A Performance Appraisal Tool Using Behaviorally Anchored Ratings

Martin CJ Mongiello

John F. Kennedy School of Psychology and Social Sciences, National University

PSY-8407 v1: Industrial/Organizational Psychology (1526540070)

Doctor Lisa Teeter, Ph.D.

September 20, 2025

Introducing Performance Management and Behaviorally Anchored Rating Scales (BARS)

Performance management is a continuous process of identifying, measuring, and developing individual and team performance to align with organizational goals (Aguinis, 2019; DeNisi & Murphy, 2017; Landy & Conte, 2013; Soe et al., 2025). Formal performance appraisal—a core element of performance management—provides documented evaluations, developmental feedback, and data for decisions about pay, promotion, or training (Aguinis, 2019).

Among appraisal formats, Behaviorally Anchored Rating Scales (BARS) are prized for balancing quantitative ratings with qualitative richness. BARS tie numerical scores to specific, observable behaviors, minimizing rater subjectivity (Landy & Conte, 2013).

This paper focuses on specific analysis of an Industrial/Organizational (I/O) psychologist and is constrained to four tasks given to them to perform, only, for the strict focus of a museum foundations needs. It is not a comprehensive, broad-based 27-point review of an I/O psychologist that O*Net Online provides (National Center for O*NET Development, n.d.). "O*NET OnLine is sponsored by the U.S. Department of Labor, and developed by the National Center for O*NET Development" (U.S. Department of Labor, n.d.). O*NET provides detailed, up-to-date information about job tasks, skills, knowledge, and work requirements across hundreds of occupations. Industrial/Organizational psychology, bridging human behavior and workplace systems, provides the theoretical underpinning for evidence-based performance management (Levy, 2017).

Traditional graphic rating scales (Landy & Conte, 2013), can be vague like comments ("good communicator, folks find him funny, she does well with the team"), inviting bias or halo effects (DeNisi & Murphy, 2017; Yusoff, 2025).

BARS counter this by anchoring each scale point with behaviors that exemplify varying performance levels (DeNisi & Murphy, 2017; Landy & Conte, 2013). For complex professional roles—such as an I/O psychologist, BARS clarify expectations for both raters and employees, improve rater agreement, and guide developmental coaching towards furthered excellence, success for all, and happiness (DeNisi & Murphy, 2017; Hauenstein et al., 2010; Kurniawan et al., 2024; Landy & Conte, 2013; Soe et al., 2025).

Typical BARS Development Steps

1. **Job Analysis:** Gather critical incidents, tasks, knowledge, and abilities (often using O*NET, interviews, or surveys).
2. **Identify Performance Dimensions:** Group similar incidents into core task categories.
3. **Generate Behavioral Examples:** Subject Matter Experts or incumbents describe behaviors representing outstanding, excellent, effective, and poor performance.
4. **Scale and Anchor:** Panelists rate the examples' effectiveness, and representative behaviors are chosen for each scale point (e.g., 1–5).
5. **Validate and Revise:** Pilot the scales, gather rater feedback, and refine anchors to ensure clarity and reliability (Aguinis, 2019; Hauenstein et al., 2010; Kurniawan et al., 2024; Landy & Conte, 2013; Levy, 2017; National Center for O*NET Development, n.d.).

This method produces actionable, job-relevant measures—ideal for evaluating I/O psychologists, (Aguinis, 2019; Hauenstein et al., 2010; Landy & Conte, 2013), whose work spans assessment design, organizational diagnostics, and evidence-based interventions. Hiding BARS from employees to spring on them is detrimental to any team (Aguinis, 2019; DeNisi &

Murphy, 2017; Ghani et al., 2024; Longenecker et al., 2014). Providing such year-round allows for self-measurement towards success of individuals and teams.

