Taylor, S. E., & Armor, D. A. (1996). Positive illusions and coping with adversity. *Journal of Personality, 64*, 873–898.

Van Woerkom, M., & Meyers, M. C. (2013). My strengths count! *Human Resource Management, 54*, 81–103.

Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science, 14*, 131–134.

Working With Social Comparisons in the Appraisal and Management of Performance

R. Blake Jelley
University of Prince Edward Island

Research and practice in performance appraisal and performance management seem to suffer from the same "delusion of absolute performance" that Rosenzweig (2007, p. 112) described with respect to commentators' evaluations of company performance in a competitive market economy. Commentators on business success factors have tended to speciously neglect or downplay the relative nature of performance (Rosenzweig, 2007). Downplaying the relative nature of performance is apparently the strategy endorsed by most performance appraisal scholars, too. Goffin, Jelley, Powell, and Johnston (2009) estimated that less than 4% of the published performance rating research has involved relative or social-comparative approaches, despite demonstrable advantages for relative over absolute rating formats (discussed below). Similarly, social comparison research and organizational scholarship have not traditionally been closely integrated (Buunk & Gibbons, 2007; Greenberg, Ashton-James, & Ashkanasy, 2007).

Goffin and Olson (2011) described the status of comparative appraisals as "pariah-like" (p. 51), a perception that is not likely to be improved by erroneous statements equating "relative rating" and controversial "forced ranking" as "synonymous terms" (e.g., Dominick, 2009, p. 413). Forced rankings involve relative comparisons (Jelley, Goffin, Powell, & Heneman, 2012), but relative approaches need not force any particular rating distribution or management decision. It is better to recognize that relative approaches to performance rating encourage raters to compare ratees with one another and are fundamentally different from conventional approaches wherein each ratee

R. Blake Jelley, School of Business, University of Prince Edward Island, Charlottetown, Prince Edward Island, Canada.

Correspondence concerning this article should be addressed to R. Blake Jelley, School of Business, University of Prince Edward Island, 550 University Avenue, Charlottetown, Prince Edward Island, Canada C1A 4P3. E-mail: bjelley@upei.ca

is considered individually and evaluated against some "absolute" standard that does not involve explicit social comparisons (Goffin, Gellatly, Paunonen, Jackson, & Meyer, 1996; Goffin et al., 2009; Goffin & Olson, 2011).

In the debate about abolishing or retaining formal performance ratings, both sides acknowledged relative comparisons to a limited extent (Adler et al., 2016). Arguing the position that formal (e.g., annual) performance ratings should be abolished, Colquitt, Murphy, and Ollander-Krane (CMO) recognized that "people compare themselves with others" and that social comparisons are central to influential theories in psychology and management (e.g., equity theory). CMO cautioned, however, that people's drive to engage in social comparison does not mean they welcome others (i.e., raters) to also compare them with others ratees.

Arguing that it would be a bad idea to get rid of ratings, Adler, Campion, and Grubb (ACG) noted the value of differentiated performance evaluations to support merit-based human resource decisions. ACG also cited research alleging that Millennials, in particular, are motivated by relative performance feedback and suggested that methods that differentiate performance may gain importance as Millennials become more prominent in the workforce (Adler et al.). ACG recognized research on relative rating methods (among other examples of the many rating approaches appraisal scholars have investigated), but neither side in this debate reviewed evidence that social comparisons can improve ratings. A brief summary of such evidence is provided below, followed by a broader discussion of the power of social comparisons and a call for scholars of performance appraisal and performance management to better recognize the potential benefits and challenges inherent in managing social comparisons to improve performance.

Prevalence of Social Comparisons

Social comparisons are natural, automatic, pervasive processes people use to assess their own and others' attributes, particularly (but not only) when objective indicators are unavailable (Buunk & Gibbons, 2007; Festinger, 1954; Goffin & Olson, 2011). Assessments of attributes such as one's attitudes, attractiveness, creativity, intelligence, strength, or performance are more meaningful when judged relative to other people. Goffin and Olson (2011) provided examples of the evolutionary functionality of social comparisons, including for self-protection. When deciding whether or not to fight an aggressor, the critical question is relative ("Is the aggressor stronger than me?"), not absolute ("How strong is the aggressor?" Goffin & Olson, 2011, p. 54).

