American Interdisciplinary Journal of Business and Economics
ISSN: 2837-1909| Impact Factor : 4.6
Volume. 9, Number 3; July-Sept, 2022;
Published By: Scientific and Academic Development Institute (SADI)
8933 Willis Ave Los Angeles, California
https://sadipub.com/Journals/index.php/aijbe

FORCED RANKING GOES WRONG: EXAMINING THE DISASTROUS EFFECTS OF VITALITY CURVES

van de Poll, Theo Kroese
Managing Director at Transparency Lab BV
Founder Moving-as-One

Abstract: General Electric's Jack Welch introduced the 20-70-10 rule, also known as the vitality curve, which has become a widely used employee performance management technique. However, it has faced criticism for its negative effects on teamwork and the lack of correlation between individual employee ratings and team performance. This study explores the possibility of extending the vitality curve to team performance management and addressing the criticisms of the vitality curve. The study surveyed over 1,600 teams with over 110,000 employees using questionnaires and divided the employees into three groups (Red, Amber, and Green) within each team, similar to the vitality curve. The results showed that 40% of the teams had predominantly green employees, 40% had mostly amber employees, and 20% had a large contingent of red employees. The study concludes that it is essential to evaluate teams separately to improve performance management.

Keywords: performance management, vitality curve, team performance, employee performance, performance evaluations, rank-order evaluations, forced ranking, teamwork, employee morale.

Introduction

The introduction provides an overview of the origin and implementation of the vitality curve, including its strengths and weaknesses. Some studies argue that the controversy surrounding the vitality curve is due to poor implementation rather than flaws in its design. Others have explored the effects of rank-order evaluations on employee behavior and the relationships between employees and leaders. Companies are shifting away from ratings-based performance management, and there is uncertainty about what happens after stopping the practice. The pros and cons of using the vitality curve in performance evaluations have been discussed, with some arguing that its negative effects on teamwork and employee morale outweigh the benefits of identifying and removing lower-performing employees.

Proof of the curve

Dick Grote, a management consultant and former GE employee argues for the strengths of forced ranking as a performance management tool. He asserts that the strategy's controversy arises from poor implementation rather than flaws with its design. Grote's research and experiences demonstrate how managers can use forced ranking fairly and effectively (Grote, 2005). Gill et al. (2019) studied how organizations using rank-order evaluations responded to the rank they received, mainly through the effort

they put into their jobs after being ranked. They found that the rank response function is U-shaped, meaning the hardest working employees were ranked first or last. Kwak and Choi (2015) researched how forced ranking performance appraisals influence employees, leaders, and organizations' relationships and how discrepancies in ratings relate to turnover intention and leader-member exchange. The results showed asymmetrical and nonlinear relationships. Companies continue to shift away from ratings-based performance management, but there is uncertainty in what happens after the practice has been stopped. Rock and Jones (2015) researched 33 of the 52 largest companies that had thus far eliminated performance ratings to find out how those companies dealt with performance management afterward. Cappelli and Conyon (2018) recognize the disdain towards performance appraisals in general, but they argue their importance in determining merit pay and promotions, for example. Their evidence indicates performance appraisals being a significant part of a "relational, openended view of employment," rather than to "simply settle-up contractually based employment relationships." Discussing the advantages and disadvantages of using forced ranking to remove the lower performers in a company is futile if the process of identifying those performers is ineffective (Lawler, 2002a).

Pros & cons; why companies have turned away from the curve

While forced ranking may seem appealing, Hazels and Sasse (2008) point out that there are also consequences to using the system that may not be right for every company. In 2004, it was estimated that around one-third of employers use a forced ranking system. While it seems like a fair way to make cuts, Bob Rogers, president of Development Dimensions International, asserts that using forced ranking "has a tremendous downside in terms of teamwork, culture, competitiveness, and legal problems" (Johnson, 2004). Lipman, a Mass Mutual Financial Group manager, adds to the discourse around forced rankings as a leader who followed the system. However, he recognizes that while forced ranking has its benefits, the negative issues such as lowered employee morale were not worth it (Lipman, 2012). Stewart et al. (2010) took a closer look at the advantages and disadvantages of using a forced distribution system in performance evaluations. They concluded that an organization should assess whether such a system would be compatible with its organizational culture and be fully aware of the downsides that go along with its use. Welch's performance management system that eliminates the bottom 10 percent of employees annually has been praised and implemented by many large corporations. On the surface, this seems like the most effective way to ensure a company consistently has the best employees, and Welch reasons that those employees are also better off not staying in a company that is "bad for them."