All humans go to work wanting to "get it right" and most do not look forward to saying, "I can' wait to get it wrong today!" (Aguinis, 2019; Ghani et al., 2024; Longenecker et al., 2014; Penn, 2023; Pfeffer & Veiga, 1999). According to Aguinis (2019), using Behaviorally Anchored Rating Scales as part of a continuous performance management cycle improves fairness and legal defensibility.

Using these methods ensures fair and accurate evaluations across any team, shielding organizations from costly lawsuits, bias claims, and allegations of fraud. According to DeNisi and Murphy (2017), Hauenstein et al. (2010), and Kurniawan et al. (2024), Behaviorally Anchored Rating Scales reduce bias and clarify performance expectations.

Treating an I/O psychologist as a "luxury" can backfire—poor documentation, sloppy reviews, and mishandled firings can drive away top talent and invite repeat lawsuits (Aguinis, 2019; DeNisi & Murphy, 2017; Longenecker et al., 2014; Melhem et al., 2025). Such legal hits can drain even billion-dollar companies to bankruptcy, a harsh lesson in the value of proper performance management (Cascio, 1981; Davidson, 1995; Desai, 2014; Kalev & Dobbin, 2006). Worse yet would be a former employee that appears on the side of the Plaintiff with large lawsuit that states, hypothetically, as an example, "The CEO and leadership of the company repeatedly refused to adhere to acceptable ethics guidelines and standardized BARS use. As a doctor, it was known that such behavior would someday result in a large-scale lawsuit which is why one removes themself from the company and goes to work elsewhere. Complaints in writing, and verbally, were ignored or rebuffed and made employees angry" (Wei et al., 2025; Atkin & Conlon, 1978).

Four Specific, Isolated BARS with Scale Descriptor and Behavioral Anchor

Table 1

Behaviorally Anchored Rating Scale (BARS) for Key I-O Psychology Tasks

Task (O*NET)	1 – Poor	2 – Below Expectations	3 – Meets Expectations	4 – Above Expectations	5 – Exceptional
Develop and implement selection procedures or validation studies	Ignores job analysis; uses unvalidated tools	Conducts minimal analysis; tools show weak validity evidence	Designs legally defensible, job-related tools with acceptable validity	Proactively updates tools with fresh data; trains others on validation methods	Innovates cutting-edge, evidence-based systems adopted organization-wide
Analyze job requirements to inform organizational interventions	Misidentifies key tasks/KSAOs; findings unusable	Provides partial analysis missing critical tasks or abilities	Delivers accurate, comprehensive job analysis linked to interventions	Integrates multi-method data, producing actionable, high-impact recommendations	Anticipates future competency needs; analysis shapes strategic workforce planning
Design and deliver training or development programs	Delivers irrelevant or poorly structured content	Training loosely linked to needs; limited evaluation	Creates structured, needs-based programs with clear objectives and assessments	Tailors content creatively; measures transfer of training effectively	Demonstrates measurable organizational change and ROI; programs become benchmarks
Provide consultation on organizational change strategies	Offers advice without data; resists feedback	Gives generic recommendations with minimal stakeholder input	Provides evidence-based guidance aligned with culture and goals	Builds coalitions, manages resistance, and monitors progress proactively	Leads transformational change; outcomes cited as best practice across the field

Note. Scale anchors: 1 = Poor, 2 = Below Expectations, 3 = Meets Expectations, 4 = Above Expectations, 5 = Exceptional.

Conclusion

This analysis positioned performance management as a strategic, evidence-based framework linking individual, team, and organizational outcomes. Behaviorally Anchored Rating

