



(RESEARCH ARTICLE)



CONCEPTUAL FRAMEWORK FROM PERFORMANCE MANAGEMENT SYSTEM TO MISSION CONTRIBUTION FRAMEWORK

M. Amr Sadik ^{1, 2, 3, *}

¹ Group Chief Human Resources Officer, CDIC “Katakit” – Syria & Jordan.

² Adjunct Professor, Human Resources & Strategy Management, IPE Management School, Paris.

³ International Educational Ambassador, Enigma Business School, Indonesia.

* Corresponding Author; Email: dramrsadik@gmail.com

ORCID Details

M. Amr Sadik: <https://orcid.org/0009-0007-6051-1928>

World Journal of Advanced Research and Reviews, 2026, 32(01), 847–852

Publication history: Received on 29 August 2026; revised on 10 October 2026; accepted on 12 October 2026

Article DOI: <https://doi.org/10.30574/wjarr.2026.32.1.2588>

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

Performance management systems built on individual ratings deliver little of what they promise: in large multi-rater datasets, rater tendencies explain over half of rating variance and the ratee’s performance about a quarter or less, while the process absorbs large amounts of time. This paper reviews that evidence and the motivation theory (self-determination, goal-setting, relational job design, public service motivation) that explains what ratings leave out. It proposes a Mission Contribution Framework in which work is defined, discussed and recognized by its contribution to a stated mission and to named beneficiaries, with continuous coaching, collective judgement of evidence, and decoupled pay decisions. An eight-step implementation path and the main risks are set out. The framework is a conceptual synthesis; its components have support in adjacent research, but the whole awaits empirical testing, and five studies are proposed.

Keywords: Performance management, Performance appraisal, Mission, Prosocial motivation, Self-determination, Public service motivation

1 INTRODUCTION

This paper argues that organizations should stop asking “how well did each person score?” and start asking “what did each person contribute to the mission, and to whom?” Decades of research suggest that the first question yields noisy data and weak motivation, while the second connects daily work to purpose, which the motivation literature treats as a durable driver of effort.

Dissatisfaction with performance appraisal is long-standing. The 2016 SIOP debate on ratings notes that over 25 years of research found formal system changes had no discernible effect on performance management behavior or on performance itself (Adler et al., 2016, citing DeNisi & Smith, 2014). A major review of 30 years of empirical work likewise finds that basic questions about what performance management achieves, and what a rating is worth, remain

open (Schleicher et al., 2019). Yet the practice is entrenched: in a 2013 Mercer survey cited in the same debate, 89% of companies linked pay decisions to ratings.

The paper makes three moves. First, it summarizes the evidence that the traditional performance management system (PMS) is a weak instrument. Second, it draws on motivation theory (self-determination, goal-setting, prosocial and public service motivation) to explain what a better design must supply. Third, it proposes a **Mission Contribution Framework (MCF)**, sets out an implementation path, and states its limits.

Status of the MCF. The term is this paper's own synthesis, not an established, validated model. Evidence cited for its components comes from adjacent literatures; the framework as a whole has not been tested, and Section 8 lists the studies needed to do so.

2 WHAT TRADITIONAL PERFORMANCE MANAGEMENT WAS BUILT TO DO?

A century of research has centered on appraisal mechanics. DeNisi and Murphy (2017) review 100 years of work across eight areas: scale formats, rating criteria, rater training, reactions to appraisal, the purpose of rating, rating sources, demographic differences, and cognitive processes. They observe that the field has mostly tested ideas proposed elsewhere rather than proposing new models, and they call for more research on performance *management* rather than appraisal alone.

The dominant designs share a logic: set individual targets, measure against them, rate, and link the rating to reward.

- **Annual appraisal and rating.** A single yearly judgement by a manager, used for pay and promotion.
- **Balanced Scorecard.** Kaplan and Norton (1992) proposed measuring beyond finance, across customer, internal process and learning perspectives, to manage strategy implementation.
- **Cascading objectives.** Goals flow down from corporate to individual level. The goal-setting literature supports specific, challenging goals (Locke & Latham, 2002), which gave this design its theoretical base.

Cappelli and Tavis (2016) describe the history as a swing between development and accountability, with the current shift moving back toward development through frequent feedback, driven by business needs rather than HR.

Two features matter for what follows. The PMS is built around **the individual as the unit of measurement**, and it is built around **evaluation as its core act**. The MCF changes both.

3 EVIDENCE THAT THE RATING-CENTERED PMS UNDERDELIVERS

Ratings mostly measure the rater. Scullen, Mount and Goff (2000) analyzed two datasets of managers ($n = 2,350$ and $n = 2,142$), each rated by seven raters. Rater-specific tendencies explained more than half of rating variance, more than twice what the ratee's actual performance explained.

