Precarious curve ahead: The effects of forced distribution rating systems on job performance

Sue H. Moon a,⁎, Steven E. Scullen b, Gary P. Latham c

a College of Management, Long Island University
b College of Business and Public Administration, Drake University
c Joseph L. Rotman School of Management, University of Toronto

A R T I C L E   I N F O

Article history:
Received 12 July 2012
Received in revised form 22 October 2015
Accepted 22 December 2015
Available online xxxx

A B S T R A C T

The forced distribution rating system (FDRS) is frequently used to appraise an employee's performance. The purpose of this paper is to synthesize theory and empirical research to present an integrative model for understanding the potential benefits and risks of a FDRS on the three components of job performance: task, citizenship, and counterproductive performance. A FDRS may lead to higher task performance in the relatively short-term, as it initially motivates effort as well as helps attract and retain top talent. Caution is in order, however, as a FDRS may also lead to lower citizenship performance and higher counterproductive performance through injustice perceptions and dysfunctional competition. Over time, the risks of a FDRS on job performance may increasingly outweigh initial benefits, particularly under certain task (interdependence) and group (cultural) characteristics. Implications of a FDRS for human resource management research and practice are discussed.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The forced distribution rating system (FDRS) is a performance appraisal system that forces supervisors to distribute a predetermined percentage of employees in categories based on their employees' performance relative to other employees' performance. Jack Welch is most often credited with popularizing the FDRS (Lawler, 2002; Naughton & Peyser, 2004). Under his leadership, General Electric’s (GE) FDRS involved a 20/70/10 “vitality curve” (Welch, 2001). Specifically, supervisors identified the “top 20%,” “vital 70%,” and “bottom 10%” of their subordinates. Those in the top 20% received raises two to three times greater than those given to the vital 70%, who were given “solid” increases. Those in the bottom 10%, in contrast, received nothing and were fired. Although GE’s 20/70/10 distribution is considered something of a standard (Becker, Huselid, & Beatty, 2009; Boehle, 2008), the FDRS is by no means unique to GE. For example, supervisors at both Ford and Goodyear have evaluated employees using a 10/80/10 distribution (Osborne & McCann, 2004). It is estimated that about 30% of America’s highest ranked (Fortune 500) companies have used variations of a FDRS (Ovide & Feintzeig, 2013). These companies include Amazon, Cisco, Dell, and others.

The defining characteristic of a FDRS are as follows: Supervisors are required to rate their employees’ performance relative to other employees’ performance; and, based on this appraisal, assign employees to predetermined performance categories. The number of performance categories is usually between three and five, and – comparable to a normal distribution or bell-shaped curve – there are typically smaller percentages of employees in the top and bottom performance categories, with most employees being rated in the middle categories (e.g., General Electric’s 20/70/10, Ford’s 10/80/10 distributions). These performance...
categories are, in turn, typically linked to decisions affecting compensation, promotion, and termination. Hence, the FDRS is often referred to in the business community as “forced ranking,” “stack ranking,” “the bell curve,” or “rank and yank” (Eichenwald, 2012; Holland, 2006; Kantor & Streitfeld, 2015). Despite significant controversy surrounding the practice, this FDRS continues to be used in many large organizations.

Based on computational modeling, Scullen, Bergey, and Aiman-Smith (2005) showed that a FDRS initially leads to improvement in a firm’s workforce potential. In a laboratory experiment, Berger, Harbring, and Sliwka (2013) showed that productivity on a task requiring independent work effort was higher in a forced distribution as compared to a baseline condition. While these studies contribute important insights, under-specified in the current literature is how social and psychological factors in the context of organizations influence the performance effects of a FDRS. As we further discuss in this paper, a FDRS may lead to dysfunctional competition and perceptions of injustice, and in turn less helping and more sabotage behaviors. These are critical factors to consider because job performance in real organizations often requires teamwork rather than independent work effort.

The purpose of this paper is to advance theory on the effects of a FDRS by reasoning about its potential benefits and risks on job performance, and the conditions under which its risks are expected to outweigh potential benefits. Toward this end, the paper is organized into three main sections. First, background information and existing literature is briefly reviewed. Second, an integrative model of the effects of a FDRS on the three components of job performance is proposed, with seven sets of research propositions flowing from this model. Through this integrative model, we explain how and why a FDRS may initially have positive effects on task performance, but over time have negative effects on citizenship and counterproductive performance, such that the practice may eventually become detrimental to employees’ overall job performance and by extension organizational performance – particularly when applied in organizational contexts that require high levels of task interdependence and social support among comparison others. As we further discuss below, in reasoning about the underlying processes through which a FDRS affects job performance, we assume that the organization is using the practice as a “stand alone” system (i.e., not in conjunction with other performance appraisal practices) based on valid performance criteria. Finally, based on the integrative model, we conclude with implications for human resource management research and practice.

2. Background information on forced distribution rating systems

2.1. Performance appraisal criteria

In general, Grote (2005) advocates use of the company's core values or, where possible, the list of job competencies that are important to a firm. The criteria against which performance was assessed at General Electric (GE) were the company's four E's of leadership: high energy, ability to energize others, edge to make tough decisions, and ability to execute and deliver on promises (Welch, 2001). Sanofi-Aventis used data on sales numbers, client relationships, product and service knowledge, and call reporting and documentation (Boehle, 2008). At Microsoft, employee performance was rated against previously identified job goals (Amalfie & Steinier, 2005).

2.2. Research on forced distribution rating systems

There is a paucity of research on the FDRS. Blume, Baldwin, and Rubin (2009) studied how potential applicants' attraction to an organization is influenced by four FDRS elements: (1) extent of rewards for high performance; (2) extent of consequences for low performance; (3) frequency of performance feedback; and, (4) size of comparison group. College students read descriptions of jobs in different companies that use different types of FDRSs. They then rated the attractiveness of those jobs. What the students found most attractive in a FDRS were: high monetary rewards for high performance; frequent performance feedback; and, large comparison group sizes. The severity of the consequences for low performance – particularly, job dismissal – had a negative impact on their judgment of a FDRS.

