

## Eliminating bias from performance appraisals

Author: Felicity Menzies . November 2021

[Nine in 10 human resource leaders don't believe annual performance reviews result in inaccurate information.](#) Similarly, a [survey](#) of Fortune 1,000 companies reported that 66% of employees were strongly dissatisfied with the performance evaluations they received—71% of employees perceived that their evaluations were unfair. Line managers also lack confidence in performance appraisals. In one study, [only 15% of women and 24% of men managers had confidence in the performance evaluation process, while most viewed it as subjective and highly ambiguous.](#) Despite consistent concerns expressed by HR, employees, and managers about a lack of objectivity and fairness in appraisals, a [survey of 100 large organisations reported 57% of them they weren't taking any actions to address bias in performance reviews.](#)

### The Problem with Performance Appraisals

The [idiosyncratic rater effect](#) refers to individual-level variations in assessing the performance of others. The idiosyncratic rater effect means that an individual can receive different ratings and subjective feedback from different assessors. This is because evaluations of performance are driven by individual raters' interpretations of the meaning of assessment criteria, their own sense of what 'good looks like' for a particular competency, variations in how harsh or lenient they are in judging others, and their own inherent and unconscious social and cognitive biases. The impact of the idiosyncratic rater effect on performance appraisal is significant: [58 to 72 percent](#) of an individual's performance rating reflects assessors' characteristics, not theirs.

The idiosyncratic rater effect is more pronounced when performance evaluations are weighted towards or open to subjective assessments of performance. Subjective assessments of performance are more common in professional services settings where objective measures of performance are difficult to define and capture. Psychometricians have shown that [people don't hold stable or equivalent definitions of abstract qualities](#), such as business acumen, strategic thinking, political savvy, leadership potential, assertiveness.

Subjective assessments are also common when the appraisal process employs an 'open-box approach'. An 'open-box approach' to performance appraisal involves appraisal forms that pose broad, generic questions about employee performance and offer a blank space for managers to respond with their observations and assessments. [An example of a broad open-box question is 'Describe the ways the employee's performance met your expectations.'](#) The problem with generic questions is their ambiguity—managers are left to define or interpret what the specific expectations are for that particular

employee. [Studies](#) have shown that when performance criteria are ambiguous, people are more likely to rely on stereotypes and other biases or subjective criteria such as attitudes and personality when making an appraisal, rather than making an objective assessment of how well an employee performed their assigned tasks and exhibited desired behaviours.

Rating scales, relative to open box appraisals, are not necessarily any more objective. When rating criteria are not well-defined in terms of observable behaviours and measurable outcomes, individual raters interpret criteria and rating scales differently. One assessor's score of 3, for example, might be another assessor's score of 5. Non-numerical rating scales such as 'Exceeds, Meets, Needs Improvement' are also open to individual rater interpretations.

## Consequences of Unfair Appraisals

Biased [performance appraisals](#) negatively impact organisations and employees by:

- **Favouring the advancement of some groups over others, in turn, limits the organisation's ability to build to a diverse leadership team comprising the organisation's top performers.** Examples include the glass ceiling that prevents women from advancing to executive roles and the bamboo ceiling that holds back individuals of Asian descent from attaining leadership positions at the same rate as individuals with European backgrounds.
- **Favouring some groups for certain roles, in turn, limits an organisation's ability to build diverse workgroups across the organisation.** Examples include the clustering of women in human resource functions or administrative roles and the clustering of individuals of Asian descent in technical or technology roles. Role segregation can result in a workforce that is diverse overall, but homogenous at the department or team level. Homogeneous teams report lower levels of performance compared with diverse and inclusive teams. To reap the competitive benefits of a diverse workforce, the ultimate goal of an organisation's diversity efforts should be fostering diversity and inclusion at the level of the workgroup. Many organisations focus on tracking and improving entity-level diversity and overlook where the rubber hits the road—team diversity.
- **Limiting the effectiveness of the organisation's performance management processes for developing and optimising talent.** When supervisors and their employees do not have reliable information on employee performance, setting meaningful and relevant professional development goals is not possible and an employee's potential is not realised. An example is divergent feedback given to men and women. [Studies show that men are more likely to receive specific feedback and guidance on how they can improve their performance, whereas women are more likely to receive vague feedback](#) not linked to specific outcomes and not accompanied by actionable performance development guidelines. Other research shows women leaders often receive negative feedback that is overly focused on communication style and conflicting—[told on the one hand that they're too bossy or aggressive, but on the other that](#)

---

they should be more confident and assertive. One study reported that 76% of references to being “too aggressive” happened in women’s reviews, versus 24% in men’s.

