



A mass reenlistment for soldiers with the Division Special Troops Battalion, 3rd Division Sustainment Brigade, and 541st Combat Sustainment Support Battalion is highlighted by a “fireball” detonation on an explosive ordnance disposal range at Camp Buehring, Kuwait, 8 September 2021. The Army has faced recent shortfalls in retention and recruiting, and the author opines this is partially a leadership issue and proposes a modification of its talent evaluation process for officers early in their careers. (Photo by Spc. Elorina Santos, U.S. Army)

The One-Hundred-Year War for Talent

Maj. Jeffrey T. Wilson, U.S. Army

In recent years, the Army has encountered challenges in meeting its recruiting and retention goals.¹ This raises critical questions: Could these challenges stem from our treatment of personnel, or could they be caused by our promotion decision processes? Perhaps they are influenced by factors

beyond our control. Ultimately, the root cause is likely a combination of multiple elements. However, the primary concern is not just identifying these factors but determining actionable steps to address them. While this article does not aim to resolve every issue faced by the Army, it narrows its focus to one significant

area: refining the talent evaluation process to weed out counterproductive leaders between the rank of second lieutenant to major. Later, this article delves into why these specific ranks are selected.

Measuring the intangible result of leadership in the Army can be difficult. While it has some quantitative metrics, it is largely qualitative overall and less definable of what success looks like in the long term. However, one way to quantifiably measure leadership success, or effectiveness, is through echo metrics. By looking at how the American populace views and thereby joins the Army, we can see and measure the effectiveness in our leaders to build cohesive teams. The better they build those teams, the more likely soldiers are to speak positively about their experience and, as a result, build good will with the American populace and increase the likelihood of their joining the military.

There is no one single item or data point that fixes everything. There is no “easy button.” Consider this: when planning to climb a mountain, considerable thought goes into achieving that goal. To climb that mountain, first you must train your body, gather tools, practice, plan your primary and contingency routes, and identify danger areas; then, you have to work your way up. Even the journey up the mountain is a tiered process, such as when to switch tools or when to start

using oxygen. This proposal is just one item to address the issue but not a simplistic “fix it all.”

Multi-Source Assessment and Feedback

In 2024, the Army will celebrate the one hundredth anniversary of using Form 67, the officer evaluation report series for talent evaluation, a practice that began over a century ago. Despite notable advancements since the introduction of the War

Department Adjutant General Office Form 711 in 1922, which evolved into Form 67 two years later, the Army’s approach still requires further refinement.² Unsurprisingly, the evaluation system is not perfect. The challenge of accurately rating such a vast and varied organization as the Army is both complex and substantial. This complexity only increases when attempting to standardize evaluations across diverse roles, locations, and missions. This article explores identified issues and changes in response to reform the evaluation system over the past two decades, culminating in a proposed approach for the next evolution in talent evaluation to prevent counterproductive leaders from continuing to promote and advance in the Army.

Army Regulation (AR) 623-3, *Evaluation Reporting System*, para. 3-9, states, “the senior rater will assess the rated officer’s potential compared to all officers of the same rank.”³ When comparing two officers, specifically regarding leadership, subjectivity plays a part in its evaluation. The Army needs to strike a balance in evaluations mixing objective and subjective elements.

The Army attempted to use the Multi-Source Assessment and Feedback (MSAF) 360 to address part of the subjectivity. In 2008, the Army instituted the MSAF to allow peers, subordinates, and superiors to provide assessment on performance.⁴ The MSAF was the right idea but poorly thought out and executed. First, the rated officer was the one who selected which of their subordinates to evaluate and assess. Human nature is inclined to select people we know and like and is indicative of positive reviews. Second, it required time and access to complete. At the platoon level, many soldiers, especially junior enlisted, do not have regular access to a government-issued computer. Third was the concern of backlash; while it was anonymous, it was unknown if what was written could be identified or tied to a person and used informally against them. Lastly, there was no actionable or tangible result tied to it. Everyone may have understood what it was supposed to do (provide feedback and make adjustments to leadership methods), but no one truly understood what it actually did (what tangible results it led to). And within a few years of its inception, completion of the MSAF was rescinded as a requirement for future evaluations.⁵