Scales (BARS) were highlighted as superior (Hauenstein et al., 2010; Kurniawan et al, 2024; U.S. Department of Labor, Employment & Training Administration, n.d.) due to their blend of quantitative evaluation and qualitative specificity. BARS provide explicit, observable anchors that improve inter-rater reliability, legal defensibility, and developmental dialogue (Aguinis, 2019; DeNisi & Murphy, 2017). Behavioral anchors for each level (1–5) give evaluators an evidence-based framework for determining performance (Aguinis, 2019; Hauenstein et al., 2010; Kurniawan et al., 2024; Landy & Conte, 2013; Levy, 2017). However, Atkin and Conlon's (1978) analysis leaves BARS significantly challenged. They contend that BARS are riddled with theoretical flaws: critical incidents chosen as anchors can skew ratings, scaling decisions often distort the meaning of performance levels, and raters' memory and perception biases easily override the supposed objectivity of behavioral examples. Some have emphasized that BARS' complexity cannot eliminate subjectivity or cognitive error (Yusoff, 2025). Dr. William E. Deming's famed 14 Points reject performance reviews because they create fear (Penn, 2023), undermine pride in workmanship, encourage competition over cooperation, and blame individuals for systemic issues rather than improving processes.

Despite these critiques, BARS function as more than appraisal tools: they safeguard organizational integrity, support legal compliance (Bohnet et al., 2025), and foster transparency. Integrated into a continuous performance cycle (Dalal, 2005) and communicated clearly to employees (Barber, 2007), BARS uphold industrial and organizational psychology's core values of rigor, fairness (Huang et al., 2014), and evidence-based decisions (DuVernet et al., 2015).

Institutionalizing BARS strengthens talent management (Islami et al., 2018; Parke et al., 2018) while reducing ethical, legal, and operational risks—evidence of a forward-looking commitment to excellence (Pfeffer & Veiga, 1999) and organizational stewardship.

References

Aguinis, H. (2019). *Performance management* (4th ed.). Chicago Business Press.

Atkin, R. S., & Conlon, E. J. (1978). Behaviorally anchored rating scales: Some theoretical issues. *Academy of Management Review, 3*(1), 119–128. https://doi.org/10.5465/amr.1978.4297058

Barber, A. E. (2007). Recruitment. In S. G. Rogelberg (Ed.), *Encyclopedia of industrial and organizational psychology* (Vol. 2, pp. 667–670). SAGE Publications.

Bohnet, I., Hauser, O. P., & Kristal, A. S. (2025). Can gender and race dynamics in performance appraisals be disrupted? The case of social influence. *Journal of Economic Behavior & Organization, 235,* 1–29. https://doi.org/10.1016/j.jebo.2025.107032

Cascio, W. F. (1981). *Implications of performance appraisal litigation for personnel decisions*. *Personnel Psychology, 34*, 211-226. https://doi.org/10.1111/j.1744-6570.1981.tb00939.x

Dalal, R. S. (2005). A Meta-Analysis of the Relationship Between Organizational Citizenship Behavior and Counterproductive Work Behavior. *Journal of Applied Psychology*, *90*(6), 1241–1255. https://doi.org/10.1037/0021-9010.90.6.1241

Desai, V. (2014). Learning to behave badly: Performance feedback and illegal organizational action. *Industrial and Corporate Change, 23*(5), 1327–1355. https://doi.org/10.1093/icc/dtt043

Davidson, J. E. (1995). The Temptation of Performance Appraisal Abuse in Employment Litigation. *Virginia Law Review*, *81*(6), 1605–1630.

DeNisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology, 102*(3), 421–433. https://doi.org/10.1037/apl0000085

DuVernet, A. M., Dierdorff, E. C., & Wilson, M. A. (2015). Exploring factors that influence work analysis data: A meta-analysis of design choices, purposes, and organizational context. *Journal of Applied Psychology, 100*(5), 1603–1631. https://doi.org/10.1037/a0039084

Ghani, B., Malik, M. A. R., & Memon, K. R. (2024). Effects of performance appraisal on employees' extra-role behaviors and turnover intentions: A parallel mediation model. *Personnel Review, 53*(9), 2413–2441. https://doi.org/10.1108/PR-10-2022-0707

Hauenstein, N. M. A., Brown, R. D., & Sinclair, A. L. (2010). BARS and those mysterious, missing middle anchors. *Journal of Business and Psychology, 25*(4), 663–672. https://doi.org/10.1007/s10869-010-9180-7

