

Tools for Organizations

Improving Evaluation Forms & Processes

How to Design Evaluation Forms to Interrupt Implicit Bias

Controlling Prove-It-Again Bias (“PIA bias”)

Groups stereotyped as less competent or not qualified for leadership often have to provide more evidence in order to prove their competence or leadership compared to other groups. Groups that have to “prove it again” (“PIA groups”) commonly include women, African-Americans, Latino/as, Asian-Americans (particularly in leadership roles), and individuals with disabilities.¹

How to design your performance evaluations to interrupt PIA bias:

1. Require people to pre-commit to what’s important.

Ideal scenario: Performance evaluation forms should identify the specific skills and competencies necessary for each specific job, and measure the employee against those competencies. Standards for what indicates a “poor,” “good” or “great” performance on each competency should be articulated, and individuals should be judged according to these standards based on their actual performance. Otherwise, studies show, standards shift depending on who’s being evaluated.

At minimum: Set out some skills or competencies that can be clearly defined. Create separate sections on the performance evaluation form for each, and separate these sections of the form from sections that allow global judgments. This will make it easier to assess whether bias is creeping in.

2. Global judgments without specifics are a petri dish for bias.

Ideal scenario: Performance evaluation forms should not allow global judgments. Instead, they should focus on specific skills and competencies, with the rankings of different raters averaged.

Global judgment: Overall an effective teacher

Specific competencies: Presents organized lectures, Class materials were clear and informative, Provides effective answers to student questions, etc.

At minimum, take two steps:

Step #1: Performance evaluation forms should not allow global judgments (e.g. “He has the whole package”) without specifics to back them up.

Step #2: Separate out global judgments from specific competencies—and analyze for patterns. Are some groups regularly judged to have more potential than others, even if their ratings on specific skills are the same? If so, evaluators should be made aware of this bias.

3. Hold evaluators accountable.

Evaluators should be required to provide factual backup and specific examples to justify any evaluation they make, and should be prepared to explain their reasoning in detail if asked to do so. If rules or requirements are waived for a particular individual, evaluators should be required to explain why. All of these requirements should be clearly messaged throughout the organization.

4. Evaluate performance for a specific time period.

Employees may develop reputations—good or bad—that persist for years. Performance evaluations should consider work done over a specified time period to help assure that some individuals aren't merely skating by because they've succeeded in the past, mistakes don't dog some employees forever, and successes are given appropriate weight.

**Tip:* Create a log/form to help evaluators track performance and relevant metrics over the course of the applicable review period (past year or other evaluation period); remind evaluator of specific skills and competencies being assessed, and provide space to list assignments performed and track specific examples for evidence of positive and negative feedback.

Controlling Tightrope Bias (gender, race/ethnicity, and social class)

High-status jobs are seen as requiring stereotypically masculine qualities, but some groups are not perceived nor expected to have these qualities. For instance, women are expected to be self-effacing and nice—good team players, but jobs may require ambition, self-promotion, and assertiveness. The result is women often find themselves walking a tightrope between being seen as too masculine (and so respected, but not liked) or too feminine (and so liked, but not respected). Women, more so than men, may struggle to strike a balance that allows them to be seen both as competent and as likable.²

Similar tightropes exist for some racial/ethnic groups or for class migrants (professionals who come from nonprofessional backgrounds). Stereotypes of Asians/Asian Americans may prevent them from opportunities to assert themselves. African American men may face concerns about appearing *too* aggressive. Class migrants may struggle at self-promotion because they are very concerned about job security. If these groups demonstrate such stereotypically masculine qualities, they're more likely to experience backlash.

How to design your performance evaluations to interrupt Tightrope bias:

1. Separate style from skill sets.

Tightrope bias means that women and many racial/ethnic groups who behave in masculine ways often are criticized as having personality defects, and called "aggressive," "sharp elbows," "outspoken," "too emotional," "prima donna," "not a team player," "cold and distant," "a real self-promoter," or "mean." The question is whether a broader range of behavior is accepted in men than women or in white people than in African- or Asian-Americans. Are men allowed more latitude to get angry, where an angry woman is criticized as having "really lost it?" Are African Americans penalized for anger in a way white people are not?

To control this bias, create an evaluation form with separate space for two kinds of developmental feedback: one that specifies additional skill sets needed (competencies), and a separate one for personality traits. If evaluators find themselves filling out the personality box again and again only for one group more than another, this may help increase awareness that their evaluations might be affected by bias. This also makes it easy for HR to spot bias.

