# On Normalization Performance Scores Models: An Illustrative Case Study

Mncedisi Michael Willie

## Abstract

Problem Statement: Performance Management System (PMS) applies to all companies. It is a system that has been in existence for decades and, yet Human Resources professionals and managers have the difficult task of ensuring that it produces results intended for. One of the limitations currently is that models used to measure performance are subjective and methodologies such as normalization of performance scores are not applied consistently nor have some limitations.Methodology: This study design was a retrospective case study on a one-year performance review data.The hypothesis in the current study was that the modified normalization performance scores models reduces bias and performs better than the normalization score models. Final year-end performance scores for individual employees were used to assess four models. Results:The results showed no significant differences between the four models. Therefore, the modifying normalization performance scores did not improve the model. These results also revealed precincts of forced distribution such as the size of the business unit or organization and lastly, the employeesupervisor consequence.

# 1 Introduction

erformance Management is a process of defining clear organizational objectives for employees and regularly review their actual performance against set targets. One of the vital stages in the process is to eventually reward high performers and also identify nonperformers with an objective of employing interventions to help them improve. High performers are generally rewarded in monetary or non-monetary form. Rewarding of high performing employees is subject to policies and performance standards that are defined at organizational level. Effectiveness of organizations is achieved through improving the performance of staff by continuously developing their capabilities.

Performance management remains an important aspect of connecting people management to the overall performance of the organization. There is extensive literature that links performance management to the overall strategy of the organization **??**Callaghan, 2005;Adler, 2011;Chau, 2008). According to Saravanja (2010), Performance Management has to be approached from an integrated perspective, where there is synergy between the performance management system and strategic planning. PMS is an important part of the performance management process as these systems consist of measuring and monitoring the achievement of the goals through clearly defined key performance indicators.

In recent years South African private companies and most government entities have increasingly started to link reward to performance **??**Callghan, 2005; **??**hengu, 2012). On the international front, large organizations are achieving better results and employee engagement by linking reward directly to performance (Shah et al., 2012;Armstrong, 2010). A study by Lawler et al. (2012) found that bonuses and salary increases tied to performance appraisals are associated with better organizational performance.

O'Callaghan (2005) listed factors that are often not addressed in organizations and eventually upshot in a destructive performance management processes. The author further specified that performance management

1

should be a process that incorporates the following: Another body of literature depicts performance management process asanintricate process due to some reasons, one of them being that the direct reward (or the withholding thereof) for performance may impact on the employee's motivation to perform better (or worse).

Furthermore, a performance reward management system that lacks objectivity might become unsustainable or controversial.

Leneburg (2012) discussed the methods and factors that may adversely impact the objectivity of PMS. The four rating errors described by the author include P strictness, leniency, central tendency, the halo effect and, recent events. The rating scale method is the most common method of recording and evaluating employees and for deciding promotions and annual increases. These methods continue to attract controversy due to bias as well as inconsistencies when implemented.

Normalization of scores commonly compares and standardizes performance scores of individuals belonging to different business functions in an organization. A recent study by Sarkar et al (2011) proposed a modified methodology of normalization of scores. In an illustrative example the author found that the modified methodology reduced bias in the form of association between the rank of an individual and the organization.

A study by Vaishnav and Denos (2005) discussed limitations associated with normalization of scores in the PMS. The authors warned that a PMS that employs normalization of scores methodology needs to be adjusted for supervisor or manager effect. Zewotir (2012) argued that unless the same supervisor is evaluating all employees in the organization, then there is likely a bias effect that could possibly be introduced in the process. The author further noted that the supervisor influence were a significant factor that could not be ignored in any employees' performance appraisal.

In the current study, we conducted a comparison analysis between the normalization and modified normalization of a performance score model. The modified model was proposed by Sarkar et al. (2011) as a better model that reduces bias.