However, that is not to say there are not lessons to be learned. The intent of the MSAF was good—it was largely tied to the Battalion Commander Assessment

Maj. Jeffrey T. Wilson, U.S. Army, is pursuing a master’s degree in operations research from the Air Force Institute of Technology in Dayton, Ohio. He holds a BS from Ohio State University. During his career, he served with the 2nd Battalion, 5th Special Forces Group; 3rd Brigade, 101st Airborne Division (Air Assault); 2nd Security Forces Assistance Brigade; and the Soldier Support Institute. He has deployed to Kuwait, Iraq, and Afghanistan.

Program (BCAP) process for lieutenant colonel and above competing for centralized selection list (CSL) positions. In 2019, the Army instituted the BCAP process to assess and evaluate officers as part of placement on the CSL. The CSL relied on evaluations and officer record briefs to select personnel to become battalion commanders, division staff officers in charge, and other key nominative billets.⁶ The CSL, prior to BCAP, did not take in a holistic view of the whole soldier concept, nor did it factor in information from peers or subordinate observations. BCAP does take a holistic view using Leader 360, along with physical and mental evaluations and previous board information to determine if the officer under evaluation is ready for these specially selected positions. Waiting until someone is already a lieutenant colonel is too long to identify counterproductive leaders. There needs to be a tool to identify them earlier in their careers and allow them to adjust before continuing to move up the ranks. The below proposal is based on identifying those traits at the second lieutenant to major levels before even being considered for BCAP, and this information can also be used as part of the overall BCAP process as a continual evaluation.

In “360 Degree Feedback Best Practices and the Army’s MSAF Program,” Col. James Fiscus notes that while the Army differs fundamentally from civilian organizations, it can still benefit from adopting their best practices.⁷ With nearly one million soldiers, including active duty, Army Reserve, and National Guard personnel, the Army’s scale is vastly larger than most civilian entities, posing significant challenges in devising an effective evaluation system applicable across the entire organization. Although a perfect model is unattainable and inherent flaws will persist, this does not preclude the possibility of improvement. Criticisms of the evaluation system are longstanding, yet they also underscore the potential for ongoing refinement.

Fiscus identifies eight key components when instituting new assessment tools to be used by employees; however, personnel answering assessment questions need to understand and believe in the purpose.⁸ Inherently within that is the key messaging of what the intended purposes are: evaluations, assessments, promotions, assignments, etc. As the MSAF 360 did not directly tie into any of those, it became another “check the block” item that needed to be completed,

similar to unit tasks as stated by Leonard Wong and Stephen Gerras in *Lying to Ourselves: Dishonesty in the Army Profession*.⁹ Too many tasks to do and not enough time to do them results in compulsory completion at best. The key is identifying the specific purpose and how a soldier’s assessment can provide tangible results.

An article written by Brennan Randel in 2023 titled “It’s Time to Re-Evaluate the Officer Evaluation System” discusses congressionally mandated changes to the OER [officer evaluation report] system.¹⁰ The law provides a framework to consider how to accomplish this: “(A) increase its effectiveness at accurately evaluating and documenting the performance of officers; (B) provide more useful information to officer promotion boards; and (C) provide more useful feedback regarding evaluated officers.”¹¹

Evolution of the Evaluation System and Limitations

The Army is not a public company with a product for sale; however, if it were, its product would be people. The measure of success of a leader cannot be adequately measured solely by objective metrics, especially as they become more senior in grade. As a new second lieutenant, their effectiveness might be how well they execute a range operation, live-fire exercise, Army Combat Fitness Test, etc., but as they move up in rank, assessing leadership requires a more qualitative rather than quantitative review.