- **Contributing to unfair compensation and rewards.** Because an organisation’s remuneration structure is typically linked to performance appraisal, bias in appraisal and performance management can lead to inconsistencies in pay and rewards whereby individuals are not remunerated fairly in a manner that accurately reflects their efforts and achievement. Fair compensation is often interpreted as equal pay for like roles and many companies claim to have achieved equal pay, at least for gender diversity. Fewer companies track and manage compensation by other traditionally marginalised dimensions including race/ethnicity/language, sexual orientation, and disability. Fair compensation, however, goes beyond equal pay for like roles. Pay inequity occurs when some groups have access to higher financial rewards and non-financial benefits because they belong to an identity group that is more commonly perceived to be a better ‘fit’ for higher-paid roles in the organisations. Consider, for example, the ‘Mad-Men’ (North American TV series set in the 1960s) scenario where women typically earn lower administrative salaries and the men typically earn higher commercial salaries. Sadly, almost 60 years on, gender role segregation persists globally in traditionally masculine industries. Pay inequity conveys disrespect because it implies that some groups of people are more valued than others because of their membership in a group. In turn, disrespect undermines employees’ perceptions of inclusion. An organisation truly committed to fairness must seek to address pay inequities across diversity dimensions such that there is no systematic difference in average pay across identity groups, whether that be gender, race, language or other. Only when pay equity is achieved can an organisation truly say it has achieved equality of opportunity.
- **Contributing to unfair dismissal.** Appraisal systems that do not provide an objective assessment of an employee’s performance can lead to erroneous performance concerns and even unfair dismissal with potential legal implications.
- **Damaging employee perceptions of their worth and value to the organisation, leading to disengagement or voluntary termination from the organisation.** When employees feel that their efforts and achievements are not fairly assessed and rewarded, they are less motivated and committed to the organisation. The [U.S. Department of Labor](#) reported that the leading reason employees leave their job is because they don’t feel appreciated. Turnover of underrepresented talent in an organisation is typically higher than the turnover of members of the dominant group, damaging an organisations ability to build a pipeline of diverse talent. Biased performance appraisals can also have a ‘self-fulfilling prophecy’ impact on diverse talent. When employees do not perceive they are valued by the organisation, they may actively refrain from applying their full and best effort or even engage in self-sabotaging behaviours. On the contrary, when employees perceive fairness in the evaluation processes, they are [more likely](#) to accept their evaluations, [in which case they will digest the information contained in the evaluations and motivate themselves accordingly.](#)

## Types of Bias That Impact Performance Appraisals

Cognitive neuroscience research has shown that most of the assessments we make, particularly regarding

people, are alarmingly contaminated by biases that sit beneath our level of unconsciousness and impact our assessments in ways in which we are unaware and most likely would deny. The existence of these unconscious and hidden biases mean that, although we intend to and truly believe we are treating people fairly and making objective assessments of merit, these unconscious biases cause us to favour some individuals and groups over others without us even realising it. Common rater biases that can influence performance appraisals include:

- **Affinity bias:** [Affinity bias](#) is our tendency to favour our own social group (ingroup) more than groups of which we are not a member (outgroups). [Studies](#) show that, in general, people extend greater trust, positive regard, cooperation, and empathy to ingroup members compared with outgroup members. We also make more favourable assessments of ingroup members compared with outgroup members. For example, we attribute the successes of ingroup members to positive character traits and high skill rather than to external causes but we attribute failures of ingroup members to situational causes rather than to undesirable character traits or incompetency. For outgroup members, on the other hand, causal attributions are less favourable. When outsiders experience success, we are more likely to attribute it to luck or to situational causes rather than to any positive character traits and we are more likely to attribute the failures of outgroup members to innate character flaws rather than to external causes. Affinity bias is also linked to a tendency to withhold praise or rewards from outgroup members and a swifter condemnation of any outgroup behaviours that breach social codes whereas we exhibit greater tolerance of ingroup deviance and failures.
- **Stereotypes:** Stereotypes refer to beliefs that certain attributes, characteristics, and behaviours are typical of members of a particular group of people. The way we categorise social groups is often based on visible features that provide the largest between-group differentiation and least within-group variation (for example, skin colour, gender, age). We construct stereotypes from direct personal experience or, more commonly, from other people, or via the media. The media has a large influence on stereotype formation when we have limited opportunities for meaningful exchange with people from outside our own social group. Stereotypes [influence our assessments](#) of others in ways in which we are unaware and likely would deny. Stereotyped assumptions [can occur intragroup](#) as well as intergroup. Thus the gender of the assessor does not necessarily improve appraisal objectivity. Women often remark that their harshest critics are other women.
- **The ‘halo’ effect:** The halo effect describes the tendency for a single attribute or the assessors general and overall impression of the employee to form the basis of that individual’s overall performance rating, rather than examining performance against multiple indicators. The halo effect can lead to unfair assessments by failing to consider an employee’s strengths and weaknesses fairly and can also cause assessors to overlook specific areas for development.
- **The ‘horns effect’:** This is the opposite of the halo effect and occurs when the manager’s overall impression of the employee is negative, based on a single piece of evidence or no evidence. The overall negative impression drives the assessment rather than giving a fair and balanced consideration to the employee’s performance against multiple indicators.
- **Confirmation bias:** Confirmation bias is our natural tendency to search for and more readily recall information that confirms our preconceived beliefs about people and to discount or actively reject information that is contrary to our pre-conceived judgements of people.

Confirmation bias compounds the impact of stereotypes and affinity, halo, and horns bias on our assessments of others.

- **Distance bias:** The tendency to believe that people and events that are nearer to us in space are more important than people or events further away. This bias leads to unequal weightings across locations. Employees that are located in the same office as the assessor, for example, receive relatively greater attention than the performance of geographically dispersed employees.
- **Recency bias:** Similar to distance bias, our memory favours the recall of recent performance information over performance at the beginning of an assessment period.
- **Spill-over bias:** Similar to confirmation bias, the tendency for a rater's past assessment of an employee to cloud future assessments. Expectations create blind spots that limit accurate assessments in the present.
- **Expediency bias:** The tendency to rely on information that is most readily available, often at the expense of more valuable or relevant information. For example, a manager's reliance on sales figures rather than focusing on customer satisfaction, which may be a better measure of future sales.
- **Central tendency bias:** The tendency of raters to score most of their employees in the middle of a rating scale. Some raters might also tend to rate employees at the higher end of a rating scale. A central tendency or higher rating bias might be linked to an aversion to giving negative feedback.
- **Self-rater bias:** In an effort to increase perceptions of fairness and the engagement of employees in the appraisal process, many employers ask their employees to rate their own performance and to discuss their self-assessment with their manager. However, similar to manager assessments of performance, subjective self-assessments allow for the possibility of bias influencing ratings. Subjective self-assessments of performance can be higher or lower than objective measures of performance and can have an anchoring effect on manager assessments. There are also [cultural differences in self-ratings](#) related to variations across cultures in cultural norms regarding self-promotion. Underrating one's performance is more common among members of cultural groups that value humility such as collectivist cultures. Consider the Western proverb, the '[squeaky wheel gets the grease](#)' and compare with the Japanese proverb, '[the nail that stands up gets hammered down](#)'. In many Asian, Middle Eastern and African cultures, self-promotion is frowned upon. Other individuals have a strong personal aversion to self-promotion due to their socialisation or personality. Introverted personality types are more self-aware and introspective and may rate themselves lower than extroverts. Individuals high in perfectionism or a fixed mindset orientation might rate themselves at lower scores than individuals with a lower propensity for perfectionism or a growth mindset. Members of traditionally underrepresented or marginalised groups can also come to internalise bias and stereotypes and this can lead to lowered self-assessments of performance. Studies repeatedly show that women are more likely to underrate on self-assessments while men are more likely to overrate. Also, whereas women tend to credit their achievements to the efforts of the others such as their workgroup or to luck and their failures to intrinsic flaws, men tend to credit their achievements to their intrinsic strengths and failures to external circumstances.
- **Stereotype threat:** Stereotype threat is the risk of conforming to a negative stereotype and can have negative implications for the performance of individuals from marginalised groups.
- **Institutionalised bias:** Bias can be hardwired into the performance appraisal process in the form of competency criteria or capabilities that favour some groups (usually the majority group) over