Social comparison information may affect survival beyond the individual level, too. For example, Rosenzweig (2007) outlined specific tactics a large retailer used to improve its operations, which were associated with improved inventory turns from 3.45 to 4.56 between 1994 and 2002. The 32% increase seemed favorable by absolute standards. However, its competitors performed better and improved even faster. One rival had better inventory turns in 1994 (5.14) than the soon-to-be-bankrupt retailer did in 2002, yet that rival increased its inventory turns by 57% (to 8.08) in the same 8-year period. Absolute improvements must be considered in their social context.

The valuable information provided by social comparisons may also explain their dominance in self-assessments of task performance. In a series of three studies, Van Yperen and Leander (2014) demonstrated a phenomenon they called "the overpowering effect of social comparison information (TOESCI)" (p. 676). Van Yperen and Leander (2014, p. 676) observed a consistent use of interpersonal standards (i.e., "How did I do relative to others?"), even for participants who adopted mastery-based goals (self-selected or experimentally induced) and for whom temporal, intrapersonal feedback (i.e., "How did I do relative to how I did before?") should have been most relevant to self-evaluations of performance. Instead, self-evaluations were driven by social comparisons, with temporal comparison information playing a supplementary role (Van Yperen & Leander, 2014). People recognize that the performance of employees, athletes, students, and others is judged in large part relative to others' performance. Gold medals are awarded for being the best in a relative sense, not necessarily for intrapersonal bests, even if the route to winning requires concerted improvement of personal (or team, or organizational) performance over time.

Improving Rating Validity via Social Comparisons

From its origins focusing on a drive for self-evaluation via self–other comparisons, social comparison theory (Festinger, 1954) has given rise to a large, complex field of research (Buunk & Gibbons, 2007). Goffin and colleagues (2009; Goffin & Olson, 2011) drew on developments in social comparison theory to highlight its relevance for other–other comparisons and associated implications for enhancing rating quality. Consistent with recommendations to simplify rating formats in line with raters' natural cognitive processes (e.g., ACG in Adler et al.), Goffin and colleagues (2009) argued that raters may find relative ratings simpler than absolute ratings and, ultimately, provide higher-quality ratings when social comparisons are invoked explicitly. Greenberg et al. (2007) acknowledged the prevalence of social comparisons and argued that the workplace makes assessments of relative standing particularly salient, yet with respect to performance appraisal, they were concerned about biases and speculated that social comparisons may not always enhance appraisal validity (cf. Goffin & Olson, 2011).

Goffin and Olson (2011) summarized evidence for the superiority of relative ratings over absolute ratings in the domains of job performance, attitude measurement, and person perception. Goffin and colleagues' (1996) Relative Percentile Method (RPM) overcomes some serious problems that plague traditional relative approaches (e.g., it avoids the ordinal level of measurement inherent in straight rankings). The RPM also facilitates rater calibration by prescribing a relevant comparison group. In the performance appraisal domain, Goffin et al. (1996) found the RPM generally had stronger relations than did the Behavior Observation Scale (BOS; Latham & Wexley, 1977) with conceptually-related variables such as personality, vocational interest, and cognitive ability. Goffin et al. (2009) found that job performance ratings leveraging social comparisons showed incremental validity over noncomparative ratings in relation to managerial assessment center scores, but noncomparative ratings showed no incremental validity over social-comparative ratings. Goffin and colleagues (2009) also cited previous meta-analyses in which higher validity coefficients were observed for relative ratings than for absolute ratings (i.e., Heneman, 1986, .66 vs. .21; Nathan & Alexander, 1988, .51 vs. .34; cf. Bommer, Johnson, Rich, Podsakoff, & MacKenzie, 1995, .45 vs. .36). Also, Wagner and Goffin (1997) found that relative ratings were more accurate than were absolute ratings on some indices, whereas absolute ratings showed no advantages over relative ratings. Overall, explicit social comparisons generally enhance the accuracy and validity of performance ratings.