In a subsequent study, Blume, Rubin, and Baldwin (2013) found that respondents who scored higher on a cognitive ability test indicated greater attraction to an organization using a FDRS than those who had lower scores. The researchers reasoned that this may be due to those with higher cognitive ability anticipating success in acquiring the job knowledge needed to perform well. There was a significant interaction between collectivistic values and perceptions of the fairness of a FDRS. Respondents who scored high on collectivism were particularly sensitive to perceived lack of fairness in a FDRS. Collectivists value group welfare over individual rewards, and therefore are averse to a system that they perceive unfairly rewards some team members and punishes others.

While Blume et al. (2009, 2013) examined the FDRS from the perspective of potential ratees, Schleicher, Bull, and Green (2009) examined this practice from the perspective of potential raters. MBA students and managers reported that appraising performance was more difficult and less fair using a FDRS than a performance rating system where individuals are rated against an absolute standard (e.g., meets or exceeds sales goal). This was especially true if the raters perceived relatively little performance variability among employees, and if their evaluations were linked to administrative decisions such as salary raises, promotions, or terminations.

Berger et al. (2013) compared the performance effects of a forced distribution compared to an absolute rating condition. Productivity was 6 to 12% higher under the forced distribution condition. The performance benefits of the FDRS were lower, however, when workers had previously been exposed to the absolute rating condition. In addition, workers in the forced distribution condition were twice as likely as those in the absolute rating condition to sabotage others when the option to do so was available.

At the organizational level of analysis, Scullen et al. (2005) used computational modeling to examine the impact of the FDRS on an organization's workforce potential. Over the first few iterations of a FDRS, an organization can significantly improve the performance potential of its workforce. This initial improvement, however, falls to nearly zero over the next few iterations. The reason for this decreasing rate of improvement is that as low potential employees are removed (and are hopefully replaced with higher-potential applicants), the lower end of the distribution of employee performance potential becomes increasingly higher. As this continues over subsequent iterations, it becomes increasingly difficult for an organization to attract and hire applicants whose potential is higher than that of the lowest-ranked employees who had been recently terminated.

3. Integrative model: The effects of forced distribution rating systems on job performance

3.1. Clarification of independent and dependent variables

Given that the FDRS is a performance management tool, its effects on job performance warrants thoughtful consideration. As noted earlier, job performance has three components: task, citizenship, and counterproductive performance (Rotundo & Sackett, 2002). Task performance refers to behaviors that contribute to the production of a good, or the provision of a service, and behaviors formally recognized as part of the job (Borman & Motowidlo, 1993). Citizenship performance refers to behaviors that contribute to the organization’s psychological and social well being yet are not necessarily formally recognized as part of an employee’s job (Organ, 1988). Counterproductive performance refers to volitional behaviors that threaten an organization’s well being (e.g., blaming others, stealing supplies) (Robinson & Bennett, 1995).

3.2. Overview of integrative model

A central thesis of our integrative model is that a FDRS may benefit what has traditionally been recognized as job performance (task performance), but does so at the expense of citizenship performance. Furthermore, it may lead to counterproductive performance. That is, in the short-run, a FDRS may increase employee effort as well as help attract and retain high ability employees, and therefore increase task performance. Yet, the FDRS may in the longer run decrease citizenship and increase counterproductive performance, as it may engender injustice perceptions and dysfunctional competition, especially under certain conditions. As a FDRS has negative effects on citizenship and counterproductive performance over time, and thus the psychological, social, and organizational context that serves as a catalyst for task activities and processes (Borman & Motowidlo, 1997), negative effects on task performance and overall job performance may eventually follow – particularly when a FDRS is applied to jobs or in contexts that emphasize teamwork or social interdependence. These propositions are summarized in Fig. 1.

3.3. Domain, boundary conditions, and assumptions

We conceptualize our integrative model mainly at the individual level and hence draw primarily on theory at this level of analysis. While we do not explicitly focus on group dynamics, we do consider the moderating role of task characteristics and cultural constructs on the effects of a FDRS on an individual's performance. Furthermore, as a FDRS is reasoned to affect an

![Fig. 1. The effects of the FDRS on job performance.](Image)
individual's task, citizenship, and counterproductive performance, it follows that a FDRS affects employees' overall job performance, and by extension organizational performance. Additionally, three assumptions underlie our propositions:

i. The FDRS is being used as a “stand alone” performance evaluation system, rather than in combination with other methods (e.g., absolute rating system, group-based appraisals). While we do not attempt to describe and analyze the wide variety of ways organizations may want to cluster elements of their performance evaluation system and their effects, we do discuss in the “Human resource management implications” section how organizational decision-makers may want to use a FDRS in combination with other human resource management practices to mitigate its risks.

ii. The organization has invested the resources necessary for a FDRS to be technically sound, including training raters (e.g., supervisors) and using valid performance criteria. As we discuss within our integrative model, however, the extent of objectivity in performance criteria and therefore the risk of bias and injustice perceptions will depend in part on the characteristics of the job (e.g., task identity). Similarly, we assume that the organization provides comparison others with equal access to the resources (e.g., training, equipment) necessary to perform their jobs.

iii. As is often the case with ratings from performance evaluations (Ferris, Munyon, Basik, & Buckley, 2008; Landy & Farr, 1980), we assume that ratings from the FDRS are linked to administrative decisions, particularly higher pay associated with a top performance ranking and the potential for termination associated with a bottom performance ranking (e.g., Kwoh, 2012, Welch, 2001, 2005).

No integrative model can account for all possible relations. In the interest of parsimony, we illustrate key variables based on a review of the literature. In addition, as we discuss below, there may be offsetting relations among the criterion variables. For example, Halbesleben and Bowler (2007) suggested that individuals engage in higher levels of citizenship performance when they fear their task performance falls below managerial expectations. The goal of our model is to provide insights that may stimulate future theorizing and empirical research, including assessment of the model’s boundary conditions and assumptions.

3.4. Potential benefits

3.4.1. Minimizes rater leniency and motivates employee effort

While Welch has been credited with popularizing the FDRS (e.g., Kwoh, 2012), the practice had its beginning in the years between World Wars I and II (Cappelli, 2009). After World War I, officers in the U.S. Army were rated against an absolute standard, namely the extent to which they exhibited certain traits or behaviors (e.g., “leadership,” “ability to obtain results”). Ratings were on a Likert type scale, typically ranging from 1 (unsatisfactory) to 5 (satisfactory) (Sisson, 1948). While these ratings appeared to be valid for identifying exceptionally poor performers, ratings were typically inflated.