The objective of the current research was to assess one of the key pillars of an effective performance management process, namely the rewarding of performance (O'Callaghan, 2005). The hypothesis was that the modified normalization of scores methodology reduced bias and was not coupled with factors such as job complexity, variances in job functions and the supervisors' effects. For the purpose of the current article, factors such as job complexity and the supervisors' effects were not explored in detail. Therefore, the primary objective of the study was to illustrate the use of a bell curve to assess the overall performance of employees for the 2011 financial year, secondary was to compare the ordinal normalization scoring processes and the modified methodology.

# 2   II.

# 3   Methods

# 4   a) Research Population and Sample

The investments company included in the current study was a consulting firm that consisted of over a 100 employees employed across 18 business units. As a part of the performance management assessment, employees were assessed for performance reflecting the 2011 financial year. The study included both mid-year and final assessments and the average of the two scores was used in the analysis.

# 5   b) Procedure

There is comprehensive literature on performance rating methods, a study by Stewart et al (2010) describes a plethora of performance terms. These include terms like forced distribution, forced ranking system, bell curve, group ordering and normal distribution. These are often used in performance evaluation systems to rate and rank employees performance. Many organizations make use of these rating systems where performance scores of various functions are combined, irrespective of outliers (Sarkar et al., 2011). The current research adapted a methodology employed by Sarkar at al. ( ??011) and considers grading range and corresponding incentive level as depicted in table 1 below.

Table 1 further depicts that employees who obtained scores less than 46do not meet the minimum criteria for financial incentive reward and these were denoted as underperformers. Employees that obtained performance scores of more than 80 points were regarded as outstanding performers and qualified for a performance bonus factor of 10%. c) Data Analysis Method

The study design was a retrospective case study which compared four performance models, these models followed forced (normal) distribution function. The hypothesis in the current study was that the modified normalization performance score models reduced bias and performed better than the normalization score models. In this study descriptive statistics including frequencies and mean ratings scores. Final year-end performance scores for individual employees were then used to assess the three models. Significance was at 5% level and, the analysis was conducted on both (SAS, 9.2) and Stata 12.0 statistics packages.

# 6 d) Model Specification

There is extensive literature on the use of a Gaussian (Normal) distribution to measure individual performance. These practices are particularly prevalent in the field of human resources management, organizational behavior, and industrial and organizational psychology. The assumption made was that individual performance follows a Gaussian (normal) On Normalization Performance Scores Models: An Illustrative Case Study distribution in the form of a bell curve with the majority of performers clustered around the mean. This predisposed organizational practices for a while now. The normal distribution, sometimes denoted as a forced distribution would assume that there would be a small number of non-performers and a small number of high performers. The majority of individuals would be the average performers clustered around the mean ?? As per normal distribution, high performers are selected if they scored more than the average + 'Z' times the standard deviation. The 'Z' value depicts the standardized normal variable or the Z score.

# 7 Global

For example, to identify the top 10% of employees, the Z score will be 1.28155 ??Sakar et al, 2011). The normalization of scores was the methodology employed in the current research and, scores were used to determine which employees qualified for performance incentives such as bonuses or annual increases.

Normalization of performance scores was denoted by Model 1 (M1). Model2, Model 3 and Model 4 [M2-M4] are modifications of M1 and are subject to different characteristics as depicted in Equation ??.

In Table 2 below, the Z-score in Equation 1 was derived for each business unit and, the final comparable score for the respective Models were calculated for each employee as follows:

Comparable score = overall average +Z score × overall standard deviation (1) Comparative scores based on Model 3. Re-classification of business units to attain effective size per business unit. Desired number of business units was 5. Re-classification of business units 'classes' were purely based on the size effect. Therefore job complexity between professions and professionals of the level of qualification were not accounted for.

# 8 Model 4 (M4)

Comparative scores based on Model 4. Reclassification of business units to attain effective size per business unit. Desired number of business units was 4. Re-classifications of 'classes' business units were purely based on the size effect. Therefore job complexity between professions and professionals of the level of qualification was not accounted for.