others. For example, capability frameworks that emphasise stereotypically masculine qualities like individual initiative and strong work-ethic (a criteria linked to long-hours) are biased against women who have traditionally been the primary caregiver, are more likely to work flexibly, and are assessed through the lens of prescriptive gender stereotypes that favour men over women when making assessments regarding individual initiative. Similarly, capability frameworks that omit stereotypically feminine qualities like collaboration or emotional intelligence can also drive gender bias in assessments. Other biases that are typically hardwired into performance appraisals in Western settings involve capabilities that favour extroverts, members from individualistic (expressive, direct speech and competitive) cultures, people without a physical or mental disability, native-English speakers.

- **Forced rankings and subjective calibration processes:** Although increasingly less common than in the past, compensation schemes can be linked to a calibration process that forces employee rankings into a bell curve. Forced rankings have fallen out of favour largely because [this process has been shown to negatively impact performance](#). Ranking employees against each other fosters competition and animosity within teams. Rather than working together, individuals compete against each other to receive a higher score. Calibration processes are an important part of forced rankings but are also used independently of forced rankings for the purposes of assigning performance ratings to individual employees. Calibration processes typically involve group discussions among senior staff regarding the comparative rankings of different employees. Although calibration processes are assumed to increase fairness by increasing consistency and objectivity in ratings because all manager ratings are subject to an independent and comparative review by a committee, unless calibration processes are structured to focus specifically on bias detection and elimination, these discussions can further open the door to bias. For example, if particular individuals dominate calibration discussions because of rank, personality, or communication style, the process may perpetuate rather than disrupt bias. Also, there are differences across candidates to the extent to which they are known to the calibration members and this can influence relative assessments. “Cronyism” refers to favouritism that a manager gives to employees that they are closest with on a personal level. Further, [researchers observing calibration discussions have noted the extent calibration committees and their members vary](#) regarding the weighting they assign to different criteria and that subjective assessments and gender stereotypes often play into these assessments. When calibration processes allow room for subjective assessments, they can add another layer of bias to assessments, rather than improve objectivity and fairness.

## Solutions for Fair Performance Appraisal

Employers can adopt a number of strategies to improve the objectivity of performance appraisals. Employers that practice these strategies simultaneously stand the best chance of developing, retaining, and promoting top talent and building diverse leadership teams.

**1. Provide unconscious bias training to people managers.** [Unconscious bias training assists](#) managers in understanding their implicit assumptions and prejudgments. It seeks to reduce bias in performance appraisal by transferring skills for objective assessment and developing hiring managers' ability to monitor and manage their own and others' biases. Research into [unconscious bias training](#) highlights two important considerations (i) unconscious bias training is necessary, but in itself not sufficient, for eliminating workplace bias and (ii) some unconscious bias training programs are more effective than others. Specifically, unconscious bias training is most effective when it: (i) [incorporates bias awareness, or 'a-ha', activities and](#) (ii) [transfers evidence-based bias reduction and mitigation strategies](#). Incorporating 'a-ha' activities that allow individuals to discover their biases in a non-confrontational manner is more powerful than presenting evidence of bias in employment or laboratory studies. Stereotypes and prejudices are maintained and reinforced by powerful cognitive and motivational biases that act to filter out information that contradicts or challenges our pre-existing beliefs or attitudes. We all see bias vested in others [but rarely see or admit our own biases](#). A-ha activities help participants to see how their subconscious preferences and beliefs drive their responses.

Raising awareness of bias is only the first step in helping managers to manage their biases. The [SPACE2 Model of Mindful Inclusion](#) is a collection of six evidence-based strategies that [activate controlled processing and enable individuals to detect and override their automatic reflexes](#): Slowing Down — being mindful and considered in your responses to others; Perspective Taking — actively imagining the thoughts and feelings of others; Asking Yourself — active self-questioning to challenge your assumptions (see below) Cultural Intelligence— interpreting a person's behaviour through their cultural lens rather than your own; Exemplars — identifying counter-stereotypical individuals; Expand — the formation of diverse friendships. While [prompting individuals to remember the six techniques](#), the SPACE2 acronym reinforces a key message—to manage bias, individuals must create space between their automatic reflexes and their responses.