Table 1: Sources of variance in performance ratings

Source of rating variance	Dataset 1	Dataset 2
Rater's idiosyncratic tendencies	62%	53%
Ratee's general and dimension-specific performance	21%	25%
Random error	11%	18%

Source: Scullen et al. (2000). Perspective effects (boss, peer, subordinate) were small and are omitted, so rows do not sum to 100%.

Later work points the same way. A paper on rater source effects reports that idiosyncratic rater effects accounted for 72% of rating variance in an earlier multisource study and an average of 58% across Scullen et al.'s two datasets. A multilevel study of supervisory ratings finds that rater leniency and the rating context also explain significant variance. Deloitte's internal analysis, reported by Buckingham and Goodall (2015), found that actual performance accounted for only 21% of rating variance and that the process consumed close to 2 million hours a year. In Deloitte's survey, 58% of executives believed their approach drove neither engagement nor high performance.

The case against ratings, as summarized by Colquitt, Murphy and Ollander-Krane in the SIOP debate (Adler et al., 2016), includes:

- Disagreement among raters of the same performance;
- No adequate criteria for judging ratings;

- A weak link between ratee performance and the rating received;
- Conflicting purposes (pay, development, promotion) in one number; and
- Inconsistent effects of feedback on later performance.

A caution: the debate is unsettled. The debate's background section argues that no intermittent, activity-driven system will by itself produce better performance management behaviour, so removing ratings is not a cure. Cappelli and Tavis (2016) report firms redesigning around frequent feedback and learning. Gartner survey figures circulating in practitioner blogs could not be traced to primary data here and are not relied on.

4 THEORETICAL FOUNDATIONS: WHAT MOTIVATES CONTRIBUTION

Four bodies of theory explain what a rating-centered PMS leaves out and what a mission-centered design can add.

- **Self-determination theory.** Ryan and Deci (2000) identify three innate psychological needs, competence, autonomy and relatedness, which when satisfied yield greater self-motivation and well-being and when thwarted diminish them. Their research examines social conditions that enhance or undermine intrinsic motivation. A system that mainly surveils and ranks individuals risks thwarting autonomy and relatedness.
- **Goal-setting theory.** Locke and Latham (2002) summarize 35 years of evidence that specific, challenging goals improve performance, with goal commitment and feedback as key conditions. This supports keeping goals in any new design. Its implication for the MCF is that goals should be specific about the contribution sought, not only about output volume. Evidence also suggests that on complex, unfamiliar tasks, learning goals can outperform fixed performance goals.
- **Prosocial motivation and relational job design.** Grant (2007) argues that jobs spark the motivation to make a prosocial difference through their relational architecture: the person's opportunities for impact on, and contact with, beneficiaries. In a field experiment, fundraising callers who met a beneficiary showed 142% more phone time and 171% more money raised a month later than control groups (Grant et al., 2007). Perceived impact mediated the effect in a further experiment. Longitudinal evidence from youth care professional's links changes in job contact and job impact to changes in prosocial motivation (Leiden study, listed in References).
- **Public service motivation.** Perry and Wise (1990) developed the case for motivational bases of public service as an alternative to merit pay. Perry, Hondeghem and Wise (2010) review twenty years of research, and later work concludes that public service motivation, supported by longitudinal and experimental evidence, can be an influential lever in motivational strategies.

Taken together, these literatures converge on one design requirement: people need to see that their work matters, to whom, and to have some say in how they contribute. A rating cannot supply that. A statement and evidence of contribution can.

5 THE MISSION CONTRIBUTION FRAMEWORK

Definition. The MCF is a management system in which each person's work is defined, discussed and recognized by what it contributes to a stated organizational mission and to identifiable beneficiaries, with evidence gathered continuously and judged collectively, rather than by an individual's rating against pre-set targets.