However, according to Lawler's research, creating higher performance cannot be achieved through firing those at the bottom (Lawler, 2002b). Buckingham (2013) is vocal about forced rating systems not fulfilling the role they are intended to. He argues that forced ranking instead allows for fair compensation and value-alignment between employees and companies, but it is still a poor way of accomplishing these things. After compiling evidence that companies are dissatisfied with their ratings-based performance management systems, Rock et al. (2014) conclude that these systems are overly convoluted and counter-productive. Instead of improving practices, neuroscience research suggests that these performance management systems are grounded in a misconstrued understanding of human responses and damage employee performance. At the turn of the 21st century, a forced ranking created quite a controversy, with various lawsuits filed against large companies such as Microsoft and Conoco (Boyle, 2001).

Stack ranking is widely discussed, with most people agreeing it has more negative attributes than positive ones. However, companies like IBM and Amazon continue to use stack rankings. There are significant disadvantages to using the system in performance appraisals, primarily if an organization relies on creativity and innovation. However, positive outcomes depend on the type of company (PerformYard, 2019). Forced ranking systems have been called out in various lawsuits as discriminatory, but Gary (2001) makes a case for how forced rankings can be used more effectually. Mulligan and Schaefer (2011) stated

that "systems with probationary periods before termination may realize some of the gains in workforce performance potential that traditional 'rank and yank' systems pose, while also having the potential to increase fairness perceptions." At the Society for Industrial and Organizational Psychology 2015 conference (https://www.siop.org/), a lively debate was held on performance ratings. The panelists outlined the main points for eliminating the rating systems, including discrepancies between multiple raters on the same performance and inconsistencies in the appraisals' effects on performance. But they also indicated reasons for continuing the practice – such as understanding the need to improve performance management and recognizing that evaluating performance in some way is still necessary (Adler et al., 2016).

The future of performance management

Regardless of whether companies have official performance reviews or not, employees are still being rated one way or another. Without a formal system, there is no transparent or fair way to decide whom to give raises and promotions. With data from Facebook stating that 87% of their employees approve of keeping the rating system, Goler et al. (2016) offer a defense for the ranked performance reviews, arguing that people want to know their place in a company. Among employees in the Australian Public Service, anxiety is the main feeling about performance evaluations. The research done by Blackman et al. (2015) demonstrates the need to rethink how performance management is handled in these organizations to reduce anxiety among workers. Even though the performance management revolution began years ago, many organizations still use traditional performance management like rankings and annual reviews. There are improvements, but according to PerformYard (2020), "What defines performance management as 'modern' is not your process, but your approach." Various attempts have been made to improve performance management systems, but many still see it as a nuisance and ineffective towards what it is supposed to achieve. Pulakos et al. (2015) outlined their reform to performance management, focusing on everyday practices and experiential learning. Annual employee performance evaluations continue despite ongoing evidence of their ineffectiveness and the organizational shift in companies wanting different skill-sets that do not mesh with industrial-era performance management systems. According to Ewenstein et al. (2016), the problem in adapting these systems lies with the uncertainty of what would come next. They offer insights into the origins of ranked performance management systems and suggestions on how companies may adapt to the future - warning that it is necessary to do so quickly.

Can the vitality curve be extended to teams?

