WEP Warwick Evans Publishing

Transactions on Economics, Business and Management Research
ISSN: 2960-1789, eISSN: 2960-2254 | Volume 2
MEHSS 2023

Personnel Evaluation and Rating Errors in the U.S., China, and Germany

Yunting Zhu

Technical University of Munich, Arcisstrasse 21, Munich, Germany

zhuyuntingkelly9829@gmail.com

Abstract. This paper examines performance management in Chinese, American, and German firms, highlighting the significance of knowledge-based employees in enhancing enterprise efficiency and success. It underscores the role of performance management in optimizing employee performance, aligning with strategic goals, and promoting organizational development. From both employee and enterprise standpoints, performance management offers various benefits, such as increased motivation, individual performance improvement, operational efficiency, strategic goal attainment, and workforce harmony. Nonetheless, common errors in employee performance assessment stem from cultural disparities, inappropriate appraisal methods, communication gaps, and inadequate training and supervision systems. This paper suggests pragmatic solutions to address these issues, including fostering the right evaluation mindset, clarifying assessment objectives, strengthening training and supervision mechanisms, promoting open communication, and cultivating a positive evaluation culture. These solutions are vital in enhancing the effectiveness and fairness of performance management in Chinese, American, and German companies, thereby contributing to their sustained growth and competitive advantage.

Keywords: Performance Management; Employee Evaluation; Organizational Development.

1. Introduction

Knowledge is vital for enterprise development, primarily relying on knowledgeable employees, impacting overall enterprise performance, and affecting their success or failure. Effective performance management is essential for a company's sustainability and profitability, preventing problems in the competitive market. This study explores the performance management of Chinese, American, and German companies, delving into issues and proposing optimization strategies. Despite cultural and policy differences, commonalities in management concepts and goals persist. This research offers reference and inspiration for global enterprises and aids managers in problem identification and resolution, enhancing enterprise strength. Performance management, as defined by Murphy et al. (1990), involves improving resource use to achieve cost-effective and efficient goals (Murphy et al., 1990, p. 227). Motowidlo and Scotter (1994) distinguish task and contextual performance, and performance management balances these (Motowidlo and Scotter, 1994, pp. 475-477). Performance management influences employee behavior and aligns strategic objectives with performance management systems, enhancing strategy execution (Hansi and Madbouly, 2021, pp. 31-42). It enables effective evaluation, compensation, incentives, and promotions based on performance assessments and continuous rule enhancement (Cappelli and Tavis, 2016, para. 15-20). Performance management has a longer history in the West, evolving into a theoretical framework guiding employees toward organizational goals (Verburg et al., 2018, p. 182). Notable American scholars like Lynch, Cross, Boyett, Conn, Whitman, and Presley have contributed to performance management research, emphasizing reward mechanisms, talent development, and performance appraisal (Lynch and Cross, 1991; Boyett and Conn, 1993; Presley and Whitman, 1998, pp. 2-4). Modern theory includes retrospective and prospective elements (Pulakos, 2004, pp. 8-21; Schleicher et al., 2018, pp. 2209-2215). Performance management has diverse meanings, motivating employees, enhancing individual performance (Dieleman, 2006, pp. 1-7), and improving enterprise efficiency, unifying efforts toward strategic goals, and optimizing human resources (Sahoo et al., 2012, pp. 3-5).

2. Methods and Approaches

This paper investigates common errors in company performance appraisals based on performance management theory, comparing American, German, and Chinese companies using 209 collected samples from these countries.

The report structure is as follows: it begins with an introduction providing the research background, purpose, and significance. Subsequently, the paper outlines the literature review, research theory, research topic identification, and questionnaire design. Details include information about survey participants, written responses, rating scales, data collection procedures, and software for data analysis. Following this, the paper summarizes findings and discusses implications, applications, and suggestions.

The study employs documentary research, questionnaire research, and statistical analysis methods.

1. Documentary research: A comprehensive literature review from Google Scholar and university library databases informs the theoretical foundation of the paper.

2. Questionnaire research: Questionnaires were distributed to respondents to gather practical data, employing measurement scales and questionnaires adjusted based on prior research to minimize biases.

3. Statistical analysis: The data were statistically organized and analyzed using Excel and R Studio for descriptive analysis.

Data was collected via online questionnaire tools, including Google Forms and Wenjuanxin. The study encompassed 209 participants from China, the U.S., and Germany, excluding other countries and participants with missing values. A five-point scoring scale assessed agreement levels, with lower average scores indicating the magnitude of the problem.

Regarding demographics, 122 participants were female, 82 were male, and 5 did not specify their gender. Ages ranged from 18 to over 60, with varying years of experience in their companies. Performance evaluation methods included qualitative, combined, and quantitative approaches.

3. Results

3.1. Halo Effect

The halo effect is a psychological phenomenon that refers to a strong quality or characteristic that spreads around like the halo of the moon. The concept of the "halo effect" was first introduced by the famous American psychologist Thorndike in the 1920s and was considered a form of bias in performance evaluation (Thorndike,1920). He believed that when people make perceptions and judgments about things, they often start from a partial point of view and then diffuse to get a general impression, and these perceptions and judgments are often biased. In other words, if a person is labeled as good, he or she will be surrounded by an aura of positivity that associates various good qualities with him or her. Conversely, if a person is labeled as a bad person, he or she will be surrounded by an aura of negativity and associated with all kinds of bad qualities. Since then, many scholars have studied the halo effect and many psychologists have experimentally confirmed the existence of the halo effect. According to Borman, the halo effect is an evaluation bias that occurs when the evaluator is influenced by the general impression of the measured variables and does not carefully distinguish the specific characteristics of each measured variable (Borman, 1975). Decotiis argued that the halo effect is a phenomenon of ambiguity in the evaluation of different attributes due to the evaluator's unwillingness or inability to distinguish between the attributes of the evaluated person (Decotiis,1977). Murphy and Reynolds argued that the halo effect is a redundant correlation beyond the true correlation between the attributes (Murphy & Reynolds, 1988).