Huang, J. L., Zabel, K. L., Ryan, A. M., & Palmer, A. (2014). Personality and adaptive performance at work: A meta-analytic investigation. *Journal of Applied Psychology, 99*(1), 162–179. https://doi.org/10.1037/a0034285

Islami, X., Mulolli, E., & Mustafa, N. (2018). Using management by objectives as a performance appraisal tool for employee satisfaction. *Future Business Journal, 4*(1), 94–108. https://doi.org/10.1016/j.fbj.2018.01.001

Kalev, A., & Dobbin, F. (2006). Enforcement of civil rights law in private workplaces: The effects of compliance reviews and lawsuits over time. *Law & Social Inquiry, 31*(4), 855–903. https://doi.org/10.1111/j.1747-4469.2006.00038.x

Kurniawan, D., Al-Faqih, H., & Wahid, R. L. (2024). Development of a comprehensive performance appraisal instrument using behaviorally anchored rating scales and fuzzy TOPSIS. E3S Web of Conferences, 484, Article 01007. https://doi.org/10.1051/e3sconf/202448401007

Landy, F. J., & Conte, J. M. (2013). *Work in the 21st century: An introduction to industrial and organizational psychology*. Hoboken, NJ: Wiley.

Levy, P. E. (2017). *Industrial/organizational psychology: Understanding the workplace* (5th ed.). Worth Publishers.

Longenecker, C. O., Fink, L. S., & Caldwell, S. (2014). Current U.S. trends in formal performance appraisal: practices and opportunities – Part One. *Industrial & Commercial Training, 46*(6), 321–326. https://doi.org/10.1108/ICT-03-2014-0018

Melhem, M. J., Darwish, T. K., Wood, G., & Seitanidi, M. M. (2025). When performance appraisals fail: Emotion regulation and the direction of organizational routines. *Journal of Management Studies.* Advance online publication. https://doi.org/10.1111/joms.13238

National Center for O*NET Development. (n.d.). Industrial-organizational psychologists — 19-3032.00. O*NET OnLine. Retrieved September 20, 2025, from https://www.onetonline.org/link/details/19-3032.00

Parke, M. R., Weinhardt, J. M., Brodsky, A., Tangirala, S., & DeVoe, S. E. (2018). When daily planning improves employee performance: The importance of planning type, engagement, and interruptions. *Journal of Applied Psychology*, *103*(3), 300–312. https://doi.org/10.1037/apl0000278

Penn, J. D. (2023). Forget fads! Seven principles that support a commitment to continuous improvement. *Assessment Update, 35*(4), 8–9. https://doi.org/10.1002/au.30355

Pfeffer, J., & Veiga, J. F. (1999). Putting people first for organizational success. *Academy of Management Perspectives, 13*(2), 37–48. https://doi.org/10.5465/AME.1999.1899547

Soe, H. S., Uchkempirov, M., Mueed, A. D., & Kulkarni, P. (2025). Multi-document summarization for human resources performance appraisal. In *2025 International*

Conference on Advanced Machine Learning and Data Science (AMLDS) (pp. 130–135). IEEE. https://doi.org/10.1109/AMLDS63918.2025.11159450

U.S. Department of Labor, Employment & Training Administration. (n.d.). *O*NET OnLine*. National Center for O*NET Development. https://www.onetonline.org/

Wei, Q., Sun, J., & Wang, Y. (2025). Inaccurate equals bad? The effect of motivational performance appraisal politics on organizational citizenship behavior. *Journal of Personnel Psychology, 24*(1), 44–53. https://doi.org/10.1027/1866-5888/a000354

Yusoff, A. (2025). The good, the bad and the ineffective: An appraisal of performance appraisals. *Journal of Cognitive Sciences and Human Development, 11,* Article 8481. https://doi.org/10.33736/jcshd.8481.2025