Rater Bias and Performance Evaluation at Tech Solutions Inc.

Nicole Salem

PSY 724

Dr. Gina Possin

October 31, 2026

Rater Bias and Performance Evaluation at Tech Solutions Inc.

Tech Solutions Inc.'s audit questions whether evaluations reflect performance or willingness to work overtime. Maria receives strong ratings from peers and direct reports, but her supervisor, Mark, criticizes her commitment. Alex receives exceptionally high ratings despite having weaker technical skills. The audit shows that employees who work overtime receive higher ratings even when their core performance is similar. Mark appears to let availability influence judgments about other responsibilities. Halo, selective leniency, and contrast effects help explain how this can happen.

Rater Biases

A halo effect occurs when an overall impression or strength in one area influences judgments about other areas of performance (Murphy et al., 2018). Alex's willingness to accept additional work may lead Mark to view him favorably across the evaluation, including areas where his skills are weaker. Extra contributions do not establish technical proficiency.

Elsbach et al. (2010) found that seeing employees at work outside normal business hours could lead observers to infer commitment, even without consciously making that connection. Mark may interpret Alex staying late as dedication and carry that impression into his overall rating. Maria's lower availability could then overshadow the quality of her work.

Mark's ratings also suggest selective leniency. Leniency occurs when an evaluator gives ratings that are more favorable than performance warrants (Murphy et al., 2018). Here, the generosity appears directed toward employees who work overtime. The audit strengthens this interpretation because the pattern extends beyond Maria and Alex.

Mark may excuse technical weaknesses when an employee volunteers for additional assignments while judging similar weaknesses more critically in someone who declines overtime. Employees with comparable core performance could receive different ratings because one fits his expectations of commitment.

A contrast effect occurs when another employee becomes the reference point for an evaluation instead of the job's performance standards (Murphy et al., 2018). Alex's availability may become Mark's benchmark for team spirit. Maria could meet established collaboration expectations and still receive a lower rating because she does not match Alex's willingness to work extra hours.

The case suggests, but does not confirm, this comparison. Asking him to explain what Maria should have done differently would help clarify whether he is applying job requirements or comparing her with Alex.

Impact on Employees and the Organization

Maria's lower ratings could limit raises, promotions, and development opportunities. Unexplained criticism could also damage her trust in the company. Alex may benefit initially, but inflated ratings could leave him unaware of technical weaknesses that need attention.

If availability outweighs competence, Mark's staffing and promotion decisions may place employees in roles they are not prepared to perform. Employees may respond by prioritizing visible overtime, even when additional hours do little to improve results. Elsbach et al.'s (2010) findings help explain why appearing committed can become influential without providing direct evidence of effectiveness.

The ethical concern is an unclear expectation that may be harder for some employees to meet. HR should investigate whether employees understand overtime requirements, whether Mark applies them consistently, and whether assumptions about gender or family responsibilities influence his judgments. It should also examine whether the ratings affect pay or advancement. The case raises potential discrimination concerns, but the available facts do not establish a legal violation.

Maria's favorable 360-degree feedback should be considered without treating it as proof that Mark is wrong. Different sources observe different aspects of performance (Murphy et al., 2018). The company should review its examples and determine whether the disagreement reflects different observations, different standards, or bias.

Proposed Solutions

Tech Solutions should define technical proficiency, work quality, and collaboration separately. "Team spirit" is too vague to guide improvement. Sharing information, helping resolve problems, and completing agreed responsibilities give employees clearer expectations and managers something observable to evaluate.

The company should also clarify when overtime is required and how additional contributions receive credit. An employee who resolves an urgent problem after hours may deserve recognition for that contribution. The extra hours should not automatically raise an unrelated technical rating. Research reviewed by Murphy et al. (2018) indicates that behavioral scales can improve clarity and acceptance while having limited effects on accuracy. A revised form therefore needs changes in how managers use it.

Frame-of-reference training would allow managers to rate the same performance examples, explain their reasoning, and receive feedback against shared standards. This approach helps evaluators develop a common understanding of performance dimensions and rating levels (Murphy et al., 2018).

For this company, practice examples could include a skilled employee who completes work during scheduled hours and an employee who stays late but frequently needs corrections. Managers would have to explain which evidence supports each rating. That would make assumptions about availability easier to identify and challenge.

HR should facilitate calibration meetings where managers support ratings with specific examples from the review period. For Maria, Mark should explain which behaviors demonstrate poor collaboration and how those behaviors fall below the established standard. Reviewers should compare his evidence with feedback from peers and direct reports.

Calibration should focus on consistent reasoning rather than producing a preferred distribution of scores. This approach applies shared standards while considering differences among rating sources (Murphy et al., 2018). DeNisi and Murphy (2017) review how training, rating sources, and cognitive processes contribute to appraisal, providing a broader framework for these changes.

An independent reviewer should reassess Maria's evaluation and give both Maria and Mark an opportunity to provide evidence. Unsupported ratings should be corrected. HR should then repeat the audit after training and compare ratings among employees with similar responsibilities and core performance. It should also check whether managers document concerns and provide feedback throughout the year.

The audit suggests that overtime has become a shortcut for judging performance at Tech Solutions. Clear expectations and separate performance dimensions would help employees understand what matters. Training and evidence-based reviews would help managers recognize additional contributions without letting availability overshadow the skills and results the job requires.

References

DeNisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology, 102*(3), 421–433. https://doi.org/10.1037/apl0000085

Elsbach, K. D., Cable, D. M., & Sherman, J. W. (2010). How passive 'face time' affects perceptions of employees: Evidence of spontaneous trait inference. *Human Relations, 63*(6), 735–760. https://doi.org/10.1177/0018726709353139

Murphy, K. R., Cleveland, J. N., & Hanscom, M. E. (2018). *Performance appraisal and management*. SAGE Publications.