To show the system is capable of change, the table is a brief overview of some of the issues and changes applied to them.¹²

In the current officer evaluation system, three primary ratings are utilized to assess performance: Most Qualified (MQ), Highly Qualified (HQ), and Qualified (Q). The MQ rating is designed to identify the top one-third of officers, illustrating exceptional performance. The HQ rating is intended to recognize officers performing better than the majority, marking them as above average. The Q rating, theoretically, should indicate satisfactory performance, sufficient for retention. However, in practice, a Q rating has come to imply a recommendation against promotion or retention. This issue is exacerbated by the lack of a cap on HQ ratings, leading to a situation where officers deemed adequate for retention are often rated as MQ or HQ. With MQ

Table. Evaluation System Issues and Changes

Issue	Authors	Impact	Change
Subjective Bias	David Tier	Introduction of Evaluation Entry System replacing AKO MyForms	Mandatory block checks for company grade officers
External Evaluators	David Tier	Proposing external evaluators to reduce bias	Concept deemed infeasible due to potential for misunderstanding
Opposing Force (OPFOR)/ Combat Training Center (CTC) Evaluations	David Tier	Suggested using OPFOR and O/C teams during CTC rotations for evaluations, increasing realism but adding logistical challenges	Misunderstanding of the evaluation system
Civilian Education for Officers	Paul Yingling, Scott Maucione	Advocated for encouraging officers to pursue civilian post-graduate degrees to broaden their perspectives and improve strategic capabilities	Congressional guidance addressed
All-or-Nothing Retirement/ Blended Retirement System (BRS)	David Tier	Military retirement system	Introduced BRS to provide a 401k-style plan, addressing the career risk of the all-or-nothing twenty-year retirement system
Toxic vs. Counterproductive Leader	Center for Army Learning	Counterproductive leader	Changed terminology to focus on leadership effectiveness, shifting from "toxic leader" to "counterproductive leader" to better identify and address leadership issues

(Table by author)

ratings extending up to 50 percent, distinguishing the top one-third from those just above the median becomes challenging. Restrictive limits are valuable, yet they are not without flaws. Not everyone can attain the highest evaluation rating, because if everyone is rated the best, then no one truly is.

In the context of the MQ, it is crucial to understand the imposed limitation that only one individual in a group of three can receive the MQ, as the number of MQs awarded must remain below 50 percent. This constraint necessitates a strategic approach in the evaluation process. Consider a scenario with three candidates: one with poor performance, another with average abilities, and a third who demonstrates exceptional skills and qualifications. The challenge arises when the individual of the highest caliber is evaluated first. The core challenge is that the MQ cannot be awarded to the most deserving candidate until the evaluations of the other two individuals are completed, or at least considered. A comprehensive assessment of the Army evaluation system

necessitates a balance between subjective insights and objective data. By scrutinizing both the qualitative and quantitative dimensions, we can begin to formulate a data-centric solution.

Statistics and Bias

Lt. Col. Lee A. Evans and Lt. Col. G. Lee Robinson critically examine the U.S. Army's officer evaluation system in their article "Evaluating Our Evaluations" in the January-February 2020 publication of *Military Review*. They focus on mathematical errors, statistical errors, and cognitive biases inherent in the system in the realm of objective metrics rather than subjective views.¹³ They delve into the implications of these constraints on evaluating a large number of officers, emphasizing the challenges in ensuring fairness and accuracy in performance appraisals. The article also explores the impact of cognitive biases on evaluations, underscoring the complex nature of accurately assessing officer performance and potential.

a) If the potential assessment is consistent with the majority of officers in that rank, the senior rater will place an “X” in the “Highly Qualified” box. If the rated officer’s potential exceeds that of the majority of officers in the senior rater’s population, the senior rater will place an “X” in the “Most Qualified” box. The intent is for the senior rater to use this box to identify the upper third of officers for each rank.

b) In order to maintain a credible profile, the senior rater must have less than 50 percent of the ratings of a rank in the “Most Qualified” top box. Fifty percent or more in the “Most Qualified” ratings will be processed with a “Highly Qualified” HQDA electronically generated label (see DA Pam 623–3); however, it will be charged against the senior rater’s profile as a “Most Qualified” OER if it is unresolved, and a documented senior rater profile misfire will occur. To ensure maximum rating flexibility when rating populations change, or to preclude an “Most Qualified” box check from profiling as a “Highly Qualified” rating, senior raters need to maintain a “cushion” in the number of “Most Qualified” ratings given. Senior raters may consider limiting the use of the “Most Qualified” rating box check to roughly one-third of all ratings for officers of a given rank, but this is not a requirement.

