judgments are accurate. What actions can be taken to increase the time and attention managers put into ensuring they are accurately assessing employees during talent reviews?

Potential Solutions

- Clearly communicate to managers that they will be required to justify their assessments of employees. Inform them what sort of criteria they will be expected to provide to explain their decisions and provide them training on how to prepare for the talent review.   Research suggests raters put more effort into preparing for talent reviews if they know in advance that they will have to explain their ratings to their peers (Roche, 2007).
- Reward managers based on the outcomes of talent reviews. For example, consider compensating managers for hiring and developing high potential employees who are transferred into other roles. Similar to how employees are rewarded for referring successful job candidates.
- Critically review reasons why managers might avoid making accurate assessments. For example, if a manager identifies someone as high potential and that person is promoted out of their team, how does this benefit the manager? Is the manager rewarded for developing talent and given resources to hire a new team member? Or are they punished by not receiving headcount to backfill the role?

Challenge: Assessing all employees, not just those at the extremes.

A typica process for conducting talent reviews is to start by discussing employees who are felt be making the highest or lowest levels of contributions. Then to use these conversations to anchor the discussion of those employees closer to the middle of the distribution. A common problem with this approach is employees in the middle of the performance curve often get little to no attention. This raises concerns about the impact of overlooking or downplaying the importance of talking about "average" employees.

Potential Solutions

- Rather focusing on the extremes first, use a process where a high performing employee is discussed, then a low performing employee, than an average employee. Following this sort of cycle will lessen the problem of overlooking average performers, but it could also lengthen the time needed to conduct talent review sessions.
- Educate the stakeholders participating in the talent review around the importance of "managing the middle" of the performance curve. Caution against problems that occur when employees get lost in the middle of the pack and do not receive adequate recognition and feedback.

The challenges discussed above represent a small portion of all the topics that came up during the panel session. It is clear organizations would benefit from additional research into talent reviews. This includes looking at the dynamics that occur during the actual review meetings. For example, how do different meeting processes and facilitation methods impact the discussion and assessments that occur during talent reviews?   We see this panel as an early step in what hopefully will become a major research area within the field of industrial-organizational psychology.

References

Adler, S., Campion, M., Colquitt, A., & Grubb, A. (2016). Getting rid of performance ratings:

Genius or folly? *Industrial and Organizational Psychology, 9*(2), 219-252.

Aguinis, H., Joo, H. & Gottfredson, R.K. (2011). Why we hate performance management –

and why we should love it. *Business Horizons*, 54, 503-507.

Austin, J,T. & Villanova P. (1992). The criterion problem: 1917-1992. *Journal of Applied

Psychology*, 77, 836-874.

Bloom, N., & Van Reenen, J. (2007). Measuring and Explaining Management Practices

Across Firms and Countries. *Quarterly Journal of Economics*, 122, 1341–1408.

J. W. Boudreau and P. M. Ramstad. (2007). *Beyond HR: The New Science of Human Capital*.
Boston: Harvard Business School Press

Colquitt, J., Conlon, D., Wesson, M., Porter, C., & Yee, K. Justice at the millennium: A meta-

analytic review of 25 years of organizational justice research. *Journal of Applied

Psychology, Vol 86*(3), 425-445.

Deaux, K., & Emswiller, T. (1974). Explanations of successful performance on sex-linked

tasks: What is skill for the male is luck for the female. *Journal of Personality and

Social Psychology, 29*(1), 80-85.

Gerhart, B. & Fang, M. (2015). Pay, Intrinsic Motivation, Extrinsic Motivation, Performance,
and Creativity in the Workplace: Revisiting Long-Held Beliefs. The Annual Review of
Organizational Psychology & Organizational Behavior, 489–521.

Glover, D., Pallais, A., Pariente, W. (2016). Discrimination as a self-fulfilling prophecy:

Evidence from French grocery stores. *Quarterly Journal of Economics.*

Hastie, R. (1986). Review essay: Experimental evidence on group accuracy. In B. Grofman &

G. Guillermo (Eds.), *Information pooling and group decision making* (Vol. 2, pp. 129–

157). Greenwich, CT: JAI Press

Hastings, R. (2011, December 12). Survey: Most large companies calibrate performance.

    Retrieved from https://www.shrm.org/resourcesandtools/hr-topics/employee-

    relations/pages/calibrateperformance.aspx

Heilman, M.E. & Chen, J.J. (2005). Same behavior, different consequences: Reactions to

    men's and women's altruistic citizenship behaviors. *Journal of Applied Psych, 90*(3),

    431-441.