III.

# 9 Results

# 10 a) Descriptive Analysis

The final analysis included a sub-sample of 94 employees out of a sample of 95 employees from 18 business units. This represented 98.9% of all employees. The average mean score was 70.3 with 95% CI (68.5, 72.1) for the sample and 70.6with 95% CI (68.9, 72.3) for the sub-sample. Table3 below also depicts a median score of 72 for both the sample and subsample. Figure 1 below depicts a distribution function of the total scores and, a Whisker Box plot for the sample which also shows an outlier. The sample was also assessed for normality and, we subsequently rejected the null hypothesis (p-value=0.0237). Therefore, performance scores of the total population does not follow a normally distributed. The identified outliers were further removed in the sub-sample data and, scores were re-tested for normality.

Table 4 below depicts Skewness/Kurtosis tests for normality which were not significant; therefore the subsample analysis scores followed a normal distribution. Normalization of performance denoted as M1 were compared to incentive levels given in Table 1. Model M2 was a modification of Model 1 as outlined by Sarkar et al. (2011). Models M3 and M4 were a modification of M1 and were based on the reclassification of business units 'classes'.

Models M3 and M4 were re-classified and the desired sample for each business unit was obtained.

This was done to test the size effect between the different business units.

In M2, M3 and M4 the Z-score for each business unit were computed and the final comparable score for respective Model was calculated for each employee as follows:

Comparable score = overall average +Z score × overall standard deviation (

Table 5 below depicts descriptive statistics computed for each model. There were no significant differences in the average scores between the four models: 70.6 95% CI (69.1-72.1) compared to 70.6 95% CI (68.9-72.3), 70.6 95% CI (69.0-72.2), 70.6 95% CI (68.9-72.3) of M1, M3 and, M4 respectively.

A noteworthy feature of the data was that there was less variation in M2 (SD=5.93) when compared to other models, which were significantly higher. The average number of employees per business unit was higher for M3 and M4, and the effect of reclassification of the business seemed to have had an impact only on M3. Normality tests for the four models are shown in Table 5 below. We cannot reject the hypothesis that M1, M2 and, M4 are normally distributed but we also cannot reject the hypothesis that M3 is normally distributed at 5% level. The kurtosis for M3 was 0.0228 with a p-value of 0.0229, which indicated that it was significantly different from the kurtosis of a normal distribution. However, we cannot reject the hypothesis that M3 is normally distributed on

the basis of skewness alone. Therefore, all four models follow a normal distribution. In Table 1 above and Table 7 below, we illustrated the model matrix of sample representation of performers per model and incentive level. The results showed that model M2 followed a contrary outcome when compared to the other models where only a few ratings obtained a 10% incentive reward. Another noteworthy feature of model M2 was that there were no 7% incentive rewards. The top-ranked (9% incentive) as depicted in Table 1;Figures ??, 4 and five below shows a comparison analysis between the four models to assess the effect on performance incentives. These results indicated that the interquartile range (IQR) was smaller for model M1 when compared to model M2, M3 and, M4. IQR of 11.6, 11.7 and 12.1 indicated a widening interquartile range.

The figures below indicated that they was bias in comparison of M1 and M2, and M3 and M4. The size effect was evident in the comparison analysis between M1 and M2, where the modified model M2 was more bias towards 3 of the 18 business units with an effect rate of 0.17.

# 11 The comparison analysis between M1 and M3

was more bias towards BB4 and BB5 with an effect rate of 0.4. Comparisons between M1 and M4 were bias towards business BBBU3 and BBBU4 with an effect rate of 0.5.

Overall, there were significant differences in the mean number of performers subject to incentive: M1 and M2 with 3.45 95% CI (2.316-4.59) vs. 3.36 95% CI (??2.11-4.61). Comparisons between M1 and M4 yielded a slightly higher with the average score of nearly three times more at 8.4 95% CI (2.28-14.52) vs. 9.0 95% CI (4.03-13.97). The sub-sample data on the four models do not have outliers and, its symmetric box implied that the scores appeared to be normally distributed.