In addition to the SPACE2 Model, the Centre for Work-Life Law offers a comprehensive [evidence-based list of prompts](#) to assist assessors in identifying and interrupting bias in performance appraisals. Similarly, consultant [Cook Ross](#) provides the following sample prompts:

- What kind of biases have I experienced myself? How has that affected me?
- What part of my own agenda is being served by this decision?
- Does this employee or their situation remind me of someone else? Is that association applicable to this situation?
- Are there differences in work style or approach between me and the person I am evaluating? If so, are they wrong, or just different? Might they yield the same results? Can these differences influence my rating of the employee?
- What do I imagine are this employee's career development aspirations? Is this what I imagine, or what he or she has told me?
- What strategies and tactics can I put in place to engage fully and consciously, putting my filters

aside?

Unconscious bias training can also benefit individuals being assessed by helping them to recognise how self-rater bias might be influencing their self-assessments. Assessing managers should also be alert to self-rater bias, including both over and under-evaluation of performance, and how that can anchor their assessments. They should be on the lookout for self-rater bias and, as required, coach employees on how to rate themselves objectively. The Centre for Work-Life Law provides another useful [resource on writing effective self-reviews](#) that managers can share with all members of their team to encourage objective self-assessments.

**2. Use formal prompts that encourage objectivity on appraisal forms.** Similar to the bias-disrupting prompts listed above that assessors learn during unconscious bias training, employers can include written prompts designed to disrupt bias and improve objectivity directly on appraisal forms: Examples include:

- Did you consider performance throughout the entire period of the appraisal?
- Did you consider your rating in light of the criteria listed?
- Give three specific examples of how the employee demonstrated a particular capability?

Prompts help to ensure different assessors approach each review in a consistent and objective manner.

**3. Use objective, specific and clear evaluation criteria.** Creating a rubric for evaluations is one of the best ways of structuring performance appraisals and driving objectivity and transparency in the appraisal process. A rubric defines the criteria against which the employee's performance will be assessed. The manager and employee then apply evidence from the employee's performance outcomes to assess whether they did or did not meet the expectations specified in the rubric. Well-defined and objective evaluation criteria reduce the likelihood that subjective assessments of performance will drive ratings. Objective measures of performance reference behaviours or outcomes that define mastery of the role. These can include productivity metrics such as the number of sales calls in a specific period of time as well as direct output measures such as sales figures, customer satisfaction scores (internal and external), customer retention. Performance criteria should tie back to evidence that business goals and outcomes have been met rather than broad statements about a person's general effectiveness or how they get along with team members or customers. Care should be taken to ensure evaluation criteria is inclusive and does not favour some groups over others. Evaluation criteria should be tested for gender-neutrality and other biases that favour one group over others. Constraining the open box on traditional appraisal forms reduces also helps managers to better identify relative strengths and development capabilities. This

allows managers an opportunity to provide constructive development guidance and prepare a relevant and effective development plan.

**4. Clearly communicate performance criteria and set development goals at the beginning of the performance period.** Evaluation criteria should be communicated to employees and agreed on ahead of the performance review period. This gives employees the best chance of success, allows the employer and manager to set up metrics for tracking success/performance during the performance period, and also provides employees with an opportunity to include professional development goals. Understanding what your direct reports want from their careers enables managers to find ways to better support their development.

**5. Weight rating criteria appropriately to reflect role requirements.** Where assessors have limited influence over the competency or capability models used for rating performance, employers should allow raters an opportunity to challenge the weighting assigned to a capability relative to the role requirements. For example, if strong interpersonal skills are not a critical requirement for a particular role, then the weighting given to that competency evaluation should be lower than other competencies that are deemed more critical for role mastery and job performance. However, when assessors are given scope to adjust the weightings applied to different criteria, checks should be in place to ensure that the weightings used are consistently applied by different assessors across similar roles. When unmanaged, the personal preferences of leaders as to which capabilities are most important or who they favour can lead to inconsistency and unconscious bias in assessment.