(Figure from Army Regulation 623-3, *Evaluation Reporting System*)

Figure 1. Mandated Evaluation Restrictions

Their article identifies several key issues within the Army’s talent evaluation system similar to issues identified elsewhere within this writing but from an academic standpoint. Part of the issue is based off guidance contained within AR 623-3, *Evaluation Reporting System*, itself. AR 623-3 limits rating to 50 percent for Most Qualified but also recommends keeping a profile at one-third (see figure 1).¹⁴

By having a tool that allows for a variance of roughly 17 percent provides for a larger range of potential error. A reasonably large sample, typically larger than thirty, means any ratings less than that size increase the chances of error, or in contextual terms, there is a 32.9 percent chance that there would be exactly two top one-third officers in a rating pool of five, assuming officers are randomly distributed into ratings pools. Thus, given the current profile constraint of less than 50 percent, raters could only award two “Most Qualified” evaluations to a pool of five officers. Moreover, the rater’s ability to discern the two top one-third performers is affected by cognitive biases. There are roughly ten thousand first lieutenants within the Army. That means a little more than five hundred of those who should receive an MQ will not receive an MQ rating.¹⁵

There is a challenge in objectively evaluating talent across different roles and ranks, with subjective biases often influencing outcomes. The system struggles with balancing objective metrics and subjective assessments, particularly in diverse roles ranging from ground-level soldiers to field grade officers. What is needed is a measure to objectively evaluate counterproductive leaders without losing the subjective aspect of assessing leadership.

Proposal to Counteract Bias and Subjectivity

The MSAF initiative, while no longer in use, laid a foundation in which elements can still be observed through BCAP. The number of officers assess through CSL, and therefore BCAP, is a much smaller cohort than the entirety of second lieutenant to major promotions. However, we can scale down the idea and refocus it as one piece of the evaluation process by adding a singular data point—specifically, have subordinates assess the rated officer.

An objection to the proposed change to the evaluation system might be the fear of leaders pandering to their subordinates for favorable feedback. However, this concern is misplaced; the biggest reason is this is only one small item of consideration for the board. All of the other current metrics stay in place (see figure 2), where the MQ with low enumeration results in high board scores and HQ or lower with poor or no enumeration results in low board scores. What rating an officer receives from their senior rater—enumeration, block checks, and potential—all remain relevant as the primary scoring component to determine who is or is not promoted.

Using the metrics of how a subordinate rated their leader could be as small as changing from a 5 to a 5+ or 5-. If the rated officer was viewed as a productive leader, and their file warranted a 5, then they might move up to a 5+. Alternatively, if they were rated as a counterproductive leader with that same 5 board rating, then they might move down to a 5-. And if they were deemed neither productive or counterproductive, the that same 5 rating would remain a 5.

Board File Messaging

Rating	Message rater is trying to send		What you can expect to see on an evaluation		
			Enumeration		Potential
Most Qualified	6+	Top few	#1, #2	1%	Must select BZ Resident ILE
	6	Superior performer	#3	3%	
	6-	Superior potential	#4	5%	
	5+	Select ahead of peers	#5	10%	Select BZ Promote ahead of peers Resident ILE
	5	Outstanding performer	None	20%	
	5-	Outstanding potential		25%	
Highly Qualified	4+	Select with peers		33%	Promote with peers Satellite ILE
	4	Solid performer		Top 49%	
	4-	Good potential			Bottom 51%
	3+	Select if there is a requirement Average performer/potential			
	3				
	3-				
Qualified	2+	Do not select Weak performer		Bottom 25%	Weak performer Not ready With further experience No ILE at this time
	2				
	2-				
Unqualified	1	Show cause	Explanation of derogatory information	Do not promote Do not send to ILE	

(Figure by author, not official board guidance)