In this paper, the authors consider the halo effect as a form of bias. When the evaluator forms an overall impression of a person, it leads to a bias in his or her evaluation of specific latitudes. This overall impression leads to the evaluator's evaluation of other independent dimensions converging with the overall impression, even when there is enough information to evaluate the independent dimensions.

3.2. Results Analysis

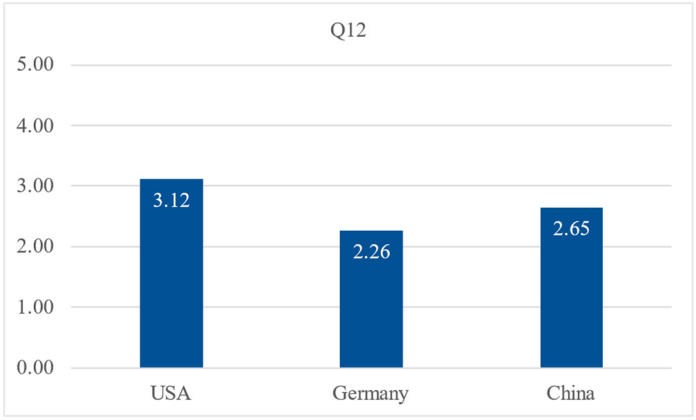

Figure 1. Mean of halo effect data

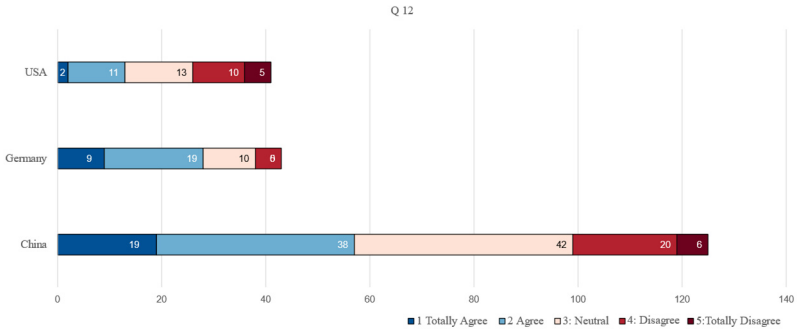

Figure 2. Halo effect collecting data

As shown in figures 1 and 2, in this survey, the writers received 41 samples from the United States. The mean of this set of data is 3.12. Simply speaking, U.S. respondents do not believe their company has a halo effect. Their managers make fair and reasonable assignments based on everyone's workability when allocating tasks.

Among the 43 samples from Germany, the average is 2.26, indicating that the halo effect exists in German companies and is more serious than in China. Specifically, more than half of German respondents agree that managers generalize their traits and extend them to all the other aspects of the work. To our surprise, no one from Germany selected "totally disagree."

Lastly, of the 125 respondents from China, 57 participants believe that if they have a clear advantage over other colleagues in one specific area, their manager will give them responsibilities in all aspects of their work, even if they are not good at other areas. However, a small group of people, accounting for 20%, still believe this error does not exist in their company. It means that their boss assigns tasks individually based on the traits of each assignment rather than on a clear advantage over other colleagues in one area. The average score of our samples from China is 2.65 out of 5. As mentioned above, the average value is less than 3, which proves that the problem is indeed in China.

3.3. Horn Effect

The opposite of the halo effect is called the horn effect. The halo effect attributes positive qualities, whereas the horn effect attributes negative qualities. Horn effect often occurs when an employee is

highly incompetent in one area, and the supervisor rates the employee correspondingly low in all areas.

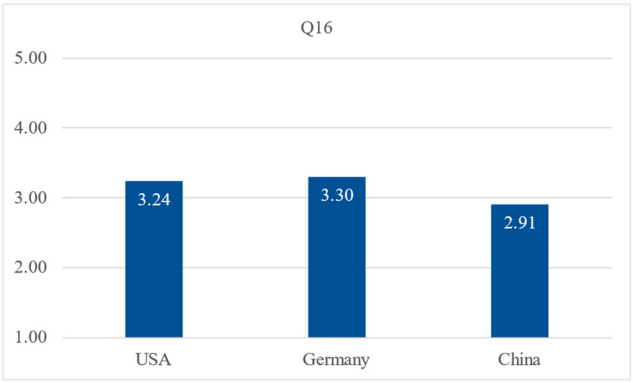

Figure 3. Mean of horn effect

As the bar char (figure 3) shows above, in the study, the average score of 125 Chinese respondents is 2.91, which is less than 3, indicating that this problem does exist in Chinese companies. Employees are concerned about making mistakes in daily work because any tiny mistake will lead their managers to deny their contribution to the team. Apart from China, the average score in Germany and America are 3.3 and 3.24, respectively. This result means that USA and Germany do not have the horn effect. Due to the mean in Germany is higher than that in America, which means that this problem is less serious in Germany.

When we look at the distribution of scores for each country in the figure 4, we can see that 45 Chinese respondents chose 1(totally agree) and 2 (agree), accounting for 36% of the 125 participants. As for German companies, more than half of the respondents concentrated their ratings on 4 (disagree) and 5 (totally disagree), reaching as many as 25. Similarly, U.S. respondents do not consider this problem seriously in U.S. companies, with just three respondents believing that they are influenced by the horn effect and getting unfair performance appraisals.