The average number reward increased further between M1 and M4 with the average number of performer: 11.25 95% CI (3.08-19.40) vs. 11.25 95% CI (??3.31-19.19).

# 12 Discussion

The purpose of this research was to review performance management models by comparing performance normalization scores to modified performance scores. The first approach in the analysis was to test the data for adequate statistical distribution, in this case is the normal distribution. Outliers were identified and removed from the final dataset as a results are duced sample of 94 observations followed a normal distribution. The total sample used represented 98.9% of all the data. Overall performance scores in all the four models followed a normal distribution.

The study showed that when adjusting for both average and above average performers; approximately 85% of the employees were considered based on M1, and 83%of the employees were constructed from M2. Performance analysis between the two models illustrated that more than two-thirds of employee's performances were related to the reward system. There were also no significant differences at business unit levels on the number of employees who qualified for performance rewards. Therefore, the four models depicted similar results overall.

These results were not consistent with the findings of the study by Sarkar et al (2011). The author found that modification of the normalization of scores The results in this study revealed that there was bias when comparison model M1 and M2, and M3 andM4. Size effect was noted when comparing model M1 to model M2, where the modified model M2 was more biased towards 3 of the 18 business units. Comparison analysis between M1 and M3 was more biased towards BB4 and BB5. The comparison between M1 and M4 was biased towards business BBBU3 and BBBU4.

It is stated in the literature that for the normalization of scores system to have statistical validity there must be a large number of employees in the pool (Stewart et al., 2010;Abelson, 2001). A sample size of 30 or more is considered appropriate; however if fewer than that, then confidence in the predictive power of the bell curve begins to diminish sharply according to the central limit theorem. Therefore, smaller companies avoid force-fitting employees to the bell curves.

The effect of size within the business units also has an effect on the manager's social preferences. A study by Willie (2014) presented that business units with a significantly small number of employees resulted in a performance rate of 100%. This potentially indicate that there may be a positive association between actual performance of the team within the unit, job complexity or to other factors such as the managers effect which was not explored further in the current research. A study by Harbringet al.(2010) found out that the manager's social preferences on ratings had a substantial impact on the rating behaviors, these social preferences were not picked up by forced distribution. This finding illustrates a need to consider other factors that introduced bias in the PMS.

Finally, the current study noted that the normalization of scores was used across the organization irrespective of the sample size of the different business units. Stewartet al (2010) warned against the use of scores across all departments, in particular, those that differ in size and job complexity. He further narrates that such practice might be problematic and maybe an unfair comparison. In the current research work, we illustrated that modification of the normalization of performance scores did not necessarily reduce bias. There is an existing literature on alternative approaches to measuring performance other than the forced distribution, Burger (2006) depicted some of these possible alternatives. We recommend that the use of forced distribution to assess

217 performance be considered in concurrence with other relevant recent methodologies, in particular when issues of
218 bias may exist.
219     V.

# 13  Conclusion

221 The current study illustrated that despite the controversies in methodological issues such as the use of
222 normalization of scores; most organizations still implement this method. This research revealed that the
223 modification of this model did not necessarily reduce bias. Therefore, the modification of the bell-curve; such
224 as the model employed in the current study needs to take into account factors, such as supervisor's/ manager's
225 effect which need to be accounted for when rewarding employees.

226     The complexity of the job and the size of the organization, inter (intra)-differences between the businesses units
227 remain a contributing factor. The size the business units were also noted as one of the critical factors. Therefore,
228 size effect of the business units need to be adjusted for in the performance reward incentive scheme; whether
229 the motive is a reward based or penalty based, this will ultimately fail in its intended purpose of improving
230 employees' overall performance.

231     A reward system for performance remains an integrated performance management process. In the current
232 study we did not conduct a comprehensive assessment of employees who underperformed. Therefore, interventions
233 employed by companies to assist underperformers still need to be explored further.