Implications for Performance Appraisal

The advantages for relative over absolute rating methods in terms of accuracy and validity have a seemingly obvious implication for collecting performance (or other) ratings for research purposes: Consider using relative ratings! Obtaining the highest-quality ratings is central to such investigations (DeNisi & Gonzalez, 2000). Implications for the practice of appraisal in organizations are less clear. The ultimate goal of an organization's performance appraisal system should be to improve performance (Adler et al.; DeNisi & Gonzalez, 2000). Stakeholder understanding and acceptance are critical for applied appraisal systems (DeNisi & Gonzalez, 2000). Although the RPM was consistently perceived most favorably among the relative scales examined by Roch, Sternburgh, and Caputo (2007), the relative scales were perceived to be less fair than were the absolute scales, especially the BOS. As noted previously, the fact that people routinely engage in social comparison does not necessarily mean they welcome being compared with others (CMO in Adler et al.).

Appraisals used to support administrative decisions and appraisals for employee development represent different paths to the same goal of performance improvement at one or more levels of analysis (DeNisi & Gonzalez, 2000). Even if people would prefer not to be evaluated at all or relative to one another, performance is always judged in some manner for these purposes (ACG in Adler et al.). Using more valid ratings to inform administrative decisions (e.g., compensation, work assignments, workforce reductions, promotions) helps managers better reinforce desired behavior, recognize achievements, signal what is valued, and place personnel in positions where they can contribute most effectively to the organization's performance. Reactions to social-comparative ratings might be more favorable if they were to be collected to support a particular between-person decision involving the allocation of scarce resources or opportunities (e.g., a promotion competition) rather than as part of an annual rating ritual. The inevitable nature of social comparisons in competitive contexts is likely to make relative ratings seem reasonable and justified. Also, explicit relative ratings should increase a rater's awareness of how high performers could be unfairly penalized if low performers were to be given inflated ratings.

With respect to the employee development path to performance improvement, Wagner and Goffin (1997) found that raters were better able to identify individual patterns of strengths and weaknesses when raters made relative (vs. absolute) and broad (vs. narrow) ratings. Unfortunately, these are exactly the opposite of the kinds of ratings seen as most valuable to help employees develop (e.g., Latham & Wexley, 1977; see also Kluger & DeNisi, 1996). So far, attempts to leverage social comparisons to enhance the quality of BOS ratings have not resulted in easy solutions to the paradox of what is best for ratees versus what raters can most accurately assess (Jelley & Goffin, 2001; Jelley et al., 2012). Fortunately, both sides in the current debate recognize that the performance management process necessarily involves more than performance rating (Adler et al.). For example, ACG noted that managers can learn to make more valid assessments and use such assessments to improve performance through frequent, constructive, actionable, and inspiring dialogues with employees. Moreover, the employee development path to performance improvement represents intriguing possibilities for strategic use of social comparisons beyond rating format considerations.

Strategic Consideration of Social Comparisons for Employee Development

People engage in social comparisons not only for self-evaluation and self-enhancement but also for self-improvement (Buunk & Gibbons, 2007). With respect to the "intervention utility" (p. 17) of social comparisons, Buunk and Gibbons noted the need for research on the motivational effects of "judiciously encouraged" or "controlled" comparisons (similar to the use of prescribed referents to improve ratings; Goffin & Olson, 2011). Likewise, health psychologists recognize that chronically-ill patients engage in social comparisons that can affect a patient's well-being positively or negatively. Consequently, health psychologists have considered developing and testing interventions designed to facilitate social comparisons strategically to improve patient adjustment, health, and well-being (Arigo, Suls, & Smyth, 2014). In the performance management domain, social comparison interventions could be targeted to managers to help them more effectively coach employees or to employees directly to help them manage their own social comparisons in a functional manner. The purpose of this section is not to advocate blindly for more social comparisons in performance management but to advocate for more strategic consideration of a dominant process that can have implications for employee performance and well-being. Just as our understanding of the roles for feedback and goals have become more nuanced (e.g., Kluger & DeNisi, 1996; Seijts & Latham, 2012), so should our understanding and management of social comparisons in the performance improvement process.