In their new book Teams That Work (2020), Tannenbaum and Salas share how in a modern workplace that is all about teams, there are ways to drive a team's effectiveness - whether a person is a team leader or a team member. In a study done by Scott and Einstein (2001) involving team-based organizations, they found that effective leadership is needed to make performance systems work in teams after identifying three types of teams and analyzing them based on performance appraisal characteristics. While many companies today praise teamwork, few have performance management systems that assess those teams. Darino and Johnson's (2020) research suggests that evaluating teams as a unit could improve performance management systems that organizations need. In a workplace study by Google, it was found that when working in teams, individual ratings did not correlate to the team's performance as a whole (Duhigg, 2016). Deloitte's_2016 study showed that today's workplace is no longer based on hierarchy but a "network of teams." Things have changed rapidly, with companies modifying everything from job descriptions to the role of leaders. Moreover, companies' main issues revolve around these new work methods (Bersin, 2016). In team-based organizations, forced ranking performance systems should be terminated, and instead, the adoption of new performance appraisals is needed that emphasize and promote the new team-based structures (Dulebohn & Murray, 2019). Despite teams being far more prevalent in the workplace, they are often ineffective.

Aguinis et al. (2013) propose six ways performance management systems should be redesigned to fit both an individual- and team-centered workplace - such as measuring and rewarding both individual and team

performance and methods for implementation. Summarizing the abovementioned authors, the vitality curve's positive contribution is its rule-of-thumb to segment the workforce in high-, medium-, and low-performers. The main criticism is the unintended consequences of taking this rough 20-70-10 guideline as an iron principle for rewards and' punishments.' And then, management attention to teams is not only about rewarding. Predominantly not. It's about understanding which team is moving well towards particular management objectives and which teams aren't. There is no firing of teams; there is an intervention, coaching, and, if needed, turn-around management. A performance rule-of-thumb for teams would be handy for dividing teams into high-, medium, or low-performing teams and focusing management attention. However, none of the abovementioned authors mention specific percentages as a basis for such a rule-of-thumb.

OBJECTIVE

This paper aims to derive a new rule-of-thumb percentage for dividing high-, medium- and low-performing teams: a vitality curve for teams. We based the division among teams on their relative position within strategic topics in their organization. And we tallied verifiable facts and behavior about the team's actual situation.

METHOD

Procedure and participants

We deemed measuring team performance on financial indicators too tricky. A five percent increase in some indicator might be a low performance for one team but a stellar performance for another team. Moreover, every team likely had its specific history and context. Hence, we wanted to objectively compare teams on their activity towards a particular management objective rather than achieving it. So, we first designed an alternative scale based on the Guttman scale (Gutman, 1950), specifically designed for employee polling (Van de Poll, 2018 and 2021).

Next, we researched 328 relatively strategic assessments that would require focusing on management's attention. These assessments included topics on - among others - employee engagement, innovation, technology adoption, digital transformation, work pressure, value adoption, team effectiveness, diversity, purpose, IT security, work processes, competencies, creative agency management, and marketing excellence. These assessments involved 1,671 teams from various industries (both profit and non-profit) in 18 countries. These teams comprised 113,454 employees, answering close to 5.8 million questions.

Measures

Comparing teams on their progress towards a management target require tallying verifiable facts or behavior, not gathering opinions or agreements with statements. Hence, we replaced the traditional Likert survey format in favor of a survey based on a Guttman scale. Guttman scaling works with "current-status data" (Diamond, McDonald, and Shah, 1986): every following answer shows more progress than the previous answer. It's a scale from not so good to better to even better, so-called breaking points (Uhlaner, 2002). Q. How do you celebrate successes?

1. *We don't*
2. *When there is an apparent reason to do so, with whoever is involved*
3. *We make it a habit to celebrate successes with the entire team*

As in this example, such answers can be considered 'objectively real' or 'a testable proposition' (Ahrens & Chapman, 2006). We eliminated adjectives and adverbs that cannot be verified (e.g., "good") to reduce interpretation bias. And we added proof-words" like, e.g., 'periodically,' 'formally,' 'measurable,' 'documented,' and 'described' to reduce self-reporting bias by the respondent (Donaldson and Grans-Vallone, 2002). Additionally, such "proof-words" help with verification and prevent employees from adding cognitive or emotional meaning (Frese & Zapf, 1988). For tallying a team's progress, we need

'binary (no/yes), numerical or categorical representations' for our intended clustering (Plewis & Mason, 2007).

Data analysis

Each Guttman-Poll question had three answers. The 'worst' answer of three (the current situation) was rated with 0—the 'middle' answer (the intermediate step) with a score of 5. And the 'best' answer (reflecting the content of the strategy that needed to be achieved) with a score of 10. An average score for a (part of a) team required averaging the respondents' scores on the individual questions. We refrained from weights among questions and answers.