Just prior to World War II, the U.S. Army needed to quickly promote top performers to serve as generals of the rapidly mobilizing forces (Sisson, 1948). Because most of the officers had been rated as excellent or superior, the absolute rating system was inadequate for identifying top performers. The Army therefore devised a “relative rating system,” whereby officers were rated relative to one another.

The initial logic for using a FDRS – that is, to curtail rater leniency – remains the same today. As Bretz, Milkovich, and Read (1992) noted, organizations using absolute rating systems with five levels to differentiate employees typically find that only three categories are used very often in practice. Often sixty to 70% of employees are rated in the top two categories, and the bottom category is often empty. In short, a disproportionate number of employees are rated as above average when an absolute rating system is used.

Welch (2005) argued that a FDRS promotes candor on where an employee stands relative to others. Research also suggests that some employees are motivated to work harder when they initially receive negative feedback (Kluger & DeNisi, 1996). In addition, employees are likely to act on feedback that is linked to human resource management decisions such as salary increases or job terminations (Boswell & Boudreau, 2002). Reliable and valid data that differentiate high, average, and poor performers may motivate greater effort toward task performance than does an absolute rating, as managers reward top performers with raises and promotions, while awarding less pay to average performers and potentially dismissing those at the bottom (Rynes, Gerhart, & Parks, 2005). Likewise, performance feedback from a FDRS may motivate effort by conveying symbolic meaning to employees with respect to their competence and status in the organization (Jenkins, Mitra, Gupta, & Shaw, 1998).

The consequences for employees in the bottom category may also signal to co-workers that management does not tolerate substandard performance. Laboratory experiments show that observing a co-worker being punished for low output increases production more than seeing a low-performing co-worker go unpunished (Schnake & Durnler, 1989). Berger et al. (2013) found that productivity was 6 to 12% higher in a forced distribution compared to absolute rating condition – even as total payout was higher in the latter condition, as there were more employees being rated in the top two of five performance categories (approximately 80% vs. 30%) as supervisors were not forced to differentiate among them, and therefore more workers receiving the rewards associated with the top two ratings (approximately $13 & $10 US). As Berger et al. (2013) stated: “The reason for this substantial gain in performance is that many supervisors in the baseline [absolute rating] setting are very lenient in their rating decisions, and hence, performance incentives are weak” (p. 65). Taken together, we predict that:

Proposition 1. A FDRS, compared to an absolute rating system, motivates greater effort toward task performance than does an absolute rating system.

3.4.2. Helps attract and retain high performers

A FDRS may have a “sorting effect” through differential attraction and retention processes (Rynes et al., 2005). Research suggests that employees who are relatively high in cognitive ability (Blume et al., 2013; Trank, Rynes, & Bretz, 2002), self-efficacy (Cable & Judge, 1994), and need for achievement (Turban & Keon, 1993) are more likely to be attracted to organizations that reward employees based on job performance. This is because high potential applicants appear to be attracted to organizations where their talents are likely to be recognized and flourish. By attracting high ability employees, this can motivate existing employees to continually challenge themselves to improve their performance to remain competitive with new recruits.

Conversely, in organizations that do not reward employees based on their performance, high performers may compare their inputs (e.g., effort, talent) and outcomes (e.g., pay, status) relative to low performers. Equity theory (Adams, 1965) suggests that they are likely to then reduce their effort or leave the organization (Carrell & Dittrich, 1978). A study of 984 engineers in high technology companies supports the contention that differentiation helps firms to retain “star” employees, while eliminating low performers. Firms that rewarded high performance had very high performers who wanted to remain, and very low performers who wanted to leave (Zenger, 1992).

As noted previously, it was found that individuals with higher cognitive ability are more attracted to a forced distribution than an absolute rating system (Blume et al., 2013). A FDRS may ensure high ability employees that a predetermined percentage of employees will be recognized and rewarded, thereby protecting them from the possibility of “stingy” management who are reluctant to administer rewards commensurate with certain individuals’ disproportionately large contributions to firm performance (Berger et al., 2013). Taken together, compared to an absolute rating system, we reason that a FDRS will initially lead to greater task performance through the attraction and retention of higher ability employees:

**Proposition 2a.** A FDRS, compared to an absolute rating system, attracts higher ability applicants and thereby motivates existing employees to improve their task performance to remain competitive with new recruits.

**Proposition 2b.** A FDRS, compared to an absolute rating system, leads to greater task performance through the retention of high performers, as these employees are more likely to view the organizations as a place where they will be recognized and rewarded for their high performance.

3.4.3. Summary of potential benefits

A FDRS may initially lead to greater task performance through incentive and sorting effects. The potential for rewards and the threat of punishment may motivate employees to maximize their potential (Rynes et al., 2005). While those who are ranked in the top performance category may be rewarded generously (e.g., via praise, pay, & promotion), those who deliver relatively low performance may face punishments such as job termination. This differentiation, in turn, may attract high potential applicants who view the organization as a place where their contributions will be acknowledged. However, these potential benefits may eventually be outweighed by risks.

3.5. Risks

3.5.1. Perceptions of organizational injustice

Rating theory (Wherry & Bartlett, 1982) explains why categorizations of employees in a FDRS may be inaccurate. Performance rankings may be largely beyond an employee’s control, as they are partially dependent on the behaviors of co-workers. In addition, a ranking reflects a supervisor’s recall of the employee’s behavior relative to other members of the team. Consequently, there is the danger of supervisors misidentifying more “competent” employees as less competent, particularly given that in any organization some teams are invariably better staffed than others (Latham & Wexley, 1994; Lawler, 2002). Specifically, competent employees may be judged as falling in the bottom category if they happen to be on a team of exceptional performers, whereas they might have been ranked in the top category if they had been on a team of relatively poor performers.

With a FDRS, a predetermined percentage of even an all-star team must be placed in the bottom performance category and thus face the possibility of job termination (Pfeffer & Sutton, 2006). This is especially problematic when each individual on a team plays a unique and integral role (Greenwald, McCalope, Marchant, & Terdiman, 2001). As a result, a FDRS is likely to be perceived as unfair, compared to an absolute rating system, because the former can lead to more inconsistent and inaccurate ratings, particularly as the composition of one’s comparison group changes (Roch, Sternburgh, & Caputo, 2007; Wherry & Bartlett, 1982).