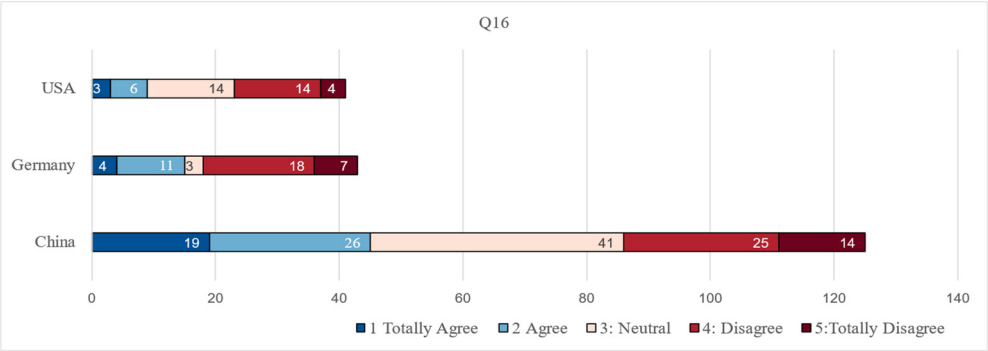

Figure 4. Halo effect collecting data

3.4. Similarity Error

Figure 5 compare whether employees in Chinese, American and German companies consider what their managers think when filling out their evaluation forms. From these three pie charts, it is clear that in the U.S. and Chinese companies, most people consider what their leaders think, with 68.29% and 56.8% of employees choosing Yes, respectively. Even so, there are differences between the U.S. and Chinese companies, with 39.02% of those who choose Yes doing so out of personal habit in the U.S. companies, while in the Chinese companies, 42.4% of people choose Yes to please their leaders. But in German companies, the situation is the opposite, most people do not care what others think, only 34.88% of employees chose Yes, and among them, only 13.95% of them intended to receive

positive evaluations from their managers. However, the majority of people chose No, which accounts for 60.47%.

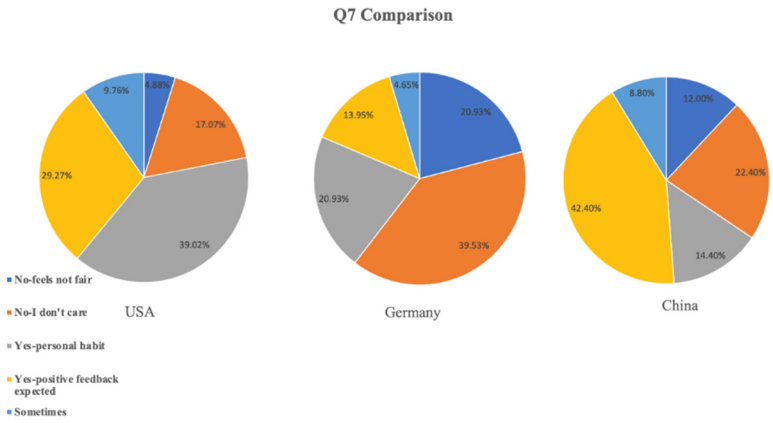

Figure 5. Comparison of similarity error

From the above, it can be concluded that as long as someone chooses the choice "positive feedback expected", there is a possibility of the existence of similarity errors, which means the evaluator tends to give positive feedback to those who are similar to himself or herself personally or professionally (MBA Skool Team, 2015, para.1). To be more specific, the problem is more serious in Chinese companies, followed by American companies, and less serious in German companies.

3.5. Stereotyping Error

The researchers want to find out if the stereotyping error is a problem in companies in different countries, which refers to the creation of a general image of members of a group based on the attributes of that person (Javidmehr et al., 2015, pp. 286-290). Therefore, to define whether this problem exists, the researchers tried to ask employees whether their managers were biased against them because of their gender, race, nationality, political beliefs, etc., and at the same time, comparing the extent of the problem in the organization in various countries.

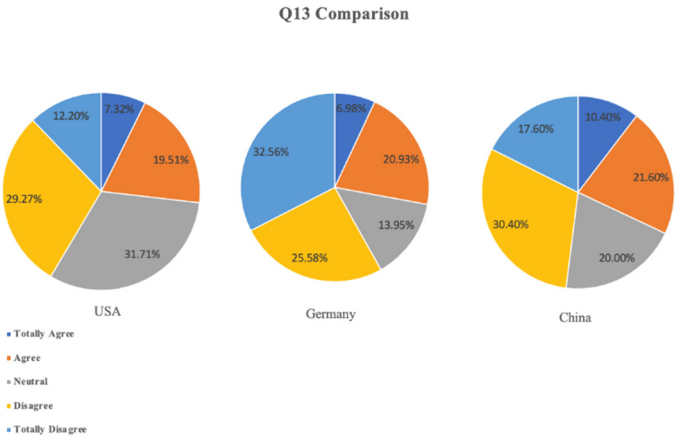

Figure 6. Comparison of stereotyping error

By looking at and comparing figure 6, it is easy to see that 26.83% of employees in U.S. companies believe that managers are biased because of their personal factors, while 27.91% and 32% of employees in German and Chinese companies agree with this statement. Meanwhile, 41.47%, 48%, and 58.14% chose to disagree with American, Chinese, and German companies respectively at the same time.