2. Make norms of self-promotion clear.

Men are expected to self-promote to show they are “men to be reckoned with”—competitive and ambitious. Women, historically underrepresented groups, and class migrants may be reluctant to self-promote if they sense they are expected to be modest, self-effacing, and even grateful employees. If your performance evaluation process includes self-evaluations, you need to send the message that *everyone* is expected to list all their accomplishments and take on new opportunities. Providing specific examples here will help make it transparent. When Google trained women that they were expected to put themselves up for promotions, the differential between men’s and women’s rates of doing so disappeared.

How To Design Evaluation Processes to Interrupt Implicit Bias

Should you abandon the traditional performance evaluation process for real-time feedback?

A growing movement advocates for eliminating traditional performance reviews in favor of real-time feedback. Unfortunately, informal feedback given on the fly is a petri dish for bias.

Does bias creep into your process after evaluations are completed?

Analyze your current evaluation practices to see if your top rated employees are receiving the best outcomes, and if this holds for different groups in your organization. And if one person or group evaluates but another makes decisions about hiring, promotion, or salary, close consultation is needed to ensure a fair translation.

Here are important tools organizations can use to detect bias in performance evaluations:

1. Compare averages.

The most simple and straightforward way is to examine your performance evaluation data. Simply comparing the mean scores of women vs. men, whites vs. historically underrepresented groups, domestic versus international employees will reveal whether some groups receive higher evaluations than others. Of course, some evaluation differences could be due to real differences in performance, and significant differences do not necessarily indicate bias. If relatively few members of some groups are present in your organization, this data may not be reliable. But it's important to keep track of and analyze why evaluations differ across groups if this occurs.

2. Train someone on staff to review performance evaluations to spot bias and eliminate it before evaluations become final.

The worksheet, *"Tools for Managers and HR – How to Interrupt Bias in Performance Evaluations"* will enable you to appoint someone in your organization to review the performance evaluations at your organization to spot bias, and interrupt it.

3. Do a Bias Climate Survey.

Stay tuned for our new Bias Climate Survey, a simple, ten-minute survey you can use to find out whether and how bias is playing out in your organization. Details coming soon on womensleadershipedge.org.

4. Conduct regression analysis.

Sometimes bias can be revealed if you do a regression analysis of your performance evaluations. In shorthand, what this means is taking any "bottom line" judgment you make of an employee (an overall numeric rating, ranking, or evaluative category) and using a statistical technique that offers a deeper and more precise analysis than just looking at averages. For example, one study found that positive comments predicted strong overall rankings in men's but not women's performance evaluations, leading men to be three times more likely to be promoted than women.

5. Conduct text analysis of narrative comments.

In addition to performance ratings, some managers provide narrative comments to employees that either explain evaluations or offer feedback/direction for future performance. These narratives can also be examined to test for group differences: E.g., are women and men receiving the same kind of feedback? Text analysis software such as LIWC (Linguistic Inquiry Word Count) can quickly calculate overall word count and sentence length, use of positive descriptors, references to social versus achievement domains, among many other categories, and can be adapted to scan for other key terms of interest (e.g., references to being “too aggressive” or “too emotional”).

Differences based on gender, racial/ethnic, or other group memberships can be easily examined. If women receive “too aggressive” or “angry” comments more than men, the “tightrope bias” may be operating. Be attuned to criticism of women that sends the message they are expected to be modest, self-effacing team players—rather than go-getters.

The text analysis also provides an opportunity to check whether performance evaluations (numeric ratings) map onto narrative comments, or whether employees may be receiving conflicting performance feedback.

¹ Any group may be a PIA if that group is stereotypically seen as “not fitting” in a work setting, or if members of the group are stereotypically seen as “not having what it takes” to meet the requirements of the position. Think about the expectations you have regarding the type of person who best fits a position, and be aware that this may lead you to ask those who don’t fit to prove it again. Technical names for prove-it-again include cognitive bias, implicit bias, descriptive bias, attribution bias, in-group favoritism, casuistry, confirmation bias, stereotype-expectancy, and recall bias.

² The technical name for Tightrope bias is prescriptive bias. Both men and women hold prescriptive stereotypes about how the “good woman” or the “man to be reckoned with” should behave.