Figure 2. Example Board Guidance

The evaluation system is multifaceted, including MQ, HQ, and Q ratings, with additional metrics for board consideration. Consistent negative feedback over time could indicate leadership issues, suggesting the leader may not merit a high rating. It is essential to communicate effectively with subordinates; failure to do so might reflect poor leadership. Additionally, the tendency to undermine others for personal gain should be a critical factor in identifying detrimental leadership behaviors. And while it is possible a leadership style, while productive, might rub a subordinate the wrong way, it will even out over a measured scale. Moreover, if a leader is not effectively communicating why something is done, that is an indication that something needs to change in their communication style. Stepping on others to make one-self look better needs to be a measurable metric.

The current evaluation system can be compared to its proposed alternative but in reverse. While the senior

rater is responsible for setting organizational goals, that does not always equate to achieving those to receive an MQ. In essence, the way our current system is set up is to not necessarily meet the goals of the organization, but simply to make your boss like you. A rated officer may even be abysmal in their job and as a leader, but they appeal to their boss so well they end up getting that coveted top block simply by being in the boss's good graces, even at the expense of stepping on the soldiers beneath them.

The complexity inherent in this model is twofold: who fills it out and how it is measured. First, it is not going to apply to every position. The initial implementation is a test run. Time is needed to evaluate if it works before full implementation to affect board scores. Time is also needed to evaluate what a positive and negative value equates to.

Whenever the evaluation is being completed, there is a message box asking the following questions: Is the

Promotable Info	
Is the Rated Officer promotable and serving in a position authorized for the next higher grade?	<input type="radio"/> Yes <input checked="" type="radio"/> No
Frocked Info	
Is the Rated Officer frocked to the next higher grade and serving in a position authorized for the rank to which he/she is frocked?	<input type="radio"/> Yes <input type="radio"/> No

(Figure from Evaluation Entry System)

Figure 3. Officer Evaluation Report Input Box

Rated Officer promotable and serving in a position authorized for the next higher grade? Is the Rated Officer frocked to the next higher grade and serving in a position authorized for the rank to which he/she is frocked (see figure 3)?

A third question needs to be added: Is this a key developmental position for the rated officer based off rank and area of concentration? If this is marked yes, then, after the officer evaluation report is completed and signed, before submitting to Human Resources Command, a questionnaire of either the binary or ordinal questions goes out to the soldiers (see figure 4). They mark their answer, and that submits the evaluation. Part of defining that measure would be the minimum number of answers required, which could also vary by position. Company commanders, compared to staff officers, have different numbers of soldiers working for them or numbers of soldiers they interact with throughout a battalion. How many the survey goes out to and how many respond are also two different metrics. It may need to be limited to basic branches as many functional areas work independently, along with the Army Medical Department, Judge Advocate General's Corps, and Army Chaplains Corps as special case scenarios. This is part of the details that needs to be worked out before implementation.

Further analysis is needed to determine the best course of action as to which option to move forward with for use. After a year of scores, it may be an Army-wide number for what equates to a productive or counterproductive leader, or there may be a number by branch as to what a "good" or "bad" score entails.

Not every position has soldiers under them. And some have more than others. Without looking at every single position available, it is worth considering most key developmental positions for basic branch officers are going to have soldiers. Both in the sense of those working for you and those working with you. That means this metric would only be used in key developmental positions.¹⁶

Implementing this modified assessment approach would likely have a minimal impact (regarding cost, effort, and time) on the prevailing methodologies used in personnel evaluations within the force. The foundational tools for this implementation are already in place, though they require updating and refinement. The assessment process utilized for the BCAP, which identifies personnel within the same unit identification code (UIC) as the officer undergoing evaluation for CSL, can serve as a model. This process can be adapted to dispatch a single-question survey to a select group of soldiers within the officer's UIC. Alternatively, the survey could be distributed to all personnel within the UIC, incorporating an additional query: "Did you work for or with this individual?" This approach could be further delineated to exclusively gather feedback from subordinates, or alternatively, to generate two distinct sets of data: one from those who worked under the officer and another from peers who worked alongside them.