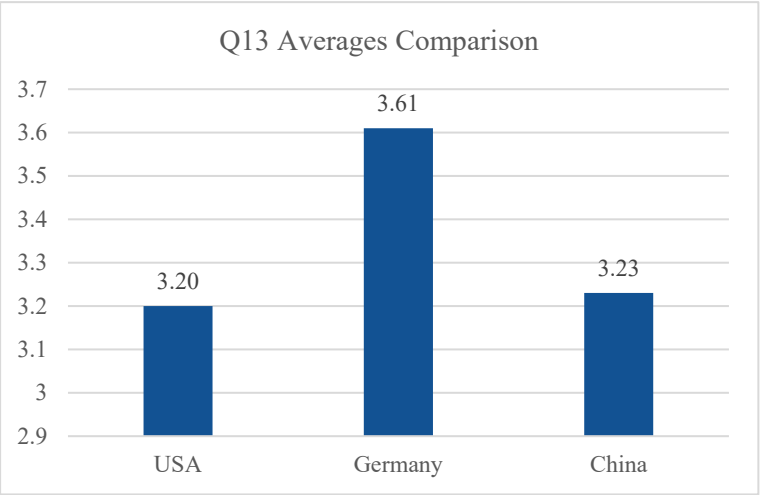

Figure 7. Averages comparison of stereotyping error

However, focusing on the ratios, as shown in the figure 7, the researchers can only conclude that stereotyping error is indeed a problem in enterprises that have a different origin, but it is difficult to compare which country's companies have this problem more severely, so a better way to determine the magnitude of this problem in these enterprises is by calculating the average. By comparing the averages (USA: 3.20, Germany: 3.61, China: 3.23), the researchers can infer that the problem of stereotyping error is not very serious, because the average of companies in these three countries are all larger than 3. However, stereotyping in American and Chinese organizations probably is more significant compared to German companies, which indicates that German companies probably are more open and democratic, although they also have this problem.

3.6. Recency Error

Ideally, performance appraisal should be based on data collected on subordinate's performance over the entire evaluation period (usually six months to a year) (Lunenburg, 2012). However, as is often the case, the supervisor is likely to put more weight on recent performance than the behaviors that occurred earlier (Lunenburg, 2012). This is so called the recency error. Failure to include all performance behaviors can skew the ratings.

The recency effect, which was first proposed by the American psychologist Luchins in 1957 (Anderson & Barrios, 1961), is one of the main phenomena in the theory of social perception bias. In the process of interaction, our most recent and up-to-date perception of others take up the dominant position, covering up the evaluation of others formed in the past (Luchins, 1957). The effect can be both positive and negative. For example, in the "Tylenol" incident in 1982, Johnson & Johnson's highly acclaimed new painkiller Tylenol has received widespread backlash after the poisoning incident. Even if consumers had a good impression of the brand at first, the original impression can be reversed due to the stimulus of the newly emerged opposite event. However, it is also because of the effective use of the positive recency effect, by taking measures including incident investigation report, media cooperation, extensive advertising, Johnson & Johnson regain Tylenol's market share within one or two years after the crisis, and even further enhance the brand image (Song et al., 2008).

Regarding the performance appraisal process, the recency error can result in a large deviation between the evaluation effect and the actual behavior or results of the evaluator (Boachie & Seidu, 2012). On the one hand, good recent performance masks substandard behavior in the past works, making the overall results qualified. Oppositely, the recent inappropriate behavior can also cover up the good or even excellent performance in the previous period. These two results can bring adverse consequences to the organization, with the former cause a falsely high appraisal results, leading to superfluous personnel within the organization, while the later bring negative emotions and even lead to dissatisfaction, making employees take a noncooperative or hostile attitude in future work. By implementing methods such as the absolute standards and results-oriented approaches, particularly

behaviorally anchored rating scale (Schwab et al., 1975) and goal setting (Locke & Latham, 1991), companies could attempt to minimize such rating errors to performance appraisal (Lunenburg, 2012).

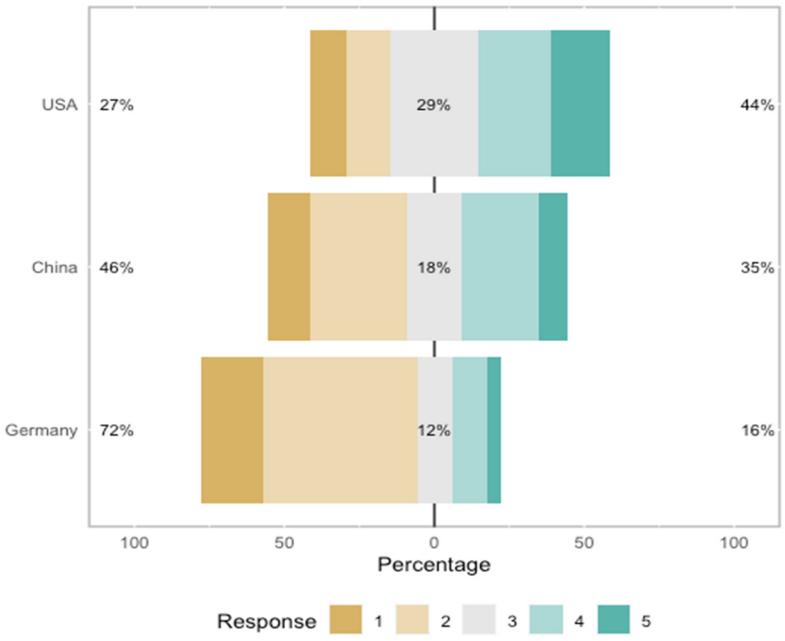

Figure 8. Recency error collecting data

The figure 8 above shows the survey results on whether recency errors happened in performance evaluation process in Chinese, the U.S., and German companies. It can be observed that as many as 72% of respondents from German companies believe that they often work harder than usual when the assessment is approaching to get more positive feedback, which shows a high possibility that a recency error existed. In the U.S. companies, however, the situation is opposite. About 44% of respondents feel that they have not been influenced by the approaching assessment to change their behavior. The distribution of ratings among Chinese companies tends to be symmetrical, with 46% of the respondents report an existence of recency error, while 35% of the respondents hold the different view.

The statistical averages in figure 9 also reveal similar information. The mean values of the Likert scale results for the U.S., Germany, and China are 3.24, 2.28, and 2.84, respectively. This suggests a recency error in Chinese and German companies, with the problem being more pronounced in Germany (Mean $_{Germany}$ < Mean $_{China}$).