Theory states that perceptions of organizational (in)justice are at least three-fold: distributive, procedural, and interpersonal (Colquitt, Conlon, Wesson, Porter, & Ng, 2001). Distributive justice refers to the perceived fairness of outcomes (e.g., an employee’s performance evaluation), while procedural justice refers to employees’ perceptions of the fairness of the procedures used to arrive at the outcome. Interpersonal justice refers to the degree to which people perceive that they are treated with respect.

Employee perceptions of organizational injustice have been linked to low citizenship and highly counterproductive performance (Cohen-Charash & Spector, 2001). Skarlicki and Latham (1996, 1997) showed that training leaders on organizational justice principles increased levels of citizenship performance among employees. Conversely, Skarlicki and Folger (1997) showed that perceptions of distributive, procedural, and interpersonal injustice interact to predict retaliation in the workplace. As a FDRS may increase perceptions of organizational injustice, their study suggests that employees may attempt to “get even” (e.g., through theft, vandalism, or sabotage). Relatedly, based on the results of more than 1000 daily diary recordings from
100 employees in various industries, Ferris, Spence, Heller, and Brown (2012) found that perceptions of interpersonal injustice are related to less citizenship performance (e.g., "refused to help a co-worker," "withheld work-related information from a co-worker") and more counterproductive performance (e.g., "purposely came to work or came back from lunch breaks late," "was nasty to a fellow-co-worker").

Colquitt (2004) reported that employees care not only about being treated fairly themselves, but also that their co-workers are treated fairly. Given the likelihood for causing perceptions of injustice, many managers find a FDRS more difficult to implement than an absolute rating system, and therefore often resist it (Schleicher et al., 2009). Even Welch (2001) acknowledged that, "Managers will play every game in the book to avoid naming their bottom 10. Sometimes they'll sneak in people who were planning to retire that year, or include others who have already been told to leave the organization" (p. 161). This is an example of both low citizenship and high counterproductive performance. Indeed, a survey of 227 human resource management professionals suggests that some managers retain incompetent employees so they can rank them in the bottom category during the next performance appraisal (Novations Group, 2004). In doing so, managers are likely to discredit the FDRS and further contribute to employees' negative perceptions of its fairness (Lawler, 2002).

Relatedly, employees may perceive the misuse of supervisors' reward and coercive power as a result of the FDRS (or supervisors' enhanced ability to influence employees' behavior through the administration of desirable [e.g., raises, bonuses] and undesirable outcomes [e.g., probation, termination]). Moreover, this may come at the expense of more effective, personal bases of power, namely, expert and referent power (French & Raven, 1959; Podsakoff & Schriesheim, 1985; Yukl & Falbe, 1991). Perceptions of procedural and interactional justice may suffer, potentially decreasing employee helping behavior and increasing sabotage behavior.

**Proposition 3.** A FDRS, compared to absolute rating system, leads to less citizenship performance and more counterproductive performance through greater perceptions of injustice.

### 3.5.2. Dysfunctional competition

A FDRS fosters a dysfunctionally competitive (vs. co-operative) situation as performance ratings are dependent on an employee's relative standing among employees, and therefore rewards for some employees prevent rewards for others (Beersma et al., 2003). Competitive situations may cause an employee to focus on the attainment of personal goals over group goals (Deutsch, 1949), and therefore reduce the propensity for "group helping" or "backing up behavior," particularly when it is apparent the group members are failing to reach the group's goals (Porter, 2005). As Deutsch (1949) theorized, not only is group-helping behavior less likely, it is also more likely that there will be hindering behavior. For example, as group members' behaviors are perceived in terms of their effects on one's relative standing, there is more likely to be distortion and conflict in the communication process.

Studies have shown that people are less willing to cooperate and more likely to dysfunctionally compete with others when they are close to attaining, or falling out of, a given performance category (Garcia & Tor, 2007; Garcia, Tor, & Gonzalez, 2006). In these studies, participants were asked to imagine being the CEO of a company ranked near the top (#3), well within (#104), or at the bottom (#500) of the prestigious Fortune 500 list of America’s largest corporations. They were asked to decide whether they would enter into a joint venture, with a company ranked one position below them (i.e., at #4, #105, or just off the list at #501) in which both companies’ profits would increase by 6%. Without the joint venture, their company’s profit would increase by 5%, and their rival’s profit would increase by 1%. Participants were significantly less willing to enter a joint venture and maximize joint gains if their company was ranked near the top (#3) or at the bottom (#500) of the Fortune 500 as compared to well within it (#104).

When employees perceive that they are close to a border between performance categories (e.g., just barely in the top performance category, or in danger of slipping into the bottom category category), they are less willing to help their co-workers because of the perceived benefits and consequences associated with being in one category versus another. Competing rather than co-operating with coworkers in such situations is arguably rational from an employee's perspective (e.g., to avoid being ranked in a lower category), especially as an individual's task performance is typically weighted more heavily than citizenship performance in performance appraisals (Bergeron, 2007). Thus, employees may become selective about the colleagues whom they choose to help. An employee near the top performance category (e.g., near the 20/70 dividing line in the case of General Electric’s 20/70/10 forced performance appraisals (Bergeron, 2007), or at the bottom (#500) of the Fortune 500 as compared to well within it (#104).

A FDRS is in many ways a “promotion tournament,” which economic theory suggests elicits undesirable behaviors (Lazear & Rosen, 1981). Promotion tournaments are common in occupations where monitoring how individuals complete work tasks is difficult (e.g., academia, law). There are three key characteristics of a promotion tournament: (1) who will win is uncertain; (2) winning is based on relative performance; and, (3) there are large differences between the payoffs for winners and losers (Ehrenberg, Smith, & Chaykowski, 2004). Becker and Huselid (1992) found that as the financial difference between winning and losing increases, so do race drivers’ performance. And when the difference is especially large, drivers are more likely to engage in hazardous behavior that jeopardizes the safety of others. Pfeffer and Langton (1993) found that the greater the pay dispersion within academic departments, the less likely it is that faculty members will collaborate on research. Pay dispersion often promote dysfunctional competition by inciting jealousy and resentment, which in turn damage social relationships in the
workplace. As noted earlier, Berger et al. (2013) found that workers were twice as likely under a FDRS, compared to an absolute rating system, to sabotage a co-worker’s performance when the opportunity was available.