Where Welch's vitality curve focused on the ends of the performance Bell-curve (the top 20% and the bottom 10%), we postulated that high performing individuals ('Green') scoring 6.0 or higher on a scale of 0 to 10 (on average, slightly above the middle answer of three). Low-performing employees ('Red') scored 3.0 or lower (roughly halfway between the worst and middle answer), and the remainder of the team was 'Amber' (scoring between 3.0 and 6.0). We described each team in three percentages (the % green, amber, and red employees). Next, we clustered the 1,671 teams for these three percentages via a K-means algorithm (randomly initialized, 20,000 iterations) and repeated this for three, four, five, and ten clusters. We correlated some questionnaire specifics (e.g., number of respondents and questionnaire length) to verify whether these control variables influence the cluster scores. Finally, we compared several scenarios to see which % green and % red would require management attention (to identify the leading and lagging teams, respectively).

RESULTS

Table 1 shows the sample size, the division in green, amber, and red respondents, and the correlation of some questionnaire specifics that might have been influencing that division. In contrast with Welch's 20-70-10 vitality curve division, we tallied 32% green respondents, 51% amber respondents, and 17% red respondents. We then performed the K-means clustering on the 1,671 teams to see whether that would lead percentages to make up a vitality curve for teams.

Table 1 : *Sample size and correlations*

					Number of clusters				
N	Min	Max	Avg.	StDev.	3	4	5	10	*Sample size*
									See table 2
Number of questionnaires	328	Number of teams	1,671						
Number of employees	113,454								
Teams per questionnaire					1	44	5.1	7.4	
Number of employees per team					3	834	68	71.2	
% respondents per team									
% Green respondents			1%	96%	32%	19%			

% Amber respondents	2%	93%	51%	16%			
% Red respondents	1%	82%	17%	13%			
Correlation							
Number of respondents in a team					-0.02	-0.05	-0.05	0.06
Number of questions in the questionnaire					0.01	0.02	0.02	-0.19
% anonymous respondents					0.02	-0.01	-0.01	0.12
% of respondents willing to share knowledge about the questionnaire					0.06	0.05	0.05	-0.09

Min.: lowest number. Max: highest number. Avg: average number. StDev: standard deviation.

Note: this table has been submitted separately as an editable Excel file

The number of respondents fuels whether larger teams (with more diverse competencies) would perform better. The number of questions could indicate whether more extensive questionnaires make it difficult for respondents to achieve a 'green' score (as there is so much work to complete). Furthermore, a high percentage of anonymous respondents and/or a low percentage of respondents willing to share knowledge about the questionnaire topics could indicate a culture of fear or a similar situation. Such a situation would likely create more underperforming 'red' teams. However, these four assessment specifics did not correlate with the clusters. We have detailed the clusters in Table 2, which shows for each cluster the class centroids. We have estimated a 'color verdict' per cluster based on these centroids.

However, giving such a verdict remains a difficult task. For example, cluster 3 in the 10-cluster analysis has 35% green respondents, 30% amber, and 34% red respondents. Would that be a draw? Aware of a risk of overfitting, we assume the rule-of-thumb division of individual respondents as 30-50-20 (rather than 32-5117), as depicted in Table 3.

Table 2: *Management scenarios, based on RAG cut-offs*

Scenario name	**Employees %**			**Division of teams (%)**			
	Green	**Amber**	**Red**	**Green**	**Amber**	**Red**	**Double**
Database							
Actual average	32%	51%	17%	35%	38%	21%	5%
Rule-of-thumb	30%	50%	20%	40%	40%	20%	
Management scenario							
Ensure top-end	50%		30%	19%		18%	
Alert on low-end							
Both ends, heavier	25%		25%	32%		12%	4%
Both ends, lighter	15%		15%	31%		18%	11%

Employees %, green: the minium percentage of'green' respondents in a team, given a scenario. Ditto for % Amber and % Red.

Double: the employee mix tags this percentage to more than one category, e.g., Amber and Red.