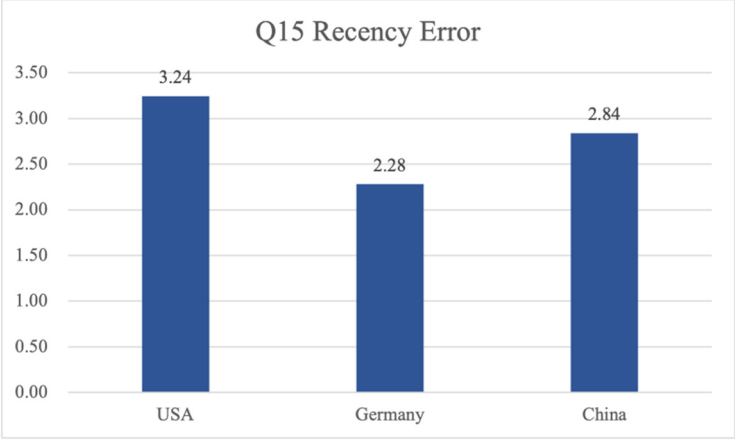

Figure 9. Means of recency error

3.7. Attribution Error

Attribution involves individuals making assumptions about the causes of someone else's behavior, typically attributing it to either internal (e.g., personality) or external factors (e.g., environment) (Heider, 1958). Attributions are based on three criteria: consensus, distinctiveness, and consistency (Kelly, 1973), aiding in determining personal, stimulus, or circumstantial attributions (Hewstone & Jaspars, 1987). The fundamental attribution error (FAE) is a common attribution error where individuals incorrectly attribute another person's actions to their personality traits rather than external circumstances (Ross, 1977). This can result from cognitive limitations, comfort-seeking, or judgment biases (Berry & Frederickson, 2015).

Organizations can also fall into this bias by explaining behaviors as dispositional rather than situational, ignoring the influence of social surroundings and institutional constraints (Swift et al., 2013). For instance, in performance appraisals, an employee's negative response may be misattributed as a negative attitude (KENJO Blog). To investigate the presence of attribution errors, respondents were asked if they were overly cautious at work due to fear that their behavior might negatively impact managerial opinions and performance evaluations. Figure 10 displays the relevant statistics.

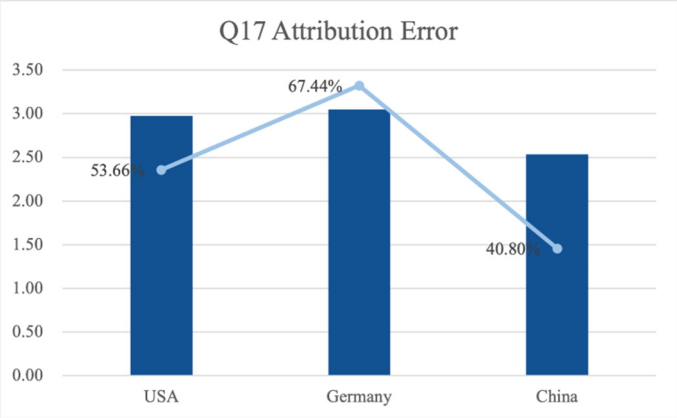

Figure 10. Mean of attribution error

From the mean value ($Mean_{US}$=2.98, $Mean_{Germany}$=3.05, $Mean_{China}$=2.54), it can be seen that attribution errors occurred in the Chinese companies. In both German and the U.S. companies the value is close to 3, which is considered to be free of significant attribution errors. When further comparing the standard deviation, it can be found that the ratings of German companies are more discrete (SD_{us} =1.21, $SD_{Germany}$=1.36), indicating a greater difference of feelings of individual employees.

In addition, in the previous survey questions, the participants also responded to whether the manager have sufficient knowledge regarding the details of the tasks when evaluating the job performance. According to figure 10, only 40.8% of the respondents from Chinese companies chose "Yes", which means that the majority of respondents think that the performance evaluators do not have enough knowledge and complete information when rating. The more information a person has, the more accurate is the inference he/she can make about that behavior (Jones & Davis, 1965). Incomplete information will likely lead to more serious attribution errors and can be a possible reason for the existing attribution error in Chinese companies. In the U.S. and German companies, the proportions are 53.66% and 67.44% respectively, which also shows a consistent trend with the Likert scale results.

3.8. Distribution Error

In the process of the questionnaire survey, we found that the following phenomenon may exist when the company conducts performance evaluation: After the appraisal, employees find that their managers' appraisal results for the same team or group of employees are very similar, and there are no major differences in either the objective description of performance within this period or the final

evaluation results, so much so that employees no longer consider the importance of appraisal a criterion for judging job completion and satisfaction.

Salary.com featured in TLNT conducted a survey, in which involved 5,970 employees, where two assessors had to supervise the work of the same group, during the same period of time. Contradictions between the two assessors were observed in 62% of all cases (Sullivan, 2011). It was generally perceived that one of the evaluators was more generous than the other (KENJO Blog). We call this phenomenon as distribution error.

Distribution errors occur when the rater tends to use only one part of the rating scale so that the results are too average, too high or too low during the evaluation process. There are three types of distribution (Denisi & Griffin, 2005):

Strictness: the appraiser evaluates all, or almost all staff, with below average ratings.

Central Tendency: the appraiser evaluates everyone with generally average ratings. No-one is actually good or bad.

Leniency: the appraiser evaluates everyone with above average ratings.