A study of golf players demonstrated that, contrary to common sense, tournaments can have the adverse incentive effect of reducing effort when there are large differences in abilities among individuals. Brown (2011) found that when competitors are similar in ability, they are motivated to expend extra effort to win a tournament. When these same players are competing against others with superior abilities (e.g., Tiger Woods during his “hot” periods), they are not as motivated to expend effort because they perceive their probability of winning as low. This study suggests an additional risk of the FDRS, namely, withholding effort when competing against star performers. Many employees are likely to conclude that they are unable to be ranked in the top performance category. However, if the top performance category is relatively large, this may lead “stars” with superior abilities to feel unappreciated. Consequently, while the size of the specific performance categories may be arbitrary (e.g., top 2% vs. 20%), the size of these categories may lead to real differences in employees’ behaviors.

Proposition 4a. A FDRS, compared to absolute rating systems, may lead to lower citizenship performance and greater counterproductive performance due to dysfunctional competition.

Proposition 4b. Dysfunctional competition is stronger among employees who perceive that their current standing places them relatively close to a meaningful cut-off point between FDRS categories.

Proposition 4c. Competent employees who perceive they are being compared to “superstars,” and believe they are unlikely to be ranked in a top category under a FDRS will withhold full effort.

3.5.3. Summary of risks

A FDRS may lead to perceived injustice and dysfunctional competition. Perceptions of injustice are in turn linked to lower citizenship performance and higher counterproductive performance (Cohen-Charash & Spector, 2001; Colquitt et al., 2001). Additionally, under a FDRS, there are negative correlations among group members’ rewards. Consequently, as dysfunctional competition ensues, employees may not help, and may even hinder, the performance of others.

3.6. Relation among task, citizenship, and counterproductive performance

Borman and Motowidlo (1997) reasoned that citizenship and counterproductive behaviors are important because they “shape the organizational, social, and psychological context that serves as the catalyst for task activities and processes” (p. 100). As citizenship behavior decreases and counterproductive behavior increases, the negative effects of a FDRS on the organizational, social, and psychological context are likely to extend to employees’ effectiveness in performing activities that contribute to the organization’s technical core: task performance. Although these relations have not been examined in prior research with respect to the FDRS and therefore their precise nature and direction are unclear, it is reasoned that task performance may increase in the short-run, but then decrease along with the practice’s detrimental effects on citizenship and counterproductive performance in the longer-run.1 Taken together, we theorize that the effects of a FDRS on task performance and overall job performance may be curvilinear over time. The offsetting relations among task, citizenship, and counterproductive performance under a FDRS over time are illustrated in Fig. 2.

As it is the combined effects of task, citizenship, and counterproductive performance and, in turn, overall job performance that should be salient to the organization, we specify the conditions (moderating variables) under which the potential benefits of a FDRS are likely to be outweighed by its costs. For example, in addition to time, we theorize that a FDRS is less likely to be effective in jobs characterized by high levels of task interdependence and contexts that emphasize teamwork and collectivism.

3.7. Moderating variables

Time and organizational contexts are reasoned to be boundary conditions that influence the effect of a FDRS on task, citizenship, and counterproductive performance.

3.7.1. Time

In most organizations that use a FDRS, the ranking process occurs annually (Grote, 2005). Welch (2001) claimed that, “Year after year, differentiation raises the bar higher and higher and increases the overall calibre of the organization” (p. 158). As noted earlier, Scullen et al. (2005), using computational modeling, found that an organization’s gains from a FDRS eventually diminish to nearly zero because initially it is relatively easy to identify and terminate poor performers. But it becomes increasingly difficult with each iteration of the FDRS to hire employees who are superior to those who were terminated.

An individual’s performance is dynamic. The typical organizational tenure-job performance relationship is characterized by an inverted U-shaped curve. Performance usually increases during an employee’s early years in an organization, but eventually decreases or plateaus (Sturman, 2003). Performance may also vary over time due to factors beyond an employee’s control, such as changing technological, social, and economic conditions. Furthermore, an employee whose level of performance is

1 We thank an anonymous reviewer for raising the possibility of offsetting relations among the model’s criterion variables over time.
consistent across time may be ranked higher at one point in time and lower at a subsequent point only because supervisory performance ratings are subjective (Sturman, Cheramie, & Cashen, 2005). Variability is especially likely when different managers are ranking the same employee’s performance (Viswesvaran, Ones, & Schmidt, 1996). In addition, as high performing employees replace relatively low performing employees, the relative standing of mid-range employees will almost certainly decrease. Employees who experience declines in their standing over time will also experience a decline in rewards. This may have negative effects on their morale, and result in the withholding of effort (Clark, Masclet, & Villeval, 2010).

Workers’ reaction to pay is often a function of where they stand in their organization’s pay hierarchy, with the most negative reactions occurring among workers whose pay is relatively low in an organization where pay dispersion is relatively wide (Clark et al., 2010; Trevor & Wazeter, 2006). Managers may not be concerned in the short-run about the negative reactions of low performers, as they may want these individuals to exit the organization. As more of the low and average performers are eliminated with successive iterations of a FDRS, however, the performance variability of the remaining employees may narrow. Consequently, a small decrease in an employee’s performance (relative to previous performance) may result in a larger decrease in rankings. Thus it may become increasingly difficult for an employee to be ranked within a higher performance category – or maintain their current one – contributing to rankings becoming less stable across time. As even Welch (2001) acknowledged, by the third iteration of a FDRS, “it’s war” (p. 161). Furthermore, a FDRS and its high stakes may contribute to a short-term focus as employees strive to acquire the relatively aggressive rewards, or avoid the relatively aggressive consequences associated with the typically annual rankings.

Consistent with the reasoning above, Berger et al. (2013) found that switching from an absolute rating to a forced distribution condition initially increased effort, but the net effect on performance became negative over time as employees realized, over iterations of the practice, that it becomes harder to achieve a ranking in the top performance category. This finding indicates that “history matters” (Berger et al., p. 66), as a FDRS is less likely to be effective if employees have been previously exposed to an absolute rating system.

Proposition 5a. The short term benefits of a FDRS on task performance may be increasingly outweighed by the costs of lower citizenship and higher counterproductive performance over time.

Proposition 5b. A FDRS and the high-stake rewards and punishments associated with this practice may contribute to an employee’s short-term (e.g., within rating period) rather than long-term focus.

Proposition 5c. A FDRS is less likely to be effective if employees have had previous experience with an absolute rating system.

