Self-ratings and bias in performance reviews - Harvard Kennedy School

Source: gnews-performance-review | **Published:** October 08, 2025 | **PMS Relevance:** PMS CORE | **Type:** Research Study | **Geography:** Multi-Region

Source URL: https://news.google.com/rss/articles/CBMivgFBVV95cUxPRHZRUWdZeld5ZzNOeURVRE9qU0h

Executive Summary

This article summarises a peer-reviewed field study published in the Journal of Economic Behavior & Organization (2025), co-authored by Iris Bohnet, Oliver P. Hauser, and Ariella S. Kristal of Harvard Kennedy School. The study investigates whether gender and race disparities in performance appraisals stem from employees' self-ratings, managers' evaluations, or the interaction between the two. Leveraging a natural experiment at a multinational financial services firm — a 2016 software glitch that prevented managers from viewing self-evaluations — the researchers analysed four review cycles (2015–2018) to isolate the causal effect of self-rating visibility on final performance scores. Key findings indicate that women, particularly women of colour, consistently assigned themselves lower self-ratings; managers rated people of colour lower regardless of self-rating visibility; and concealing self-ratings reduced anchoring effects but did not eliminate demographic disparities because managers substituted historical ratings as an alternative anchor. Notably, newcomers with no rating history who also had hidden self-ratings showed reduced gaps for women of colour. The article concludes that addressing performance rating bias requires both process-level interventions (withholding self-ratings) and manager-level interventions (structured criteria, bias interrupters, and demographic auditing).

Key Insights

1. Women, and especially women of colour, consistently self-rate lower than men — a pattern attributed to social norms and anticipated backlash against self-promotion rather than actual performance differences.

2. Race-based penalties in manager ratings persisted even when self-evaluations were hidden, indicating that managerial bias operates independently of self-rating anchoring and is not resolved by process design changes alone.

3. When self-ratings were hidden and no historical rating data existed (newcomers), women of colour achieved parity with white women and men — suggesting that anchoring on past scores is a secondary mechanism that perpetuates inequity.

Practical Takeaways

• Withholding self-evaluations from managers until after initial ratings are submitted can reduce anchoring effects, with the strongest benefit observed for employees without prior rating histories.

• Because race gaps in manager ratings persist independently of self-rating visibility, demographic outcome auditing and structured calibration processes represent a distinct and necessary layer of intervention beyond process sequencing changes.

Critical Analysis

Strengths: The natural experiment design — exploiting an unplanned software glitch — provides a rare quasi-experimental setting that allows causal inference beyond what observational data typically permits. The multi-year dataset (four cycles) and the ability to compare newcomers versus established employees add analytical depth. The study is published in a peer-reviewed economics journal, lending credibility. Limitations: The study is conducted within a single multinational financial services firm, which limits generalisability across industries, firm sizes, and national cultures. The article does not report sample sizes, confidence intervals, or effect sizes, making it difficult to assess the magnitude of reported differences. The 'modest gender effect' in manager ratings is described qualitatively, not quantitatively in this summary. The article conflates the Harvard Kennedy School news summary with the research findings, and the policy-oriented framing in the latter sections of the article edges toward advocacy rather than neutral reporting. Biases: The article is published by Harvard Kennedy School, which has institutional incentives to highlight faculty research favourably. The framing emphasises actionable organisational fixes, which may overstate the certainty of causal conclusions from a single-site study.

Organizational Practice

A multinational financial services firm used a performance review process in which employees submitted self-evaluations visible to managers before final ratings were assigned. Between 2015 and 2018, the firm collected demographic data including gender and race as part of its review cycles. In 2016, a software glitch inadvertently prevented managers from viewing employee self-evaluations prior to rating, creating a one-year disruption that functioned as a natural experiment for researchers studying the influence of self-rating visibility on final performance scores.

Strategic Implications

The findings point to two structurally distinct sources of bias in performance management systems — one located in the self-rating stage (influenced by identity-based self-presentation norms) and one in the managerial evaluation stage (reflecting persistent race-based penalties). Process interventions such as sequencing the visibility of self-ratings address the former but leave the latter largely intact. The research implies that organisations relying solely on process redesign to address equity gaps in performance ratings may be addressing only part of the problem, while the more durable race-based managerial bias remains. The newcomer finding further suggests that historical performance data, commonly used to calibrate current ratings, can function as an additional transmission mechanism for prior inequities — a dynamic with implications for how rating histories are used in calibration sessions.

Validation Status: passed