After calculating the mean of ratings from the Likert Scale, it is evidently that the mean values of the above three countries were 2.95, 3.81 and 2.92 respectively (figure 11). This data indicates that German-based companies have large probability that they do not have the distribution error in the performance evaluation process, while both Chinese and American companies are more inclined to have it, albeit not a serious one. When comparing US and Chinese companies, it is apparent that the problem is more serious in the evaluation process of Chinese companies (Mean $_{China}$ < Mean $_{US}$). The bar chart about the mean of the ratings from three countries is shown below:

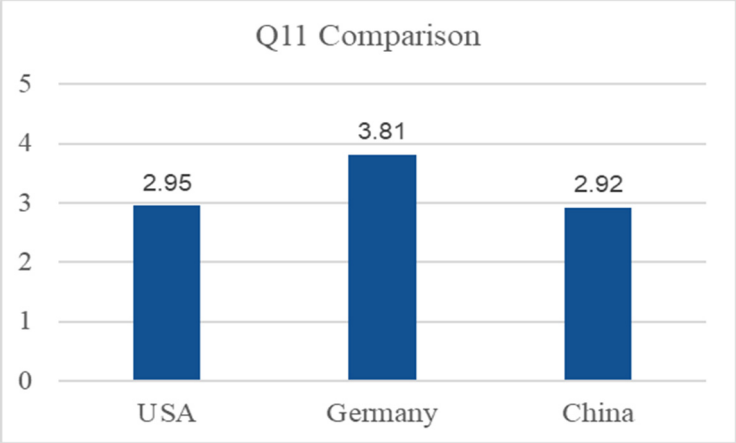

Figure 11. Mean of distribution error

In addition to judging the existence and severity of the problem from the mean value, the authors believe that the severity of the problem can also be inferred from the proportion of every rating (i.e. 1~5) from employees of the company in different countries. For this purpose, the author used the Likert package in RStudio to make a heat visualization diagram of the percentage of different ratings, which is shown in figure 12:

From the heat map it is clear to see that the absence of this distribution error in German companies has a high degree of confidence, as the ratings of the questionnaire participants are concentrated in 4 (disagree), 5 (totally disagree) with 39.5% and 25.6% respectively. Employees from Chinese companies gave a relatively even distribution of ratings, with the highest number of participants choosing ‘agree’. As for U.S. companies, although the highest percentage of criterion is 4 (disagree), we still believe that the probability of existence ofdistribution error in the evaluation process in this country’s companies is relatively high based on the basic rule of Likert Scale. However, we can see that the 26.8% of criterion 4 is an explanation that the problem is not serious for U.S. companies.

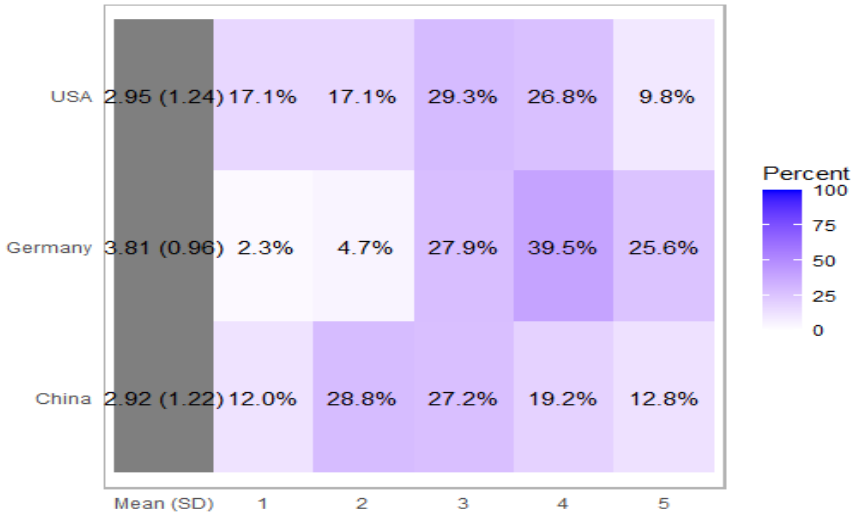

Figure 12. Distribution error collectiong data

In summary, we conclude the following that distribution error is most probably made in the personnel performance evaluation process by Chinese and American companies, but the possibility of making this error is rather slight for German companies.

3.9. Contrast Error

Contrast error is a type of rating error in which the evaluation of a target person in a group is affected by the level of performance of others in the group. When the others are high in performance, there may be a tendency to rate the target lower than is correct. When the others are low in performance, there may be a tendency to rate the target higher than is correct (APA Dictionary of Psychology). It is an error where a person sets a certain benchmark, which affects the appraisal of the candidate being interviewed (MBA Skool team, 2016).

At the same time, the authors found from the data of questionnaire that there may be contrast error in the evaluation process, as shown by the following phenomena: An appraiser compares the performance of two employees instead of using absolute performance measurements for each one. One employee who rates as ‘excellent’ could make another with a ‘good’ rating seem mediocre (KENJO Blog). In this regard, we asked the participants whether they would deliberately avoid being in the same group as a particularly good performer in their company's appraisal process, otherwise they would get a lower result than they expected.

The authors use the same thought to verify whether there is a contrast error in the personnel performance evaluation process between Chinese, American and German companies. By calculating its mean and plotting the heat map, we can see in figures 13 and 14:

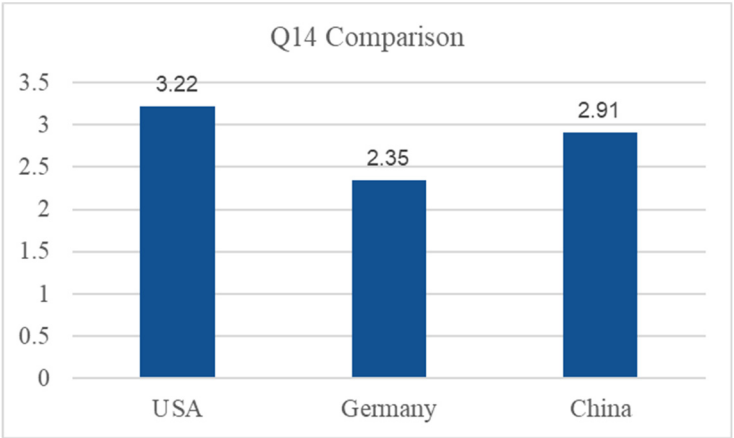

Figure 13. Mean of contrast error

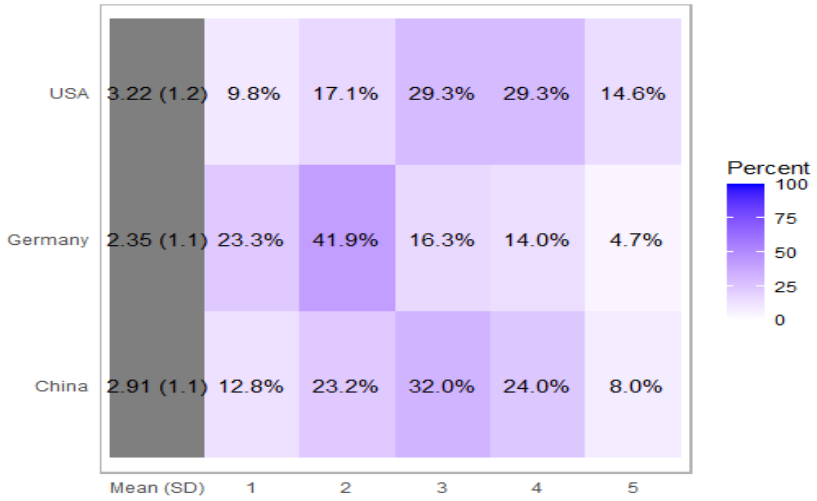

Figure 14. Contrast error collecting data

From the calculation of the mean (USA: 3.22, Germany: 2.35, China: 2.91 respectively), it suggests that both Chinese and German companies have a significant possibility that they suffer from this problem, and Germany is relatively more serious. The analysis of the heat map shows that more than 50% of the participants in the German companies chose 'totally agree' (criterion 1) or 'agree' (criterion 2), which means that more than half of employees believe that their company is conducting the evaluation process in a way that is problematic. The Chinese companies have a lesser degree of contrast problems, because although the mean is less than 3, a relatively high number of participants chose criterion 'disagree' (criterion 4). As to U.S. companies, close to 50% chose criterion 4 (disagree) and 5 (totally disagree), meaning that the problem is very limited among companies in that country.

In general, there is a high possibility that contrast error exists in the performance evaluation process for both German and Chinese companies, more so in German companies, and the error is limited in the valuation process for US companies.

4. Conclusion

From the above analysis, it becomes evident that various types of errors are prevalent in the context of employee performance evaluation within Chinese, American, and German companies. These errors encompass instances such as the similarity error, distribution error, and others. Nonetheless, it is noteworthy that certain errors, which are conventionally regarded as more frequent, are in actuality less prevalent or are confined to a limited number of companies within the assessment procedures of the aforementioned three countries. Notable among these relatively rarer errors are stereotyping errors and the horn effect.

References

[1] American Psychological Association (n.d.). Contrast Error. APA Dictionary of Psychology. https:// dictionary. apa. org/ contrast-error/.

[2] Anderson, N. H., & Barrios, A. A. (1961). Primacy effects in personality impression formation. The Journal of Abnormal and Social Psychology, 63(2), 346–350. doi: 10.1037/h0046719.

[3] Berry, Z., & Frederickson, J. (2015). Explanations and implications of the fundamental attribution error: A review and proposal. Journal of Integrated Social Sciences, 5(1), 44-57. Retrieved from https://www.jiss.org/.

[4] Boachie-Mensah, F. O., & Seidu, P. A. (2012). Employees' perception of performance appraisal system: A case study. International journal of business and management, 7(2), 73. doi: 10.5539/ijbm.v7n2p73.

[5] Borman W C. Effects of instructions to avoid halo error on reliability and validity of performance evaluation ratings. Journal of Applied Psychology, 1975, 60(5): 556-560.

[6] Boyett, Joseph H. & Conn, Henry P. (1993). Maximum Performance Management: How to Manage and Compensate People to Meet World Competition. Glenbridge Publishing.

[7] Cappelli, Peter & Tavis, Anna. (2016). The Performance Management Revolution. Harvard Business Review, Volume 94, 58-67. Retrieved from: https://hbr.org/2016/10/the-performance-management-revolution.

[8] Decotiis T A. An analysis of the external validity and applied relevance of three rating formats. Organizational Behavior & Human Performance, 1977, 19(2): 247-266.

[9] Denisi, A. S., & Griffin, R. W. (2005). Human Resource Management (2nd ed.). Dreamtech Press.

[10] Dieleman, M. & Toonen, J. & Touré, H. (2006). The match between motivation and performance management of health sector workers in Mali. Human Resource Health, Volume 4, Issue 2, 1-7. Retrieved from: https://doi. org/10. 1186/1478-4491-4-2.

[11] Habsi, Nouf & Madbouly, Araby. (2021). The Effect of Performance Appraisal Systems on Employees and Organizations in Omani Private and Governmental Institutions. Journal of Finance, Business and Management Studies, Volume 1, Issue 1, 31-42. Retrieved from: https://doi.org/10.26713/jfbms.v1i1.1753.

[12] Heider, F. 1958. The Psychology of Interpersonal Relations. New York: Wiley. doi: 10.1037/10628-000.

[13] Hewstone, M., & Jaspars J. (1987). Covariation and causal attribution: A logical model of the intuitive analysis of variance. Journal of Personality and Social Psychology, 53(4), 663-672. doi: 10.1037/0022-3514.53.4.663.

[14] Javidmehr, M., & Ebrahimpour, M. (2015). Performance appraisal bias and errors: The influences and consequences. International Journal of Organizational Leadership, Volume 4, Issue 3, 286–302. Retrieved from: https://doi. org/ 10.33844/IJOL.2015.60464.

[15] Jensen, M. C., & Murphy, K. J. (1990). Performance Pay and Top-Management Incentives. Journal of Political Economy, Volume 98, Issue 2, 225–264. Retrieved from: http://www.jstor.org/stable/2937665.