3.7.2. Organizational context and job characteristics

A survey involving 278 organizations revealed that only a third (32.9%) found a FDRS to be at least “moderately effective” (Bernthal, Rogers, & Smith, 2003). We propose that the effects of a FDRS on job performance will be affected by the organizational context, particularly the organization’s structure and culture, and the resulting way that organizational tasks and decision-making authority are formally and informally divided and coordinated among its members. For example, characteristics of a bureaucratic organizational structure – such as centralized decision-making, standardized job tasks through formalized policies and procedures, and close supervision through narrow spans of control – result in reduced personal discretion and performance variability among most employees. Hence, the differentiation that is inherent in a FDRS will be harder to justify, and therefore more likely to lead to perceptions of injustice. Relatedly, when individuals are expected to handle the same input in the same way as they follow standardized systems that co-ordinate collective efforts to yield consistent and uniform output (e.g., assembly line work, fast food preparation), deviations from standard performance may be undesirable, and “above standard” performance by an individual is less likely to have a large, positive impact on overall firm performance.

On the other hand, a FDRS may be more appropriate in the context of a decentralized organizational structure and thus wider spans of control, because there is less risk of injustice perceptions when the size of comparison groups are large and there is significant performance variance among employees. In addition, a decentralized organizational structure is associated with higher

levels of job autonomy. Hackman and Oldham (1975) defined autonomy as the amount of freedom and independence job incumbents have in completing organizational tasks. For example, scheduling one’s work and determining the methods to perform tasks (Morgeson & Humphrey, 2006). As jobs high in autonomy involve, by definition, personal judgment and discretionary effort, there is more likely to be performance variability among job incumbents. For such jobs, there may be greater competition for top talent – as top talent, coupled with intense and persistent discretionary effort, may have a significant impact on firm performance (e.g., significantly increasing revenue or decreasing costs through new product or process innovations). Under such conditions, a FDRS and the associated differentiation of ratings and rewards may help attract, motivate, and retain top talent.

While elements of organizational structure can guide and align employees' behaviors, a strong organizational culture may achieve similar ends through different means (O'Reilly & Chatman, 1996). Organizational culture refers to a system of shared values and norms that help distinguish one organization from another, and help define what are appropriate attitudes and behaviors for its members (Chatman & Jehn, 1994; Hofstede, Neuijen, Ohayv, & Sanders, 1990). Without necessarily specifying exactly when and how organizational tasks are to be completed through formalized systems or procedures, a strong organizational culture may nonetheless function as a social control mechanism, as widely-shared and deeply entrenched values and norms guide and align employees’ attitudes and in turn behaviors.

We suggest that a FDRS is inconsistent with organizational cultures that place strong emphasis on people and teamwork, and involve high levels of social support and task interdependence among comparison others for peak performance. For example, Siegel and Hambrick (2005) noted that members of top management teams in high technology settings must constantly engage in scanning behaviors, convey what they discover to team members, and participate in discussions and debates about the interpretation and implications of information. Hence, team members need to intensively coordinate and collaborate on activities to be effective. But, a FDRS and the competitive situation it fosters are likely to undermine the rapid and candid communication of information, while increasing information distortion and dysfunctional competition and conflict (Deutsch, 1949). In addition, those who receive far less rewards may respond with jealousy, while those who receive far more rewards may respond with condescension, aloofness, and social distancing towards seemingly less worthy group members (Siegel & Hambrick, 2005). Furthermore, a FDRS could reduce team members’ tendency to engage in frequent negotiations and mutual adjustments – behaviors needed for innovation (Siegel & Hambrick, 2005). As discussed earlier, a FDRS and the shift to a competitive situation are likely to impair, rather the foster, collaboration and innovation related behavior. Indeed, Eichenwald (2012) reported that Microsoft’s FDRS undermined teamwork and in turn firm performance as “a lot of Microsoft employees did everything they could to avoid working alongside other top-notch developers” (“The Bell Curve,” para. 7) and “not only tried to do a good job but also worked hard to make sure their colleagues did not” (“The Bell Curve,” para. 8). A FDRS may also undermine an organization’s ability to attract and retain team players, and top talent who seek an organization that supports collaboration and close relationships with coworkers. In short, a FDRS is less likely to be effective in jobs characterized by complementarity or “connectedness” of work tasks among comparison others – either as work flows from one’s job to other jobs or is affected by work received from other jobs (Kiggundu, 1981).

On the other hand, when there is minimal need for assistance from or collaboration with others and performance is dependent on individual work effort, a FDRS may motivate employees to improve their separate buffered domains. Under such conditions, there may be fewer opportunities for social loafing, and less need for social monitoring. Under these conditions, it may be easier to parcel out or identify an individual’s output, making evaluations under a FDRS less subject to political influence (e.g., personal liking, ingratiation) (Kepes, Delery, & Gupta, 2009). In addition, perceptions of injustice are less likely to emerge for jobs in which reliable, valid, and agreed-upon measures of output are available. Jobs high in task identity – the degree to which a job involves completing a whole piece of work, the result of which can be easily identified (Hackman & Oldham, 1975; Morgeson & Humphrey, 2006) – are more likely to be amenable to objective, valid performance criteria. That is, there is less likely to be ambiguity in the contribution of an individual employee when the job consists of creating an entire product or completing an entire unit of service. As a result, there will be greater fairness perceptions associated with a FDRS when jobs are high in task identity, and there are common, valid, and objective performance criteria available for the work performed by comparison others.

Finally, a FDRS is more likely to be sustainable to the extent that there are opportunities to move up (e.g., assistant to associate; analyst to supervisor or manager) or to another comparison group (or organization) as job incumbents show high (or low) performance over successive iterations of a FDRS. This helps ensure that the remaining group members are comparable in terms of ability and experience. The prospect of promotion, or movement up the organizational or occupational hierarchy, may also provide individuals with further motivation for continuous improvement and commitment. In addition, it increases the likelihood that individuals who occupy higher levels of the organizational hierarchy have direct experience with at least some of the benefits and risks of a FDRS. Past experience with a FDRS may put raters in a better position to mitigate some of the risks of a FDRS (through, for example, expert power or influence from knowledge and experience). Thus, we propose that:

**Proposition 6a.** A FDRS, compared to an absolute rating system, is less likely to be effective when applied to jobs that are highly standardized and low in autonomy.