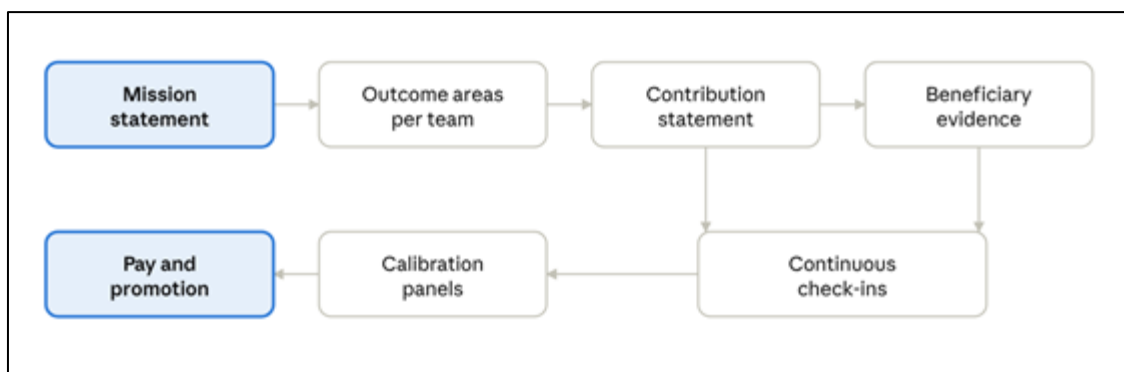


Figure 1: Mission becomes team outcome areas and individual contribution statements; beneficiary evidence and check-ins feed calibration panels, which inform pay and promotion

- **Design principles.** Each follows from the evidence above.
- **Mission line of sight.** Every role links to a mission statement specific enough to guide choices (answers the need for meaning in Grant, 2007, and Perry & Wise, 1990).
- **Contribution over score.** The core artefact is a short contribution statement: what I am trying to change, for whom, and how we will know. Goals stay specific and challenging (Locke & Latham, 2002) but are framed as outcomes for beneficiaries.
- **Beneficiary contact.** Staff see the people and results their work affects, through site visits, user stories and feedback loops (Grant et al., 2007).
- **Continuous conversation.** Frequent coaching check-ins replace the annual event, in line with the shift Cappelli and Tavis (2016) describe and Deloitte's weekly check-ins (Buckingham & Goodall, 2015).
- **Collective judgement.** Decisions use several pieces of evidence and group calibration instead of one manager's score, to reduce the idiosyncratic rater effect (Scullen et al., 2000).
- **Decoupled consequences.** Development conversations are separated from pay and promotion decisions, which draw on accumulated contribution evidence at set points. Deloitte likewise separated compensation decisions from day-to-day performance management.

Table 2: How the MCF differs from a traditional PMS

Dimension	Traditional PMS	Mission Contribution Framework
Core question	How well did you meet your targets?	What did you contribute, and to whom?
Unit of analysis	Individual	Individual within a team and mission
Main artefact	Annual rating form	Contribution statement plus evidence log
Rhythm	Annual, with mid-year check	Continuous check-ins; periodic contribution reviews
Evidence	Manager's rating	Beneficiary feedback, outcomes, peer input
Link to pay	Direct, via the rating	Indirect, via calibrated review of evidence
Motivational basis	Extrinsic: reward and accountability	Autonomy, competence, relatedness, prosocial impact
Main risk	Rater bias, gaming, dread	Vague mission, weak accountability

The last row is deliberate. The MCF trades one set of risks for another, which Section 7 addresses.

6 IMPLEMENTATION ROADMAP

Cappelli and Tavis (2016) advise organizations to find their own rhythm through a cultural discovery process before changing the system, and to settle executive sponsorship, goal alignment, reward design, handling of poor performers, and legal exposure. The sequence below follows that advice.

- **Diagnose.** Count the hours spent on the current PMS, test how much rating variance reflects the rater (as Deloitte did), and survey managers and staff on what the process achieves.
- **Define the mission.** Write a mission statement and three to five outcome areas that teams can translate into their own terms. Test whether staff can say how their work connects to it.
- **Pilot contribution statements.** Start with volunteer teams. Each person drafts a short statement with their manager: change sought, beneficiary, evidence.
- **Install the rhythm.** Weekly or fortnightly check-ins on contribution and obstacles; quarterly contribution snapshots, as in Deloitte's per-project or quarterly snapshots.
- **Build beneficiary loops.** Arrange contact with clients, citizens, patients or internal users, and collect their feedback as evidence.
- **Redesign consequences.** Run calibration panels that review evidence for pay and promotion, with documented criteria so decisions are defensible.
- **Train managers.** Coaching and evidence-based judgement are the critical skills; do not assume managers can award ratings or glean insight from reviews without training.
- **Measure and adjust.** Track engagement, perceived impact, decision consistency and time spent. Compare pilot and non-pilot teams before scaling.

Technology can lighten the evidence log (prompts, feedback capture, dashboards), but it should not turn contribution evidence into another score. Any algorithmic summarization of evidence needs human review and transparency about what is used.

