Testing Alternative Job Evaluation Schemes to Improve Civil Servantsâ€™ Work Performance in China

Researchers:

Guojun He

Alain de Janvry

Elisabeth Sadoulet

Shaoda Wang

Qiong Zhang

Sector(s): Political Economy and Governance

Location: Henan Province and Guangdong Province, China

Sample: 3,785 college graduate civil servants (CGCS) from two provinces in China

Initiative(s): Governance Initiative (GI)

Target group: Civil servants

Outcome of interest: Elected Official Performance

Intervention type: Administrative reform Performance-based pay

AEA RCT registration number: AEARCTR-0002621

Partner organization(s): Co-Innovation Center for Social Governance of Urban and Rural Communities

Employers often rely on subjective performance evaluations by supervisors to gauge the performance of workers–particularly in the public sector, where civil servants' work performance is hard to measure. However, relying on the opinions of local supervisors could cause subordinates to prioritize pleasing their supervisors rather than focusing on productive tasks and work responsibilities. Researchers conducted a randomized evaluation testing the impact of two alternative subjective performance evaluation schemes on subordinates' work performance. Hiding the identity of their evaluator from subordinates improved overall work performance and reduced subordinates' tendency to serve the interests of their evaluator, relative to a scheme in which subordinates knew who their evaluator was.

Policy issue

Employers often rely on subjective performance evaluations by supervisors to gauge the performance of workers. This is particularly ubiquitous in the public sector, where civil servants' work achievements are hard to measure using objective and standardized indicators given the multi-dimensional and vaguely-defined nature of their jobs.

However, relying on the opinions of local supervisors could cause subordinates to get involved in influence activities that favor their supervisors (for example, by running errands and chores for the supervisor) rather than focusing on productive tasks and work responsibilities. Little rigorous evidence exists on the potential consequences of subjective performance evaluations and on ways to potentially mitigate these consequences.

Context of the evaluation

In 2006, the Ministry of Human Resources and Social Security of China launched the "3+1 Supports" program, which hires more than 30,000 college graduates annually to work in rural township governments on two-year contracts with the aim of improving local governance. Admission to the program is highly competitive. Once hired, college graduate civil servants (or CGCS) assume four types of positions: township government clerks focusing on poverty alleviation, township government clerks focusing on agricultural support, teachers in township primary schools, and nurses in township clinics. More than one CGCS can work under the same "work unit", or branch led by the same team of supervisors. Upon finishing their two-year terms, CGCS may get promoted for a highly-sought permanent position as civil servants based on their performance. Their performance evaluation relies solely on their evaluating supervisor's subjective assessment.

Under China's unique dual-leadership governance structure, every government organization has two leaders: a party leader and an administrative leader. While the party leader is often perceived to have an edge in authority, at lower levels of government the division of labor between the two leaders often becomes less clear, and there tends to be substantial overlap in their roles. As such, every CGCS reports to two supervisors who both assign job tasks and provide performance feedback on a regular basis. However, under the status quo, only one of the two supervisors is chosen to be responsible for evaluating the CGCS' performance at the end of the year through an opaque assignment process. The CGCS is explicitly told who their evaluating supervisor is at the start of their job. The CGCS, therefore, knows whose opinion matters the most for their career development.



Photo credit: Drevs | Shutterstock.com

Details of the intervention

Researchers conducted a randomized evaluation testing the impact of two alternative subjective performance evaluation mechanisms on subordinates' work performance. In collaboration with governments of two of the biggest provinces in China, researchers randomly assigned 2,378 work units containing 3,785 CGCS into two groups between September 2017 and June 2018.

1. *Revealed scheme*: In work units containing two-thirds of the CGCS, researchers randomly assigned one of the two supervisors to be responsible for evaluating the performance of the CGCS of their work unit at the end of the year. As in the status quo, researchers informed the CGCS about the identity of their evaluating supervisor at the beginning of the evaluation cycle.
2. *Masked scheme:* In work units containing the remaining one-third of CGCS, researchers also randomly assigned one of the two supervisors to be the evaluator. However, researchers did not inform the CGCS about the identity of their evaluating supervisor until the end of their evaluation cycle.

Under both schemes, supervisors were not told whether they were chosen as the evaluator. In addition, colleagues of the CGCS also did not know who was evaluating the CGCS. The new performance evaluation schemes received high-level endorsement from the government partners; this endorsement ensured that the vast majority of CGCS were well aware of the high stakes involved in the evaluation outcomes under the new schemes.

By leveraging the randomization of CGCS into the "revealed" and "masked" schemes, researchers were able to assess whether evaluator anonymity improved CGCS performance. Researchers used assessments given by CGCS' colleagues (who were not hired through the "3+1 Supports" program) to measure CGCS performance. To complement colleague assessments, researchers used other indicators of performance such as the salary bonuses received by CGCS and whether CGCS received a promotion recommendation from their supervisors and colleagues for a permanent position in their work unit.

Results and policy lessons

Hiding the identity of their evaluator from subordinates–the "masked" scheme–improved overall work performance and reduced subordinates' tendency to serve the interests of their evaluator. In contrast, the "revealed" scheme induced subordinates to engage in evaluator-specific influence activities, which were visible to co-workers.

Influence activities: When researchers revealed the identity of a CGCS' evaluator at the beginning of their evaluation cycle, the evaluating supervisor gave a significantly more positive assessment of the CGCS' performance at the end of their cycle than their non-evaluating counterpart. In these instances, evaluating supervisors gave the CGCS a score 0.31 higher on average than non-evaluating supervisors along an overall 1 (bottom) – 7 (top) ranking of CGCS performance. Conversely, under the masked scheme CGCS received similar assessment scores from both supervisors. Further, researchers found colleagues of CGCS in the revealed scheme group were 5.8 percentage points more likely than colleagues in the masked group to predict that the evaluating supervisors' assessment would be more positive than the non-evaluating supervisor's assessment (a 10 percent increase over an average likelihood of 54 percent of colleagues in the masked group). These results suggest CGCS carried out influence activities when they knew who their evaluating supervisor was, and that these influence activities in turn led to more positive assessments from the evaluating supervisor–an impact which was visible to CGCS' colleagues.

Work performance: Hiding the identity of the evaluator led to better CGCS work performance. Colleague assessment scores for CGCS under the masked scheme were 0.22 points higher than colleague assessment scores for CGCS under the revealed schemes along a 1-7 performance scale (a 4 percent increase over a mean score of 5.13 under the revealed scheme). Researchers estimate this effect on colleague assessment scores is larger than the effect of replacing three-year community college graduates with four-year regular college graduates (an estimated 0.15-point increase in CGCS colleague assessment scores) and larger than the effect of being a member of the Chinese Communist Party (an estimated 0.17-point increase), two characteristics generally associated with higher ability. Evaluator anonymity also significantly increased the number of colleagues who recommended the

CGCS for tenure–from 86.4 percent of colleagues in the revealed group to 89.6 percent of colleagues in the masked group (a 3.2 percent increase). Evaluator anonymity also led to better CGCS performance assessments from non-evaluating supervisors. In addition, CGCS under the masked scheme earned, on average, 2.3 percent higher salary bonuses than those under the revealed scheme, which were usually given to the best performing employees as a reward.

These results suggest that slight refinements of performance evaluation practices within bureaucracies–particularly where high-stakes rewards are linked to the subjective opinions of designated evaluators–can lead to substantial and cost-effective improvements in bureaucratic performance.

Researchers have continued to work with the two provincial governments of this study to determine how to best continue the new performance evaluation practices and to explore new randomized evaluations.