**Proposition 6b.** A FDRS, compared to an absolute rating system, is less likely to be effective when applied to jobs that significantly benefit from task interdependence and when an organizational culture strongly emphasizes teamwork.

**Proposition 6c.** A FDRS, compared to an absolute rating system, is less likely to be effective when applied to jobs low in task identity.

Proposition 6d. A FDRS, compared to an absolute rating system, is less likely to be effective when applied to jobs in which valid and objective performance criteria for making comparisons among workers are not available.

3.7.2.1. National culture. While there are variations in values across as well as within (e.g., departments) organizational boundaries, there are even greater variations in values across national boundaries (Hofstede et al., 1990). Individualism versus collectivism distinguishes members of different national cultures in terms of how they view their relationship with others (Hofstede, 1980; Triandis, 1995). In individualistic cultures (e.g., Canada, Great Britain, the U.S.), workers tend to value achievements that can be attributed to an individual. In collectivistic cultures (e.g., China, India, Japan), workers tend to see themselves as connected to or interdependent with others; hence, they typically focus on shared objectives and commonalities among their group's members.

Individualistic versus collectivistic cultural values influence attraction to and acceptance of various human resource practices (Ramamoorthy & Carroll, 1998). National cultures with individualistic values prefer performance feedback that is tied to an individual’s rewards and consequences (e.g., pay, promotion, termination). In contrast, national cultures with collectivistic values tend to reward loyalty and commitment of a group’s members. As a result, they tend to emphasize seniority over an individual's achievements; and job security over monetary incentives. Because a FDRS conflicts with collectivistic values, it is less likely to be effective in a collectivistic society.

Furthermore, a FDRS may be less beneficial in cultures that Hofstede (1980) identified as valuing: (1) equality (vs. hierarchy) because superiors have significantly more power than subordinates under a FDRS; (2) security (vs. risk-taking) because of the risk inherent in a FDRS, particularly the threat of negative consequences for relatively low performance; (3) femininity (vs. masculinity) because the decisive and aggressive treatment of performance problems may conflict with a preference for compassion and sensitivity; (4) long-term (vs. short-term) orientation because performance ratings generally do not reflect behavior beyond the most recent performance rating time period, and because a FDRS does not support long-term relationships among employees.

Proposition 7a. A FDRS, compared to an absolute rating system, is less likely to contribute to job performance within a national culture that is collectivistic.

Proposition 7b. A FDRS, compared to an absolute rating system, is less likely to contribute to job performance within a national culture that places high value on equality, security, femininity, and is long-term oriented.

3.8. Summary of integrative model

Our theoretical model integrates and extends on disparate theory and evidence to help unpack the “black box” between the FDRS and job performance. This model explains how and why a FDRS may motivate effort as well as help attract and retain high ability employees, and thereby increase task performance. In doing so, however, a FDRS may also provoke injustice perceptions and dysfunctional competition, which may in turn decrease citizenship performance and increase counterproductive performance – especially over time. As the psychological and social work environment deteriorate due to the practice’s detrimental effects on citizenship and counterproductive performance, task performance and overall job performance are likely to be negatively affected. Furthermore, we describe the conditions under which the risks of a FDRS are particularly likely to outweigh potential benefits, such as the extent to which peak performance on the job requires cooperation and collaboration among comparison others, and organizational members value teamwork and collectivism. Fig. 1 summarizes the mechanisms (e.g., mediators, moderators) through which a FDRS is theorized to influence job performance. Fig. 2 summarizes the combined effects of a FDRS on task, citizenship, and counterproductive performance over time.

4. Human resource management implications

4.1. Theoretical implications and suggestions for future research

Much of the discussion on FDRSs in the business community has been on pros and cons of the practice (e.g., Lipman, 2012, Sahadi, 2015, Sher, 2012, White, 2015). Over the past few decades, the FDRS has gone in and out of fashion, only to come back again (Cappelli, 2009). In the academic literature, evidence from Scullen et al. (2005) suggests that the practice may initially increase a firm’s workforce potential. This study was widely cited by the media as evidence for the effectiveness of the FDRS (e.g., Jones, 2005, McGregor, 2006). As we suggest through the integrative model, however, FDRSs may have numerous unintended consequences. Researchers have found that FDRSs may be perceived as less fair by potential raters (Schleicher et al., 2009) and ratees (Blume et al., 2009) and lead to more sabotage behavior (Berger et al., 2013). But whether a FDRS will be on balance beneficial or harmful for overall performance has been an open question – and the answer may be “it depends.”

But it depends on what? Given the considerable impact that FDRSs may have on the performance as well as the social and psychological well-being of employees and organizations, it is unfortunate that only a few studies have been conducted on FDRSs and the most of what has been done has employed laboratory experiments, computational modeling, university students

and vignettes (e.g., Berger et al., 2013; Blume et al., 2009, 2013; Giumetti, Schroeder, & Switzer, 2015; Schleicher et al., 2009). As a result, the academic literature has provided few empirically based insights into the impact of FDRSs in the context of real organizations.

In this paper, we aim to reconcile divergent perspectives on the efficacy of the FDRS by reasoning about the conditions under which its risks are expected to outweigh potential benefits. We encourage future empirical, field research based on our integrative model and the propositions flowing from this model. More specifically, researchers may examine the effects of job redesign in reducing risks such as injustice perceptions and dysfunctional competition, or how cultural values (using, for instance, the Organizational Culture Profile by Caldwell & O’Reilly, 1990) moderate the FDRS’s effects on job performance. For instance, while past research has found that a forced distribution may lead to greater sabotage behavior on a task requiring independent work effort (Berger et al., 2013), future research may seek to examine how a FDRS affects not only task and counterproductive performance but also citizenship performance, particularly when applied to jobs requiring varying levels of interdependency or in contexts emphasizing teamwork or collectivism. The passage of time, or perhaps more significantly the number of FDRS cycles that are experienced, is another important moderator in our integrative model. A theory that can be leveraged for further exploring the moderating effects of time is Bandura’s (1986) social cognitive theory.