Comparisons with those who are better (i.e., upward comparisons) and with those who are worse (i.e., downward comparisons) can invoke both favorable and unfavorable affective responses (Buunk, Collins, Taylor, Van Yperen, & Dakof, 1990). Adapting insights from Buunk et al. to the performance management domain, upward comparisons may reveal that (a) an individual does not perform as well as everyone else and (b) it is possible for the individual to perform better than he or she presently does. Similarly, downward comparisons can show that (c) others are worse than an individual and (d) it is possible for the individual's performance to decline in the future. Distinct neural patterns have been found with respect to upward and downward social comparisons of self-performance on a personally-meaningful knowledge quiz (Lindner et al., 2015). Downward comparisons tended to activate brain regions associated with reward processing and self-related information, whereas findings for upward comparisons were "principally consistent with the notion that an upward comparison is an emotionally salient event that can motivate a change in goal-directed behavior" (Lindner et al., 2015, p. 572). Performance improvement involves the recognition or creation of discrepancies between an individual's current state and a desired future state to motivate change. Some discomfort is inherent, and desirable, in that process (ACG in Adler et al.). Judicious consideration of both upward and downward targets may help employees balance motivation for self-improvement with appropriate self-enhancement (Buunk et al., 1990).

Applying findings from the social comparison literature to individual employee development may include helping an employee frame social comparisons constructively, choose relevant upward comparisons as sources of inspiration and task strategies (i.e., role models), and invoke downward comparisons when the employee is in need of reassurance of his or her capability. We must recognize that "[s]ocial comparison pervades our interactions with others, informing us of our standing and motivating improvement, but producing negative emotional and behavioral consequences that can harm relationships and lead to poor health outcomes" (Swencionis & Fiske, 2014, p. 140). Any use of social comparisons during performance management discussions should attempt to build and maintain perceptions of support and trust, rather than adopt a negative tone (e.g., avoid "why can't you be more like your coworker?").

Both managers and employees need to recognize that social comparisons can have strong, "overpowering" effects that do not easily align with the intrapersonal, temporal-improvement goals that are seen as most beneficial for performance improvement, intrinsic motivation, and collaboration (Van Yperen & Leander, 2014, p. 676). This consideration complements ideas based on feedback intervention theory (Kluger & DeNisi, 1996) and the strategic use of learning and performance goals (Seijts & Latham, 2012). Consideration of TOESCI highlights that this dominant tendency may take explicit, repeated reminders to overcome (Van Yperen & Leander, 2014) or manage functionally. Instead of ignoring social comparisons, it would seem wise to recognize that social comparisons are pervasive and can be useful but then help people focus on their own behavior and performance improvements over time.

"How can people who are in need of self-enhancing feedback make use of better-off others to facilitate eventual change in their standing?" B. P. Buunk and colleagues posed that question in 1990 (p. 1248). It may be time to more seriously investigate that and related questions to help manage performance in organizations. This will necessarily be only one aspect of an improved approach to the appraisal and management of performance. The widespread dissatisfaction with the impact of performance management to date (e.g., Adler et al.) suggests it is time to make better use of a natural, dominant, and potentially functional tendency to better appraise and manage performance.

References

Adler, S., Campion, M., Colquitt, A., Grubb, A., Murphy, K., Ollander-Krane, R., & Pulakos, E. D. (2016). Getting rid of performance ratings: Genius or folly? A debate. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 9*(2), 219–252.

Arigo, D., Suls, J. M., & Smyth, J. M. (2014). Social comparisons and chronic illness: Research synthesis and clinical implications. *Health Psychology Review, 8*(2), 154–214.

Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & MacKenzie, S. B. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology, 48*, 587–605.

Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes, 102*, 3–21.

Buunk, B. P., Collins, R. L., Taylor, S. E., Van Yperen, N. W., & Dakof, G. A. (1990). The affective consequences of social comparison: Either direction has its ups and downs. *Journal of Personality and Social Psychology, 59*(6), 1238–1249.

DeNisi, A. S., & Gonzalez, J. A. (2000). Design performance appraisal systems to improve performance. In E. A. Locke (Ed.), *The Blackwell handbook of principles of organizational behavior* (pp. 60–72). Malden, MA: Blackwell Business.

Dominick, P. G. (2009). Forced rankings: Pros, cons, and practices. In J. W. Smither & M. London (Eds.), *Performance management: Putting research into action* (pp. 585–625). San Francisco, CA: Jossey-Bass.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140.

Goffin, R. D., Gellatly, I. R., Paunonen, S. V., Jackson, D. N., & Meyer, J. P. (1996). Criterion validation of two approaches to performance appraisal: The Behavioral Observation Scale and the Relative Percentile Method. *Journal of Business and Psychology, 11*, 37–47.