Note: this table has been submitted separately as an editable Excel file

We then counted the teams with at least 30% green respondents, the teams with at least 50% amber respondents, and those with at least 20% red respondents. We then came to a 35% green team division, 38% amber teams, and 21% red teams. Another 5% of teams would qualify for more than one color. E.g., a team with 60% amber respondents and 40% red respondents would be eligible as both amber and red. Further analyzing the 5% double-count and wanting to achieve a rule-of-thumb division (rather than overfitting our model), we arrive at our 40-40-20 division. The original vitality curve (for individuals) was developed to focus management attention (i.e., promoting and firing). A variety of reasons can drive management attention for teams. Table 3 shows four different scenarios of how management would like to

focus their attention, scenarios that help deviate from the 'standard' 40-40-20 rule-of-thumb. For instance, management may focus on the teams where at least 50% of the respondents are green. In such a situation, they have to focus on 19% (say, a rule-of-thumb of 20%) of their teams. Figure 1 helps identify the approximate percentage of teams given a starting percentage of either green, amber, or red respondents.

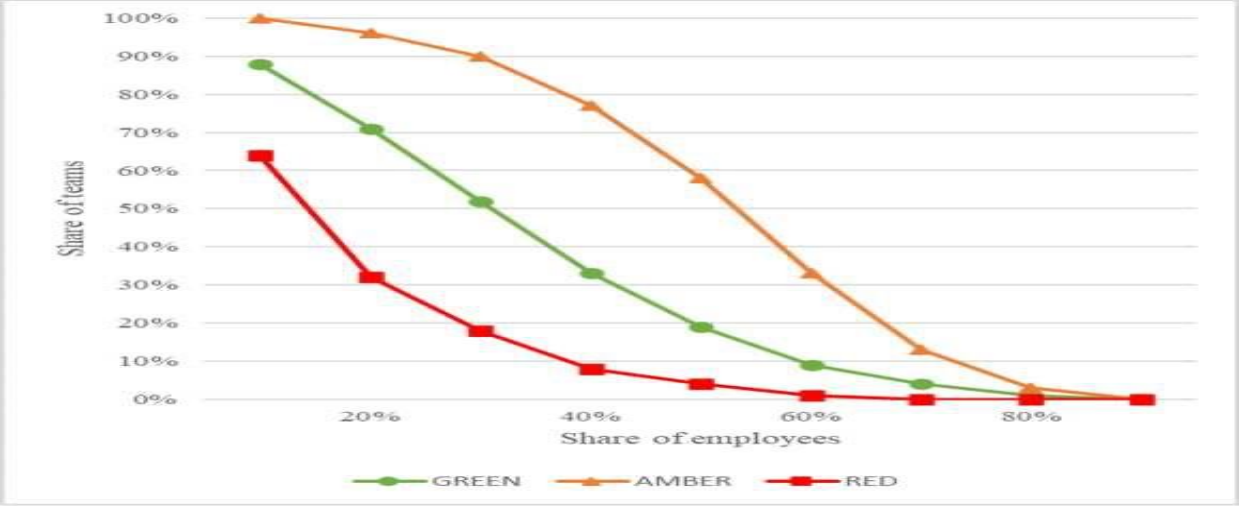

Figure 1. Linking employees to teams

DISCUSSION

Individual employees' percentual differences in the original vitality curve are based on a composite index of hard performance figures (say, reaching a sales target) and softer ones (e.g., client satisfaction or comparison with peers). As mentioned in the introduction, this has led to criticism about the ineffective identification of performers due to, among others, discrepancies between multiple raters on the same performance. Hence, our use of a survey scale focused on verifiable facts and behavior comparable within and among teams.

The consequences of this 40-40-20 rule-of-thumb are different from the original vitality curve. The main reason: teams as a whole – contrary to employees – do not get fired. Firing the bottom 20% of the teams makes no sense. As mentioned before, individual ratings did not correlate to the team's performance as a whole (Duhigg, 2016). Maybe a manager gets replaced. In the long run, bad performing functions can be outsourced. Yet, in all cases, this 40-40-20 rule-of-thumb helps to focus management attention, whatever their attention scenario.