Social cognitive theory posits three variables that predict, explain, and influence behavior, namely, goal setting, outcome expectancy, and self-efficacy. Goal setting makes concrete what an individual should be exerting effort and persistence to attain, such as being and being seen as a high performer in the organization. Outcome expectancy emphasizes the importance of an individual seeing the relationship between what one is doing and achieving the goal of being ranked a high performer. Self-efficacy refers to self-confidence that one can do what is necessary for goal attainment. Because the emphasis in a FDRS is on ranking an individual’s performance against other employees’ performance, an individual may not see a direct relationship between what she or he is doing and attaining the goal of being ranked a high performer. Hence, outcome expectancy may be weak. Even when outcome expectancy is moderate or strong, an individual may lack the confidence that she or he will be ranked higher than others in the group. Hence, the person may learn a sense of helplessness and give up.

We believe that our integrative model, coupled with social cognitive theory, suggests potentially important future research topics concerning the impact FDRSs might have on the top talent who are often sought by the organizational leaders who implement these systems. As discussed within our model, high achievers may be more likely than the typical applicant to embrace the philosophy underlying the FDRS. However, a FDRS, by definition, limits the number of employees who can obtain the rewards associated with being ranked in the top category. Thus, an organization may have more employees who self-identify as high achievers than a FDRS will allow them to recognize. How might this affect outcome expectancy, self-efficacy, and in turn the motivation of members of this select group? Ideally those who fail to secure a top ranking would redouble their efforts to perform at a higher level. Perhaps, at least in the short term, many of them would do exactly that, and some will likely succeed. But to the extent that they do, others who see themselves in that light must fall out of the top group.

How will motivation change, especially after a few cycles of the FDRS, in the employees who are sometimes in and sometimes out of the top ranking category – or in those who never achieve the top ranking? Will their views on the relationship between “what they do” and “what they get” change? Will they see this as distributively and procedurally unfair? Will self-efficacy suffer? Will they be any more (or less) likely than other employees to turn to dysfunctional competition? We urge researchers to study how members of this important group tend to deal with these issues and what implications this might have on their future task, citizenship, and counterproductive performance. These risks may be exacerbated by the fact that employees’ self-ratings of performance tend to be higher than ratings from others (e.g., Mount, 1984). Hence, the number of people who see themselves as outstanding performers likely exceeds the number of people who actually do perform at that level. This suggests that the issues raised here are not limited to those who truly do perform at the highest level, but are not recognized for it. It probably extends to others who mistakenly believe that they have performed at that level as well.

Future research might also examine the role of supervisors. For instance, if supervisors are overly lenient or stingy when rating performance, a FDRS may preserve incentives by forcing differentiation. As noted in our integrative model, however, a FDRS may enhance supervisors’ reward and coercive power, but it may be at the expense of expert and referent power. Therefore, researchers might also examine how a FDRS relates to supervisors’ bases of power, leader-member exchange, and managerial ethics. In addition, a FDRS may be inconsistent with values related to not only organizational and national culture, but also demography. For instance, Hofstede et al. (1990) argued that demography is a source of deeply held values because it reflects social memberships acquired early in life. Indeed, lack of fit between values related to demography and a FDRS may help explain why this practice has led to numerous lawsuits by women, older workers, and ethnic minorities alleging adverse impact (Amalfi & Steiner, 2005). Therefore, a further suggestion for future research is examining how the FDRS philosophy resonates with values that differ along demographic lines, and how this might affect an organization’s ability to attract and retain a diverse workforce. Finally, the model can be expanded to include how and why a FDRS affects outcomes beyond job performance, such as job satisfaction and organizational commitment as well as stress, burnout, and turnover.

4.1.1. Limitations of the model

In theorizing about underlying processes and how and why contextual factors influence the effects of a FDRS on job performance, we assume that the organization has invested the resources necessary for the mechanics of the practice to be technically sound – including training raters (supervisors) and using reasonably valid performance criteria. While beyond the scope of the current model, future research may seek to examine how the design of the FDRS itself and surrounding human resource practices (e.g., 360-degree feedback) can mitigate or perhaps increase risks of the practice.

4.2. Practical implications

As noted throughout this paper, the risks of a FDRS are likely to outweigh potential benefits to the extent that the practice is applied to jobs in which peak performance involves task interdependence and social support from comparison others. A practical implication is to avoid a FDRS for such jobs; or redesign the organizational structure and nature of work such that workers’ jobs are high in autonomy and task identity. Relatedly, it may be more sensible to adopt a variation of a FDRS. For instance, the practice might be applied to the subset of workers’ activities that require independent work effort, and can be measured using valid and objective performance criteria, though we note that this will likely shift workers’ behavior to activities that are measurable to the detriment of other aspects of work behavior.

In addition, a FDRS is more likely to be beneficial for an organization to the extent that serious attention is paid to other human resource practices. A successful implementation of FDRS requires that an organization not only consistently attracts, but also selects a sufficient number of qualified individuals and places them into jobs that match their knowledge and skills. Additionally, it requires the development and application of valid performance criteria. Further, a FDRS on its own does not convey sufficient information on what an employee must start doing, stop doing, or do differently to improve performance. It merely collapses performance into a single performance category. Therefore, a FDRS may represent a step back from a shift in emphasis from performance appraisals to performance management, from evaluating employees to coaching and developing them (Latham, Almost, Mann, & Moore, 2005).

To mitigate risks of the practice and support employee development, it may be beneficial for an organization to engage in human resource practices such as 360-degree feedback (e.g., peer, supervisor, customer feedback) and coaching as well as rewarding productive team behavior (e.g., via profit or gain sharing). In addition, a simulation study by Mulligan and Schaefer (2011) suggests that a FDRS with a probation period, rather than immediate termination of employees ranked in the bottom category, may help organizations realize some of the initial benefits of a FDRS, while reducing issues with fairness perceptions and voluntary turnover. In a related vein, applying the FDRS after two, three or even five years, instead of annually or bi-annually, may mitigate the risk of short-term oriented behavior and encourage more long-term planning and action. Finally, a FDRS is more likely to be sustainable to the extent that there are promotion opportunities or well-defined career ladders for those who exhibit high performance over time.

5. Conclusion

In this paper, we first reviewed research related to the FDRS. Second, we introduced an integrative model on the effects of the FDRS on an employee’s job performance. Seven propositions flow from this framework. In doing so, we proposed potential benefits as well as risks of using a FDRS. Boundary conditions were offered for better predicting, explaining, and managing the performance effects of a FDRS. Finally, based on our integrative model, we put forth implications for human resource management research and practice. In integrating and building on insights and evidence from multiple literatures, we seek to stimulate future empirical research and support for an informed choice regarding an appropriate performance appraisal system.

References