234     Finally, alternative statistical methods can also be applied as an alternative to normalization of performance
235 scores. Advanced statistical methods such as linear mixed modelling have been applied in annual performance
evaluations. These methods have been shown to reduce supervisor's/managers based effects.    [1] [2] [3]

**1**

| Grading range Incentive level | |
| --- | --- |
| [0-45] | 0% |
| [46-55] | 7% |
| [56-69] | 8% |
| [70-79] | 9% |
| [80+] | 10% |

Figure 1: Table 1 :

**2**

| Model | Adjustments |
| --- | --- |
| Model 1(M1) | None |
| Model 2 (M2) | Comparative scores based on Model 2 |
| Model 3 (M3) | |

Figure 2: Table 2 :

**3**

| N Median Mean | | | Lower 95% CL for Mean | Upper 95% CL for Mean |
| --- | --- | --- | --- | --- |
| 95 | 72 | 70.3 | 68.5 | 72.1 |
| 94 | 72 | 70.6 | 68.9 | 72.3 |

Figure 3: Table 3 :

236

**4**

| Variable | n | Pr(Skewness) | Pr(Kurtosis) | Adjchi2(2) | P-value |
|----------|-----|--------------|--------------|------------|---------|
| M1 | 94 | 0.057 | 0.3867 | 4.46 | 0.1074 |

Figure 4: Table 4 :

**5**

| Model | Class level | | | Total score | | | |
|-------|-------------|---|---|-------------|---|---|---|
| | Number of Business functions | Average Number of Employees per Business function | Range | Mean score | | Std. Dev. | Range (Min-Max) |
| M1 | 18 | 5 | 2-9 | 70.61 | 8.23 | | 51-87 |
| M2 | 18 | 5 | 2-9 | 70.64 | 5.93 | | 57-82 |
| M3 | 5 | 19 | 8-39 | 70.63 | 8.01 | | 53-84 |
| M4 | 4 | 24 | 14-36 | 70.61 | 8.07 | | 51-87 |

Figure 5: Table 5 :

**6**

| Variable | Pr(Skewness) | Pr(Kurtosis) | adj chi2(2) | P-value |
|----------|--------------|--------------|-------------|---------|
| M1 | 0.057 | 0.3867 | 4.46 | 0.1074 |
| M2 | 0.1202 | 0.3118 | 3.54 | 0.1703 |
| M3 | 0.1122 | 0.0228 | 7.06 | 0.0293 |
| M4 | 0.1428 | 0.3022 | 3.31 | 0.1914 |

Figure 6: Table 6 :

**7**

| Incentive Level | M1 | M2 | M3 | M4 |
|-----------------|------|------|------|------|
| 0 | 0% | 0% | 0% | 0% |
| 7 | 4% | | 4% | 5% |
| 8 | 36% | 45% | 37% | 37% |
| 9 | 48% | 53% | 45% | 48% |
| 10 | 12% | 2% | 14% | 10% |

Figure 7: Table 7 :

## .1  Acknowledgements

## .2  Competing Interests

The Author declares that there is no financial or personal relationship which may have influenced him inappropriately in writing this article.

## .3  Authors' Contributions

The Author was responsible for the data analysis and drafting of the article. The Author has proofread the final manuscript.

[Burger ()] *An Alternative Method to Forced Distribution as a Means of Evaluating Employee Performance. Engineering Management Field Projects. The University of Kansas Masters thesis. United States*, Shannon M Burger . http://hdl.handle.net/1808/1046 2006.

[Mujtaba and Shuaib ()] 'An Equitable Total Rerwards Approach to Pay Form Performance Management'. B G Mujtaba , S Shuaib . *Journal of Management Policy and Practice* 2010. 11 (4) .

[Willie ()] 'Analysis of performance reward models: A Multinum Investments case study'. M M Willie . *Journal of Strategy and Performance Management* 2014. 2 (3) p. .