[16] Ji, T. (2021). Exploring the problems in employee performance appraisal of SMEs. Human Resources. https:// doi. org/10.13768/j.cnki.cn11-3793/f.2022.0483.

[17] Jones, E. E., & Davis, K. E. (1965). From acts to dispositions the attribution process in person perception. Advances in experimental social psychology, 2, 219-266. doi: 10.1016/S0065-2601(08)60107-0.

[18] Kelley, H. H. (1973). The process of causal attribution. American Psychologist, 28(2), 107-128. doi: 10.1037/ h0034225.

[19] Liu, X. (2021). Analysis of the construction of employee performance appraisal index system in SMEs. Management Informatization.

[20] Locke, E. A., & Latham, G. P. (1991). A theory of goal setting & task performance. The Academy of Management Review, 16(2), 212-247. doi: 10.2307/258875.

[21] Luchins, A. S. (1957). Primacy-recency in impression formation. The order of presentation in persuasion, 1, 33-61.

[22] Lunenburg, F. C. (2012). Performance appraisal: Methods and rating errors. International journal of scholarly academic intellectual diversity, 14(1), 1-9. Retrieved from https://commons.erau.edu/ww-mgmt-314/2.

[23] Lynch, Richard L. & Cross, Kelvin F. (1991). Measure up!: yardsticks for continuous improvement: how to measure corporate performance. Blackwell.

[24] M. S. T. (2016). Contrast Error - Meaning & Definition. Mba Skool. https://www.mbaskool.com/business-concepts/human-resources-hr-terms/15132-contrast-error.html.

[25] MBA Skool Team. (2015). Similarity Error - Meaning & Definition. Retrieved from: https://www. mbaskool. com/ business-concepts/human-resources-hr-terms/15222-similarityerror. html#:~:text= What%20is%20 Similarity%20 Error%3F, or%20herself%20personally%20or%20professionally.

[26] Mitchell TR, Kalb LS (1981) Effects of outcome knowledge and outcome valence on supervisors' evaluations. J Appl Psychol 66: 604–612. Available: http://dx.doi.org/10.1037/0021-9010.66.5.604. Accessed 3 August 2011.

[27] Motowidlo, Stephan & Van Scotter, James. (1994). Evidence That Task Performance Should Be Distinguished from Contextual Performance. Journal of Applied Psychology. Volume 79, 475-480. Retrieved from: http://dx.doi.org /10.1037/0021-9010.79.4.475.

[28] Murphy K R, Reynolds D H. Does true halo affect observed halo. Journal of Applied Psychology, 1988, 73(2): 235-238.

[29] Presley, Adrien & Whitman, Lawrence & Liles, Donald. (1998). A Methodology for Enterprise Performance Management. Retrieved from: https://www.researchgate.net/ publication/ 2513218_A _Methodology_ For_ Enterprise_ Performance_Management.

[30] Pulakos, E.D. (2004). Performance Management: A Roadmap for Developing, Implementing and Evaluating Performance Management Systems. The SHRM Foundation, USA.

[31] Rath, Adarsh & Candidate, Doctoral (2018). Evolution of Performance Management System: A Review of Literature. International Journal of Creative Research Thoughts, Volume 6, Issue 2, 874-884. Retrieved from: https://www.ijcrt.org/papers/IJCRT1813200.pdf.

[32] Ross, L. (1977). The intuitive psychologist and his shortcomings. Advances in Experimental Social Psychology, 10, 173–220. doi: 10.1016/S0065-2601(08)60357-3.

[33] Sahoo, Chandan & Mishra, Sukanta. (2012). Performance management benefits organizations and their employees. Human Resource Management International Digest, Volume 20, Issue 6, 3-5. Retrieved from: http://dx. doi.org/ 10. 1108/09670731211260771.

[34] Schleicher, Deidra & Baumann, Heidi & Sullivan, David & Levy, Paul & Hargrove, Darel & Barros-Rivera, Brenda. (2018). Putting the System into Performance Management Systems: A Review and Agenda for Performance Management Research. Journal of Management, Volume 44, Issue 4, 2209-2246. Retrieved from: http://dx. doi.org/ 10.1177/0149206318755303.

[35] Schwab, D. P., Heneman III, H., & DeCotiis, T. A. (1975). Behaviorally anchored rating scales: A review of the literature. Academy of Management Proceedings, 1, 222-224. doi: 10.1111/j.1744-6570. 1975. tb01392. x.

[36] Song X., Zhang H., & Yin D. (2008). Trust recovery and attitude selection in corporate brand crisis from the perspective of recency effect. Business Times, 13(2), 33-34. doi: 10.3969/j.issn.1002-5863.2008.13.018.

[37] Sullivan, J., Dr. (2011, January 31). The Top 50 Problems With Performance Appraisals. TLNT. Retrieved January 8, 2023, from https://www.tlnt.com/the-top-50-problems-with-performance-appraisals/.

[38] Swift, S. A., Moore, D. A., Sharek, Z. S., & Gino, F. (2013). Inflated applicants: Attribution errors in performance evaluation by professionals. PLoS One, 8(7), e69258. doi: 10.1371/journal.pone.0069258.

[39] The 9 most common errors in performance appraisal. KENJO Blog. Retrieved January 8, 2023, from https://blog. kenjo.io/the-9-most-common-errors-in-performance-appraisal.

[40] Thorndike, & E., L. (1920). A constant error in psychological ratings. Journal of Applied Psychology, 4(1), 25-29.

[41] Verburg R. M. & Nienabe A. & Searle R. H. & Weibel A. & Hartog N. D. & Rupp D. E. (2018). The Role of Organizational Control Systems in Employees' Organizational Trust and Performance Outcomes. Journal Article, Volume 43, Issue 2, 179-206. Retrieved from: https://doi.org/10.1177/1059601117725191.