The incorporation of this assessment is just one of several factors that require careful consideration, and additional contemplations will be discussed toward the conclusion of this article. Another critical aspect to determine is the timing of these evaluations. Drawing from the MSAF model, which recommended

<p>Option 1a</p> <p>Binary: Would you deploy with this person? Yes/No</p> <p>Option 1b</p> <p>Binary (plus): Would you deploy with this person?</p> <p>Yes, No, No Preference</p>	<p>Option 2</p> <p>Ordinal: Would you deploy with this person?</p> <p>1 – I would fight to deploy with this person</p> <p>2 – I would deploy with this person</p> <p>3 – No preference</p> <p>4 – I would not deploy with this person</p> <p>5 – I would fight to not deploy with this person</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

(Figure by author)

Figure 4. Proposed Options Following an Evaluation to Assist in Identifying Counterproductive Leaders

assessments every three years, it is proposed that more frequent assessments could yield a more comprehensive understanding. For instance, lieutenants, who do not hold key developmental positions as defined in Department of the Army Pamphlet 600-3, *Officer Talent Management*, might benefit from annual reviews. In contrast, for ranks captain and major, this process is more applicable exclusively in key developmental roles.

Lastly, a significant consideration is the accessibility of the gathered data. Limiting access to this information at the division level, analogous to iP-ERMS (Interactive Personnel Electronic Records Management System), warrants examination. After the completion of evaluations, subordinate units could request access to this information, enabling the senior rater to provide informed feedback to the rated officer. This feedback could either affirm the current course of action if evaluations are positive or suggest modifications in response to negative assessments. Furthermore, this data could complement performance metrics, offering a more holistic view of an officer's performance as perceived by their senior rater.

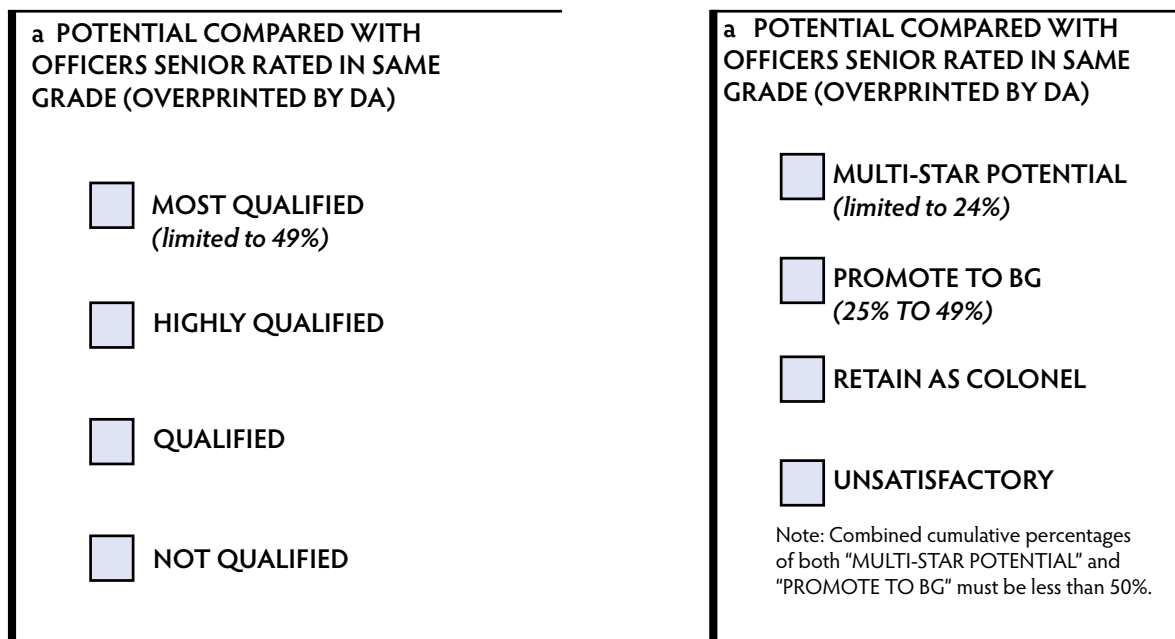
Proposal to Counteract Mathematical Error

The application of MQ ratings varies significantly among senior raters. Some may assign an MQ to an officer they consider in the top 5 percent or 10 percent, while others may use a #3 enumeration for the same rating. This inconsistency results in varied interpretations of an officer's standing within the rating pool. The need for a more precise demarcation among MQ, HQ, and Q ratings is evident, as the current system

allows for disparities in senior raters' interpretations. For instance, one senior rater may grant an MQ rating to an officer they deem in the top 20 percent, whereas another may use similar criteria for an HQ rating.