[Armstrong ()] *Armstrong's Handbook of Reward Management Practice: Improving Performance through Reward*, M Armstrong . http://www.ukessays.com/essays/marketing/the-factors-behind-performance-management-in-companies-business-essay.php#ixzz2QutytlF3 2010. (London: Kogan Page Limited)

[Abelson (2001)] *Companies turn to grades and employees go to court, The new York Times*, R Abelson . 2001. March. 51697. p. 150.

[Sarkar et al. ()] 'Comparison of Performance Appraisal Score: A Modified Methodology'. A Sarkar , A R Mukhopadhyay , S K Ghosh . *Research and Practice in Human Resource Management* 2011. 19 (2) p. .

[Stewart et al. ()] 'Forced distribution performance evaluation systems: Advantages, disadvantages and keys to implementation'. S M Stewart , M L Gruys , M Storm . *Journal of management and Organisation* 2010. 16 (1) .

[Desimone et al. ()] *Human Resource Development*, R L Desimone , J M Werner , D M Harris . 2002. Harcourt College Publisher. (3rd Ed)

[Svaranja ()] *Integrated Performance Management Systems'. PhD dissertation*, M Svaranja . http://etd.uwc.ac.za/usrfiles/modules/etd/docs/etd_gen8Srv25Nme4_1579_1298278773.pdf 2010.

[Zewotirt ()] 'On employees' performance appraisal: the impact and treatment of the rater's effect'. Zewotirt . *South African Journal of Economic and Management Sciences* 2012. 2012. 15 (1) .

[Baron ()] 'Out of the Tick Box'. Armstrong M Baron , A . *People Management* 1998. 4 (15) p. .

[Lenenburg ()] 'Performance Appraisal: Methods and Rating Errors'. F C Lenenburg . *International Journal of Scholarly Academic intellectual diversity* 2012. 4 (1) .

[Harbring et al. ()] *Performance Appraisals and the Impact of Forced Distribution: An Experimental Investigation*, C Harbring , J Berger , D Sliwka . 2010. (IZA Discussion Paper No 5020)

[Adler ()] 'Performance management and organisational strategy: How to design systems that meet the needs of confrontation strategy firms'. R W Adler . 10.1016/j.bar.2011.08.004. http://dx.doi.org/10.1016/j.bar.2011.08.004 *The British Accounting Review* 2011. 43 (4) p. .

[Lawler et al. ()] *Performance Management and Reward System. Centre for Effective Organisation*, E E Lawler , G S Benson , M Mc Dermatt . 2012. Los Angeles.

[Saxena ()] 'Performance Management System'. S Saxena . *Global Journal of Management and Business Research* 2010. 10 (5) p. .

[O'callaghan (2005)] *Performance Management: SDFs in the Fasset Sector*, A O'callaghan . 2005. July. (Unpublished paper delivered at an IIL Conference)

[Armstrong and Baron ()] *Performance management: the new realities. London: Institute of Personnel and Development*, M Armstrong , A Baron . http://www.cipd.co.uk/NR/rdonlyres/64B9C9FD-8168-4EDC-9B71-DD3D8B5C2B90/0/performance_management_discussion_paper.pdf 1998.

[Shah ()] 'The Impact of HR Dimensions on Organisational Performance'. Shah . *African Journal of Business Management* 2012. 6 (4) p. .

291 [Zhang ()] *The impact of performance management system on the employee: Analysis with WERS*, Y Zhang .
292     `http://essay.utwente.nl/62260/1/Daisy-master_thesis.pdf` 2012. 2004.

293 [Bengu ()] *The perceived effectiveness of employee share options as a mechanism of talent management in South*
294     *Africa*, B N Bengu . 2009. M.Com research report. Johannesburg. University of Johannesburg

295 [Chau ()] *The Relationship of Strategic Performance Management to team strategy, company performance and*
296     *organisational effectiveness*, V S Chau . 2008. (Emerald 14)