**6. Use specific, well-defined rating scales.** Rating scales, similar to rating criteria, should be specific and clear. Avoid ambiguous and vague terms like 'exceeds expectations', 'meets expectations', etc. and replace them with measurable achievements. For example, a rating of 1 is sales between x and y, a rating of 2 is sales between y and z etc. Similarly, a rating of 5 could mean consistently meets deadlines, 4 meets deadlines most of the time, 3 meets deadlines in roughly half of all assignments etc. As above, for like roles, rating scales must be applied consistently across assessors.

**7. Limit the rating scale.** There is evidence that decreasing rating scales can help to eliminate bias in assessments. [As noted in a recent HBR article](#), in [a study published in the American Sociological Review](#), researchers studied one school of a large, North American university that changed its faculty teaching evaluation system from a 1-10 to a 1-6 scale. In total, the researchers looked at 105,034 student ratings of 369 instructors in 235 courses. The research found that, under the 10-point system, men received significantly higher ratings than women in the most male-dominated fields, but switching to a 6-point scale entirely eliminated the gender gap. To explain the results, the researchers ran an

experiment where they gave 400 students identical transcripts of a lecture, which they were told was given by either a male or female instructor. They then randomly assigned whether students would rate the instructor on a 10-point or 6-point scale and asked students to write down the words that first came to mind when they thought of the instructor's teaching performance. Again, researchers found a large gender gap in ratings under the 10-point system, which disappeared under the 6-point system. They discovered that, when using the 10-point scale, students readily assigned 10s to John Anderson, but they were reluctant to do so for Julie Anderson, instead of giving her 8s and 9s. When analysing the words that students used to describe the instructor's performance, the researchers found that the top score on the 10-point scale evoked images of brilliant, extraordinary performance. They also found that the raters tended to associate that kind of performance with John rather than Julie. The top score on the 6-point scale, in contrast, did not come with such strong performance expectations. To receive a 6/6, it was enough for instructors to be perceived as very good; they didn't need to be seen as brilliant or extraordinary. As a result, though students using the 6-point scale were still more likely to use superlatives to describe John's teaching performance, they were just as willing to assign 6/6 marks to Julie as to John. Under the 6-point the expression of bias was limited, and the gender gap disappeared.

**8. Abandon forced rankings and shift your focus from comparing employees to their peers to assessing an individual's performance over time.** [Researchers have found](#) that temporal comparison evaluations, involving the comparison of an individual employee's current performance with their past performance and evaluating how much employees have (or have not) made progress over time, are considered to be fairer than social comparison evaluations. Specifically, when an employee's current performance was discussed relative to their own past performance, they perceived that the evaluations were more individualised, discerning and accurate and that they had been treated in a more respectful way. In contrast, employees whose performance was compared with another person's performance considered the evaluations to be less accurate and fair. The researchers noted differences in perceived fairness was independent of the favourability of the evaluations.

**9. Involve multiple perspectives.** Asking several people to evaluate individuals engages multiple data points and encourages a broader perspective on performance, both of which act to reduce bias. As an example, 360-degree reviews seek performance feedback from a variety of sources, including supervisors, direct reports, peers, customers, suppliers and other stakeholders. This increases the information available for assessing employee performance, giving you a complete picture of actual performance, while also minimising the impact of any individual rater's bias. Multi-rater reviews also help to engage stakeholders—individuals feel listened to in the process of providing constructive feedback on what they valued in the relationship and how they could be better served. Multi-rater reviews also provide a contextual element for an employees performance and assist managers in identifying development needs across different aspects of the employee's role. When soliciting input from others about an individual's review, beware of the risk that confirmation bias might encourage you to reject or ignore views that are opposite to your own. To temper confirmation bias, actively seek to understand perspectives different to your own, adopt an open mind, and foster a sense of curiosity. For

good practice, multi-rater responses should be weighted by how much exposure the raters have to the person they are rating.