CONCLUSIONS

The original 20-70-10 vitality curve, developed for employee performance management, has done many good things (from an efficiency perspective) and many bad things (from a human perspective). Although there is abundant literature suggesting that evaluating teams as a unit could improve performance management systems that organizations need, specific percentages to drive a team's effectiveness have not been proposed so far. In this study, the vitality curve for teams helps to focus management attention further. Similar to Jack Welch's vitality curve for employee performance, we divide team performance into three groups as Red, Amber, Green. Roughly 40% of the teams had predominantly green employees, 40% mostly amber employees, and 20% had a large contingent of red employees. The number of teams, respondents, and countries represented in our database could indicate sufficiently broad applicability of this 40-40-20 rule-ofthumb. Future studies will investigate whether technology-related questionnaires usually had 'greener' teams than process- or people-related topics.

ACKNOWLEDGEMENT

We would like to thank Dr. Jasna Duricic for her constructive comments.

REFERENCES

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

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Avoiding a "me" versus "we" dilemma: Using performance management to turn teams into a source of competitive advantage. Business Horizons, 56(4), 503-512.

Ahrens, T., & Chapman, C. S. (2006). Doing qualitative field research in management accounting - positioning data to contribute to theory. Accounting, Organizations and Society, 31, 819-841. DOI: 10.1016/j.aos.2006.03.007

Bersin, J. (2016). New research shows why focus on teams, not just leaders, is key to business performance. Forbes. https://www.forbes.com/sites/joshbersin/2016/03/03/why-a-focus-on- teams-not-just-leaders-is-the-secret-tobusiness-performance/?sh=51ed44f424d5

Blackman, D., West, D., O'Flynn, J., Buick, F., & O'Donnell, M. (2015). Performance management: Creating high performance, not high anxiety. In J. Wanna, H. Lee, & S. Yates (Eds.), Managing Under Austerity, Delivering Under Pressure: Performance and Productivity in Public Service (pp. 79102). ANU. Press.

Boyle, M. (2001). Performance reviews: Perilous curves ahead. Fortune Magazine.https://money.cnn.com/magazines/fortune/fortune_archive/2001/05/28/303851/index.htm

Buckingham, M. (2013). Performance management and the pony express. Harvard Business Review Digital. https://hbr.org/2013/11/performance-management-and-the-pony-express

Cappelli, P., & Conyon, M. J. (2018). What Do Performance Appraisals Do? ILR Review, 71(1).

Darino, L., & Johnson, C. (2020). From me to we: the next shift in performance management. McKinsey & Company Organization Blog. https://www.mckinsey.com/business- functions/organization/our-insights/the-organizationblog/from-me-to-we-the-next-shift-in- performance-management

Diamond, I. D., McDonald, J.W., & Shah, I.H. (1986). Proportional hazards models for current status data: application to the study of age at weaning differentials in Pakistan. Demography 23(4), 607-620.

Donaldson, S. I., & Grans-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. Journal of Business and Psychology 17(2), 245-260.

Duhigg, C. (2016). What Google learned from its quest to build the perfect team. The New York Times Magazine. https://www.nytimes.com/2016/02/28/magazine/what-google-learned- from-its-quest-to-build-the-perfectteam.html?_r=0

Dulebohn, J. H, & Murray, B. (2019). Leadership, performance management and team centric organizations: The importance of alignment. In D. L. Stone, & J. H. Dulebohn (Eds.), The only constant in HRM today is change. IAP.

Ewenstein, B., Hancock, B., & Komm, A. (2016). Ahead of the curve: The future of performance management. McKinsey Quarterly, 2, 64-73

Frese, M., & Zapf, D. (1988). Methodological issues in the study of work stress: Objective vs subjective measurement of work stress and the question of longitudinal studies. In: C. L. Cooper, & R. Payne (Eds.), Causes, Coping, and Consequences of Stress at Work (pp. 375-411). Wiley & Sons, Chichester.

Gary, L. (2001). The Controversial Practice of Forced Ranking. Harvard Management Update, 6(10). https://hbswk.hbs.edu/archive/for-whom-the-bell-curve-tolls-the-controversial-practice-of-forced-ranking

Gill, D., Kissova, Z., Lee, J., & Prowse, V. (2019). First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision. Management Science, 65(2), 494.