Joshi, A., Son, J., & Roh, H. (2015). When can women close the gap? A meta-analytic
test of sex differences in performance and rewards. *Academy of Management
Journal*, *58*(5), 1516-1545.

Keeping, L. M., & Levy, P. E. 2000. Performance appraisal reactions: Measurement,

    modeling, and method bias. *Journal of Applied Psychology*, 85(5): 708–723.

Kerr, N. L., MacCoun, R. J., & Kramer, G. P. (1996). Bias in judgment: Comparing individuals

    and groups. *Psychological Review, 103*, 687–719

Koch, A.J., D'Mello, S.D., & Sackett, P.R. (2015). A meta-analysis of gender stereotypes and

    bias in experimental simulations of employment decision making. *Journal of Applied*

    Psychology, 100(1), 128-161.

Lance, C. E., Hoffman, B. J., Gentry, W., A., & Baranik, L. E. (2008). Rater source factors

    represent important subcomponents of the criterion construct space, not rater bias.

    Human Resource Management Review, 18(4), 223-232.

Levy, P.E., Cavanaugh, C.M, Frantz, N.F., Borden, L.B., & Roberts, A.A. (2018). Revisiting the

    social context of performance management: Performance appraisal effectiveness.

    Chapter to appear in N. Anderson, C. Viswesvaran, H.K. Sinangil,& D. S. Ones

(Eds.).*Handbook of Industrial, Work, and Organizational Psychology*, 2nd Edition, London: Sage Publications.

McGrath, J. E. (1984). *Groups: Interaction and performance.* Englewood Cliffs, NJ: Prentice-Hall.

Martell, R. F., & Borg, M R. (1993). A comparison of the behavioral rating accuracy of groups and individuals. *Journal of Applied Psychology, 78*, 43-50.

Martell, R. & Leavitt, K. N. (2002). Reducing the performance-cue bias in work behavior ratings: Can groups help. Journal of Applied Psychology, 86, 1032-1041.

McKay, P. F. & McDaniel, M. A. (2006). A reexamination of black-white mean differences in work performance: More data, more moderators. *Journal of Applied Psychology, 73*, 91, 538-554.

O'Boyle, E., & Aguinis, H. (2012).   The best of the rest: revisiting the norm of normality and individual performance. *Personnel Psychology, 65,* 79-119.

Roch, S. G. (2006). Performance evaluations conducted by rater teams: Increasing behavioral and rating accuracy. *Human Performance, 19,* 91-115.

Roch, S. G. (2007). Why convene rater teams: Investigation of the benefits of anticipated discussion, consensus, and rater motivation. Organizational Behavior and Human Decision Processes, 104, 14-29.

Schaubroeck, J., Shaw, J. D., Duffy, M. K., & Mitra, A. (2008). And under-met and over-met expectations model of employee reactions to merit raises. *Journal of Applied Psychology, 93(2),* 424-434.

Shaw, J.D. (2014). Pay dispersion. The Annual Review of Organizational Psychology and Organizational Behavior, 521–44.

J., Shaw, J. D., Duffy, M. K., & Mitra, A. (2008). And under-met and over-met expectations model of employee reactions to merit raises. *Journal of Applied Psychology, 93(2),* 424-434.

Schleicher, DJ, Baumann, HM, Sullivan, DW, Levy, PE, Hargrove, DC & Barros-Rivera, BA (2018). Putting the system into performance management systems: a review and agenda for performance management research. *Journal of Management, 44,* 2209-2245.

Scullen, SE., Mount, MK, & Goff, M (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology, 85*, 956-971.

Scullen, S., Bergey, P., & Aiman-Smith, L. (2005). Forced distribution rating systems and the improvement of workforce potential: A baseline simulation. *Personnel Psychology, 59,* 1-32.

Stauffer, J.M. & Buckley, M.R. (2005). The existence and nature of racial bias in supervisory ratings. *Journal of Applied Psychology, 90*(3), 586-591

Viswesvaran, C. & Ones, D. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment, 8,* 216-228.

Vollrath, D. A., Sheppard, B. H., Hinsz, V. B., & Davis, J. H. (1989). Memory performance by decision-making groups and individuals*. Organizational Behavior and Human Decision Processes, 43*, 289-300.