**10. Structure the calibration process.** While employees perceive benefits in calibration processes, they are not completely satisfied with the system of calibration, in part because they perceive favouritism to be an issue. Perhaps not surprisingly, higher-performing employees report higher levels of perceived fairness and satisfaction with the calibration system and less perceived favouritism relative to lower-performing employees. In my work, managers have similarly expressed concerns about the objectivity of the calibration process, noting how subjective assessments of performance often dominate discussions. Structuring the calibration processes can help to reduce the potential for bias to influence outcomes. As a start, employers should formalise a process for identifying and discussing bias throughout the calibration process. This might include acknowledging the potential for bias upfront and ensuring all members of the calibration committee verbally commit to engaging in objective assessments and challenging their own and others' biases. Decision-making processes can also be formalised to mitigate the potential for some voices to override others in the process. For example, everyone should be given a voice in the calibration discussion and reach an agreement on the final rating. Also, contrary or dissenting views should be encouraged and considered fairly. Other guidelines should ensure that comments on performance are supported by objective evidence such as examples of desired behaviours or quantitative measures of performance and productivity. Prompting aids such as those provided above can be used by the group to facilitate objective discussions and to challenge final assessments for bias. Calibration members can be encouraged to write down their biased concerns that arise during discussions and these should be shared and discussed. To encourage the calling out of bias, the bias red flags of individual raters can be collected and shared anonymously. It can also be beneficial to appoint an individual to act as overseer of the calibration process. The overseer's role is to encourage objective discussion and decision-making.

**11. Adjust the frequency of performance reviews.** To avoid recency and distance bias, assessors can either take notes throughout the performance appraisal process to reduce reliance on their memories or they can engage in more frequent appraisals. While taking notes of performance during an appraisal period is helpful for promoting fairness because it helps to fill in memory gaps or correct distortions, increasing the frequency of reviews by implementing real-time feedback systems achieves the same outcomes but has additional benefits for performance development. Real-time feedback supports performance development in a way that retrospective appraisals don't because, perhaps counterintuitively, real-time feedback significantly reduces the time spent on appraising employees allowing managers to focus their efforts instead on performance management and also provides employees with real-time feedback that they can apply immediately to improve their performance rather than waiting for an annual review before concerns are noted and development goals are defined. When Deloitte discovered the organisation was spending close to *2 million hours a year* on completing appraisal forms, holding annual performance reviews meetings, and creating the ratings, it radically reinvented its performance management system to encourage a shift in the investment of time away from

leaders talking about ratings among leaders to leaders talking to their people about their performance and careers. Frequent leader feedback is a defining feature of high performing teams. When leaders engage in frequent, brief conversations with employees on setting expectations, reviewing priorities, providing feedback and coaching, and identifying and engaging the employee immediately in initiatives to address developmental needs, employees are better equipped with the information and support they need to do their best work. The Deloitte system encourages leaders to check in with employees weekly so that the conversation remains focussed on coaching future performance rather than assessing past performance. Microsoft, Gap, Adobe, Cisco, Adidas and Accenture are among several other companies that have restructured their evaluation processes in recent years, replacing annual performance reviews with a system where managers give feedback on a more regular basis. [Deloitte research](#) indicates that the impact of these new performance practices is positive: 90 percent of companies that have redesigned performance management see direct improvements in engagement, 96 per cent say the processes are simpler, and 83 percent say they see the quality of conversations between employees and managers are improving. Further, Deloitte report ninety-one percent of companies that have adopted continuous performance management say that they now have better data for people's decisions, making major progress in removing bias and discretion in promotion and advancement. Because continuous feedback and 360-degree feedback are not mutually exclusive, real-time performance conversations give employers an opportunity to gather upwards performance feedback regarding managers as well as downward feedback on individual contributors at multiple points across an appraisal period. Peer feedback can also be incorporated into real-time performance systems.

**12. Monitor ratings and feedback and actively search for patterns that could indicate bias.** It is important to regularly examine performance ratings and feedback for patterns across diversity dimensions. When patterns in performance ratings or feedback are detected that suggest bias might be influencing assessments, appraisal and feedback systems and procedures should be scrutinised to identify how bias is creeping into the system and corrective action is taken swiftly to address weaknesses. [Performance management software](#) that makes it easy to look at company-wide trends, while maintaining the ability to zoom in on individual employee or individual rater data, is critical for facilitating this kind of analysis.

<https://thehumancapitalhub.com/articles/eliminating-bias-from-performance-appraisals>