Goffin, R. D., Jelley, R. B., Powell, D. M., & Johnston, N. G. (2009). Facilitating social comparisons in performance appraisal: The Relative Percentile Method. *Human Resource Management, 48*, 251–268.

Goffin, R. D., & Olson, J. M. (2011). Is it all relative? Comparative judgments and the possible improvement of self-ratings and ratings of others. *Perspectives on Psychological Science, 6*(1), 48–60.

Greenberg, J., Ashton-James, C. E., & Ashkanasy, N. M. (2007). Social comparison processes in organizations. *Organizational Behavior and Human Decision Processes, 102*(1), 22–41.

Heneman, R. L. (1986). The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. *Personnel Psychology, 39*, 811–826.

Jelley, R. B., & Goffin, R. D. (2001). Can performance-feedback accuracy be improved? Effects of rater priming and rating-scale format on rating accuracy. *Journal of Applied Psychology, 86*, 134–144.

Jelley, R. B., Goffin, R. D., Powell, D. M., & Heneman, R. L. (2012). Incentives and alternative rating approaches: Roads to greater accuracy in job performance assessment? *Journal of Personnel Psychology, 11*(4), 159–168.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*, 254–284.

Latham, G. P., & Wexley, K. N. (1977). Behavioral observation scales for performance appraisal purposes. *Personnel Psychology, 30*, 255–268.

Lindner, M., Rudorf, S., Birg, R., Falk, A., Weber, B., & Fliessbach, K. (2015). Neural patterns underlying social comparisons of personal performance. *SCAN Social Cognitive and Affective Neuroscience, 10*, 569–576.

Nathan, B. R., & Alexander, R. A. (1988). A comparison of criteria for test validation: A meta-analytic investigation. *Personnel Psychology, 41*, 517–535.

Roch, S. G., Sternburgh, A. M., & Caputo, P. M. (2007). Absolute vs. relative performance rating formats: Implications for fairness and organizational justice. *International Journal of Selection and Assessment, 15*, 302–316.

Rosenzweig, P. (2007). *The halo effect . . . and the eight other business delusions that deceive managers*. New York, NY: Free Press.

Seijts, G. H., & Latham, G. P. (2012). Knowing when to set learning versus performance goals. *Organizational Dynamics, 41*(1), 1–6.

Swencionis, J. K., & Fiske, S. T. (2014). How social neuroscience can inform theories of social comparison. *Neuropsychologia, 56*, 140–146.

Van Yperen, N. W., & Leander, N. P. (2014). The overpowering effect of social comparison information: On the misalignment between mastery-based goals and self-evaluation criteria. *Personality and Social Psychology Bulletin, 40*(5), 676–688.

Wagner, S. H., & Goffin, R. D. (1997). Differences in accuracy of absolute and comparative performance appraisal methods. *Organizational Behavior and Human Decision Processes, 70*, 95–103.

Rating Performance May Be Difficult, but It Is Also Necessary

Steven T. Hunt
SAP SuccessFactors

The company I work for is one of the leading providers of performance management technology (Jones & Wang-Audia, 2013). This technology is used by more than 3,000 organizations worldwide, including several of the companies mentioned in Adler et al. (2016). The technology is highly configurable. It is currently being used to support performance management processes with no annual manager ratings, processes with traditional annual rating evaluations, processes that only evaluate competencies, processes that only evaluate goal accomplishment, processes that mix goals and competencies, processes that require forced-ranked comparisons between employees, processes that make no direct comparisons between employees, and much more. The capabilities of this and other human resources (HR) technology systems are allowing companies to radically rethink performance management because they enable companies to do things far differently from what was possible when they were constrained to more fixed electronic or paper forms (Hunt, 2011, 2015a). The result is an explosion in the diversity of approaches being taken toward performance management design.

My company naturally believes in the value of using performance management technology, but we do not have a strong opinion on what sort of performance management process companies should use. For example, it does not matter to us whether customers do or do not choose to collect annual manager ratings. What does matter is that whatever performance management processes they use add value to their organization, as this directly affects the value they get from using our performance management technology.

Steven T. Hunt, SAP SuccessFactors, Sherwood, Oregon.

Correspondence concerning this article should be address to Steven T. Hunt, 16958 Richen Park Circle, Sherwood, OR 97140. E-mail: s.hunt@sap.com