To address these issues, setting limits on HQ ratings and adjusting the MQ percentage is crucial. MQs should be reserved for officers considered for Below Zone or Early Consideration promotions, signifying superior performance. In contrast, HQs should be seen as indicators of officers suitable for standard promotion timelines. The Q rating, under this proposed structure, would be reserved for officers who meet the basic requirements but are not yet in the running for immediate promotion—a critical signal for improvement, particularly for newer officers such as second lieutenants. By establishing MQs at approximately 24 percent to 30 percent and adjusting HQs to encompass between 50 percent to 60 percent, a clearer understanding of an officer's relative performance within the top, middle, and bottom thirds is achievable. This approach would not only provide clarity for officers receiving their evaluations but also ensure a more objective and transparent assessment process.

As illustrated in the preceding figures, the concept of employing distinct metrics is not a novel practice. At the level of colonel, there exists a delineation between the equivalents of MQ and HQ, in contrast to the singular MQ metric (see figure 5). Implementing such a change in the evaluation system would be a considerable undertaking, necessitating extensive efforts. This would involve not only substantial modifications to the existing system but also securing the endorsement of senior leadership. Additionally, it would require a



(Figure from DA Form 67-10-3, Strategic Grade Plate Officer Evaluation Report)

Figure 5. Excerpt from Strategic Grade Plate Officer Evaluation Report

comprehensive reset of profiles, akin to the initial implementation of the evaluation entry system.

The current evaluation system is commendable for its simplicity. With the limitation of awarding MQ status to less than 50 percent of the candidates, the system allows for straightforward management. Following the first three HQ evaluations, every subsequent assessment can confer an MQ status, facilitating ease of administration and immediate calculation of compliance with established limits. However, transitioning to a system that restricts evaluations to thirds (or similar), although it addresses certain mathematical inaccuracies inherent in the current system, would demand considerably more effort and strategic planning. Such a shift would necessitate significant alterations to the existing system, revisions to regulations, and a thorough communication strategy to inform and guide the entire force. Moreover, this change has the potential to provide subordinates with more clearly defined feedback on their performance, enabling them to make more informed decisions regarding their careers based on this input.

Closing Remarks

The central focus of this article is the imperative evolution of the Army's evaluation system with a

particular emphasis on the identification of counterproductive leaders. While striving for continual improvement, two key recommendations are proposed for system enhancement: (1) the introduction of a mechanism to specifically identify counterproductive leaders through feedback from subordinates, and (2) adjusting the limitations on HQ and MQ ratings.

The first recommendation is pivotal and feasible in the short term. It involves a modest modification to the Evaluation Entry System, providing a crucial data point for boards to identify leaders who negatively impact their units. This focus on counterproductive leadership is crucial for maintaining the integrity and effectiveness of our forces.

The second recommendation, addressing the MQ and HQ metrics, serves to refine the evaluation process and correct mathematical errors. While important, this change is more complex and long-term in nature. However, it supports the primary goal by contributing to a more holistic and objective assessment of officers.

This approach directly aligns with the congressional mandate, prioritizing the identification and management of counterproductive leadership within the Army. It offers a more precise and effective method

for evaluating and documenting officer performance. These enhancements will not only aid in pinpointing counterproductive leaders but also in tracking performance trends. For the officers themselves, this refined

system will provide vital feedback on their leadership capabilities and areas for improvement, thereby guiding their professional growth and decision-making in their careers. ■

Notes

1. Manuela López Restrepo, "The U.S. Army Is Falling Short of Its Recruitment Goals. She Has a Plan for That," NPR, 5 October 2023, <https://www.npr.org/2023/10/05/1203766333/us-army-military-recruit-pentagon-air-force-navy>.