7 RISKS, CRITICISMS AND BOUNDARY CONDITIONS

- **Vague mission.** If the mission is generic, contribution statements become slogans. Mitigation: require observable change and a named beneficiary in every statement.
- **Weak accountability.** Dropping ratings can hide underperformance. The ratings debate (Adler et al., 2016) cautions that no single system change fixes performance management behavior, so manager capability and clear standards still matter.
- **Subjectivity moves, not disappears.** Judging “contribution” is still judgement. Collective calibration and explicit evidence reduce rater idiosyncrasy but do not eliminate bias, and need testing for demographic differences, a long-standing concern in appraisal research (DeNisi & Murphy, 2017).
- **Legal defensibility and pay.** Most firms tie pay to ratings, so decoupling needs documented criteria, consistent decisions and records that withstand challenge.
- **Thwarted impact.** Stressing mission raises expectations. Research on “thwarted prosocial impact” suggests that when staff believe the organization limits the difference they can make, performance and retention can suffer; the MCF must remove obstacles, not only inspire.
- **Context.** Mission is most natural in public, health, education and nonprofit settings, where public service motivation is strong. In commercial firms the mission must be credible, and fit with the culture needs checking first.
- **Evidence base.** The components are supported by separate literatures; the combined framework is untested. Claims here are propositions, not findings.

8 CONCLUSION AND RESEARCH AGENDA

The evidence supports moving the center of gravity from scoring individuals to developing and recognizing contribution: ratings carry more rater than ratee signal, the process is costly, and the motivation literature points to autonomy, impact and purpose as stronger drivers than evaluation. The MCF is a way to organize that shift, with the caveat that it remains a proposal.

Five studies would test it:

- A field experiment comparing pilot teams on the MCF with control teams on a standard PMS, measuring engagement, performance and turnover.
- A variance decomposition of contribution judgements, to see whether collective calibration reduces the rater share found by Scullen et al. (2000).
- A test of whether beneficiary contact raises perceived impact and persistence in non-fundraising roles, extending Grant et al. (2007).
- A study of pay decoupling and its effects on fairness perceptions and legal disputes.
- A boundary-condition study across public, nonprofit and commercial settings.

REFERENCES

- [1] Adler, S., Campion, M., Colquitt, A., Grubb, A., Murphy, K., Ollander-Krane, R., & Pulakos, E. D. (2016). Getting rid of performance ratings: Genius or folly? A debate. *Industrial and Organizational Psychology*, 9(2). Cambridge Core
- [2] Buckingham, M., & Goodall, A. (2015). Reinventing performance management. *Harvard Business Review*, April, 40–50. HBR
- [3] Cappelli, P., & Tavis, A. (2016). The performance management revolution. *Harvard Business Review*, October. Speaker profile citing the article
- [4] DeNisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology*, 102(3), 421–433. doi:10.1037/apl0000085
- [5] DeNisi, A. S., & Smith, C. E. (2014). Cited in Adler et al. (2016); not independently checked.
- [6] Grant, A. M. (2007). Relational job design and the motivation to make a prosocial difference. *Academy of Management Review*, 32(2), 393–417. PDF
- [7] Grant, A. M., Campbell, E. M., Chen, G., Cottone, K., Lapedis, D., & Lee, K. (2007). Impact and the art of motivation maintenance: The effects of contact with beneficiaries on persistence behavior. *Organizational Behavior and Human Decision Processes*, 103(1), 53–67. Abstract page

- [8] Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard: Measures that drive performance. *Harvard Business Review*, January–February. Citation record
- [9] Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), 705–717. PDF.pdf)
- [10] Perry, J. L., & Wise, L. R. (1990). The motivational bases of public service. *Public Administration Review*, 50(3), 367–373.
- [11] Perry, J. L., Hondeghem, A., & Wise, L. R. (2010). Revisiting the motivational bases of public service: Twenty years of research and an agenda for the future. *Public Administration Review*, 70(5), 681–690. doi:10.1111/j.1540-6210.2010.02196.x
- [12] Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. doi:10.1037/0003-066X.55.1.68
- [13] Schleicher, D. J., Baumann, H. M., Sullivan, D. W., & Yim, J. (2019). Evaluating the effectiveness of performance management: A 30-year integrative conceptual review. *Journal of Applied Psychology*, 104(7), 851–887. doi:10.1037/apl0000368
- [14] Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, 85(6), 956–970. doi:10.1037/0021-9010.85.6.956
- [15] Rater source effects are alive and well after all (extract). PDF
- [16] Systematic sources of variance in supervisory job performance ratings: A multilevel analysis of between-rater and between-context variance. North Carolina State University repository. [Link](#)
- [17] Relational job characteristics and prosocial motivation: A longitudinal study of youth care. Leiden University. [Link](#)
- [18] Prosocial impact undermines employee... (thwarted prosocial impact), working paper, Purdue University. [Link](#)