Goler, L., Gale, J., & Grant, A. (2016). Let's not kill performance evaluations yet. Harvard Business Review, 94(11), 9094.

Grote, R. C. (2005). Forced ranking: Making performance management work. Harvard Business School Press.

Guttman, L. (1950). Chs. 2, 3, 6, 8 and 9 in SA Stauffer, L. Guttman, EA Suchman, PF Lazarsfeld, SA Star and JA Clausen. 1950. Studies in Social Psychology in World War II, 4, 46-90.

Hazels, B., & Sasse, C. M. (2008). Forced ranking: A review: Quarterly journal. SAM.Advanced Management Journal, 73(2), 35.

Johnson, G. (2004). Forced ranking: The good, the bad, and the ALTERNATIVE. Training, 41(5), 24-34.

Kwak, W. J., & Choi, S. B. (2015). Effect of rating discrepancy on turnover intention and leader member exchange: APJM. Asia Pacific Journal of Management, 32(3), 801-824.

Lawler, E. E., III. (2002a). Getting rid of the bottom 10%, sounds good but... CEO Publication. https://ceo.usc.edu/wp- content/uploads/2018/03/8_Getting_Rid_of_the_Bottom.pdf

Lawler, E. E., III. (2002b). The folly of forced ranking. Strategy+business (Booz and Company). https://www.strategy-business.com/article/20290?_ref=&pg=0

Lipman, V. (2012). The Pros and Cons of Forced Rankings: A Manager's Perspective. Forbes Media LLC http://www. forbes. com/sites/victorlipman/2012/07/19/the-pros-and-cons-of-forced-rankings-amanagers-perspective.

Mulligan, J. R., & Bull Schaefer, R. A. (2011). A new hope for rank and yank. Journal of Leadership & Organizational Studies, 18(3), 385-396.

PerformYard. (2019). Stack rankings: What they're good for and where they fail. https://www.performyard.com/articles/stack-rankings-what-theyre-good-for-and-where-they-fail#

PerformYard. (2020). What is modern performance management. https://www.performyard.com/articles/what-is-modern-performance-management-1

Plewis, I., & Mason, P. (2007). What works and why: combining quantitative and qualitative approaches in large-scale evaluations. International Journal of Social Research Methodology, 8(3), 185-194.

Pulakos, E. D., Hanson, R., Arad, S., & Moye, N. (2015). Performance management can be fixed: An on-the-job experiential learning approach for complex behavior change. Industrial and Organizational Psychology, 8(1), 51-76.

Rock, D., & Jones, B. (2015). What really happens when companies nix performance ratings. Harvard Business Review Digital. https://hbr.org/2015/11/what-really-happens-when-companies-nix-performance-ratings

Rock, D., Davis, J., & Jones, B. (2014). Kill your performance ratings. strategy+business. https://www.strategy-business.com/article/00275?pg=all

Scott, S. G., & Einstein, W. O. (2001). Strategic performance appraisal in team-based organizations: One size does not fit all. Academy of Management Perspectives, 15(2), 107-116.

Slater, R. (1998). Jack Welch and the GE Way: Management insights and leadership secrets of the legendary CEO McGraw-Hill.

Stewart, S., Gruys, M., & Storm, M. (2010). Forced distribution performance evaluation systems: Advantages, disadvantages and keys to implementation. Journal of Management & Organization, 16(1), 168-179.

Tannenbaum, S., & Salas, E. (2020). Teams that work: The seven drivers of team effectiveness. Oxford University Press.

Uhlaner, L. M. (2002). The use of the Guttman scale in development of a family business index. (No. H200203). EIM Business and Policy Research.

Van de Poll, J. (2021). An Alternative to the Likert Scale When Polling Employees. International Journal of Business and Management, 9(5), 239-244.

Van de Poll, JM (2018) Ambition patterns in strategic decision-making, Doctor of Philosophy, Industrial Engineering and Innovation Sciences, Technical University Eindhoven

Welch, J., & Byrne, J. A. (2001). Jack: Straight from the gut. Warner Books.

Welch, J., & Welch, S. (2005). Winning: The ultimate business how-to book. Harper Business.