2. Lee A. Evans and G. Lee Robinson, "Evaluating Our Evaluations: Recognizing and Countering Performance Evaluation Pitfalls," *Military Review* 100, no. 1 (January-February 2020): 89–99, <https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/January-February-2020/Evans-Rob-Evals/>.

3. Army Regulation (AR) 623-3, *Evaluation Reporting System* (Washington, DC: U.S. Government Publishing Office [GPO], 14 June 2019), 43, https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/ARN14342_AR623-3_FINAL.pdf.

4. Anthony Francois Cerella, "Multi-Source Feedback in the U.S. Army: An Improved Assessment" (PhD diss., University of Southern California, December 2020), <https://www.proquest.com/openview/9750c0d15d8a27a61c58f139481f7bf5/1?pq-orig-site=gscholar&cbl=18750&diss=y>.

5. James M. Fiscus, "360 Degree Feedback Best Practices and the Army's MSAF Program" (Carlisle, PA: U.S. Army War College, 4 April 2011), 13, <https://apps.dtic.mil/sti/pdfs/ADA559989.pdf>.

6. Joseph P. McGee, *Preparation Guide for BCAP and CCAP* (Fort Knox, KY: Army Talent Management Task Force, August 2020), 2, <https://talent.army.mil/wp-content/uploads/2020/08/CAP-Preparation-Guide.pdf>.

7. Fiscus, "360 Degree Feedback Best Practices and the Army's MSAF Program," 3.

8. *Ibid.*, 5–10.

9. Leonard Wong and Stephen J. Gerras, *Lying to Ourselves: Dishonesty in the Army Profession* (Carlisle, PA: U.S. Army War College, 1 February 2015), 12, <https://press.armywarcollege.edu/cgi/viewcontent.cgi?article=1465&context=monographs>.

10. Brennan Randel, "It's Time to Re-Evaluate the Officer Evaluation System," *War on the Rocks*, 13 April 2023, <https://warontherocks.com/2023/04/its-time-to-re-evaluate-the-officer-evaluation-system/>.

11. James M. Inhofe National Defense Authorization Act for Fiscal Year 2023, Pub. L. No. 117-263, § 509C, 136 Stat. 2562 (2022), <https://www.congress.gov/117/plaws/publ263/PLAW-117publ263.pdf>. See § 509C for the U.S. Government Accountability Office review of certain officer performance evaluations.

12. Todd C. Lopez, "New Army OER Means Fewer Boxes, More Accountability for Raters," *Army News Service*, 29 March 2013, <https://www.jble.af.mil/News/Article-Display/Article/257909/new-army-oer-means-fewer-boxes-more-accountability-for-raters/>; David J. Tier, "Loss of Confidence: The Failure of the Army's Officer Evaluation and Promotion System and How to Fix It," *Small Wars Journal*, 30 August 2015, <https://smallwarsjournal.com/jml/art/loss-of-confidence-the-failure-of-the-army-s-officer-evaluation-and-promotion-system-and-ho>; Paul Yingling, "A Failure in Generalship," *Armed Forces Journal*, 1 May 2007, <http://armedforcesjournal.com/a-failure-in-generalship/>; Scott Maucione, "How Army's Archaic Evaluation System Is Hurting the Service," *Federal News Network*, 1 November 2016, <https://federalnewsnetwork.com/army/2016/11/army-archaic-evaluation-system-hurting-service/>; "Counterproductive Leadership," *Center for Army Leadership*, 7 April 2023, <https://cal.army.mil/Developing-Leaders/counterproductive-leadership/>.

13. Evans and Robinson, "Evaluating Our Evaluations."

14. AR 623–3, *Evaluation Reporting System*, 42–43.

15. Evans and Robinson, "Evaluating Our Evaluations."

16. Author's note: There is a separate argument for removing key developmental positions altogether and having branch managers manage which positions must be filled, but that's a separate discussion altogether.