

2025 MajGen Harold W. Chase Prize Essay Contest: Honorable Mention

Promote the Promotion System

Officer talent management for the 21st century

by Maj Jeremy Goldstein

The character of war has fundamentally changed, necessitating bold adjustments to force design and development. *Talent Management 2030* was intended to be the reorientation needed to enhance the force, but several years later, it has still not yielded meaningful change. Marines of Information Age warfare will need to be smarter, more competent, and more versatile. The amount of data that military leaders need to process has increased, while the speed at which they need to make decisions has decreased. It is an undeniable truism that modern warfighting has grown more complex, requiring a more talented cadre of professionals to manage violence in the 21st century. To meet the urgency of this military epoch, the mechanism by which the Service identifies, incentivizes, and elevates talent needs to be smarter as well. The next war will be fought in the future; our manpower system cannot languish in the past.

To guide this transition, seven principles are recommended to underpin a modern talent management system.

Align the incentives of system participants with desired outcomes.

The desired reporting process outcomes are well-written, timely reports with an accurate characterization of the raters' evaluation of a particular Marine. In addition, that Marine must be sufficiently contrasted from their peers to enable meaningful comparison on

>Maj Goldstein is an Operational Planner at 1st Marine Logistics Group. Lest the following be interpreted as the ravings of an aggrieved failure to select, it is noted that he has not been passed for promotion or command selection—at least not yet anyway.

a board. As it stands, none of that is objectively evaluated or incentivized. The only formal control mechanism is a fitness report (FITREP) attribute (evaluations) where the reporting senior (RS) evaluates the evaluation proficiency of their subordinates using subjective criteria that may differ from these intended process outcomes. Furthermore, the extent to which the RS or reviewing officer (RO) has differentiated reports on their profiles is not evaluated, and neither is informed whether their intended communication was interpreted appropriately by the board. Assuming that the rater's goal is to influence the board in a particular direction, there is no feedback mechanism to inform them of how successful they were. In other words, if you systematically wrote ambiguous FITREPs that presented muddled observations to a boardroom, the system is designed so that you never find out.

Ensure the underlying quality of the report data used to make a determination is sound.

The current method for calculating evaluations is highly flawed, and numerous researchers have documented statistical issues that introduce more noise

than signal. For instance, relative values (RV) only tell the intended story under a set of assumptions that are rarely present in the real world.¹ Mismatches between written descriptions and rankings are common. Superlative adjectives like “#1 Officer” only correlate to a corresponding 100 RV around one-third of the time.² Reviewing officers barely make use of all eight blocks, and 70 percent of Marines Reported On (MRO) are lumped into just two of the eight available blocks, barely enough differentiation to tell a concrete story to a board.³ An RS's evaluation of a Marine is occasionally functionally duplicated in a report due to the RO having insufficient personal observation of the MRO but submitting RO comments anyway.⁴ There have been issues with RSs ghost-writing remarks for ROs going back at least two decades.⁵ Lastly, all reports, regardless of duration, the depth and concentration of a rater's profile, and the competency of the rater, carry the same weight in the overall calculation. A four-month report from a terminal captain with a highly concentrated profile of four Marines will carry the same weight as a twelve-month report from a future general officer with a dispersed profile of twenty Marines. Clearly, the



The evaluation and selection processes for promotion are entirely subjective. (Photo by Cpl Emmanuel Ramos.)

data in both reports should not bear the same significance, and yet, today, they do.

Increase the transparency of the process and outcome, pursuant to statutory limitations.

Promotion and selection boards are notoriously opaque, partially for a reason. There are statutory limitations on the information that can be divulged from a promotion board.⁶ Despite that limitation, there is still room to release data helpful to prospective populations beyond the paltry statistics currently released. Furthermore, the same statute does not apply to non-statutory boards, like education and command, and there are fewer restrictions as a result.

Minimize surprise and unpredictability using modern technology.

Despite their availability, generative artificial intelligence and predictive data analytics tools are not used in either the writing of a report, its analysis, or the collection and interpretation of relevant statistics. For instance, very little data that pertains to a board is viewable to the general population despite the predictive value of that data. It has been demonstrated that certain factors are highly predictive of promotion.⁷ Unsurprisingly, according to a study from

2008, “a one-point increase in the Cumulative Relative Value would result in a 1.1 percent increase in promotion to major and an 8.7 percent increase in promotion to lieutenant colonel.”⁸ There are no tools available to a Marine that aggregate these factors into a probabilistic chance of selection, despite the relative ease of creating one with today’s technology.⁹

Maximize consistency by improving objectivity and decreasing luck as an influential factor.

Luck plays a crucial role today, from the depth of an RO profile, the writing chops of the RS, and the average proficiency of the board members to tease out obscure but important data points, to name but a few examples. To take the example of RO profile depth, due to the simple arithmetic average that is used to calculate the above, with, and below statistics for in processing and cumulative RO values on the Master Brief Sheet, one report from an RO with a large profile will overwhelm the data in other reports from ROs with small profiles. There are theoretically legitimate reasons to design a system like this; a larger sample size should yield more trustworthy data. But since ROs are bunching observations anyway, that does not apply here. In addition, it is

likely that the RO with a profile of ten had far more intimate understanding of a Marine’s performance than an RO with a profile of 200. This could be why there is so much clumping in the first place: ROs of large profiles do not have enough opportunity for careful observation and therefore struggle to discriminate among those they write on. Whatever the reason, the result is that large profiles are given disproportionate weight in determining a Marine’s career while smaller profiles are nearly statistically useless.

Furthermore, outside of objectively measured individual training requirements (which count for less than the subjective components of performance), the evaluation and selection processes are highly prone to personal idiosyncrasies and peculiarities. The common refrain that “every board has its own personality” is a tacit recognition that our boards naturally yield systemic inconsistencies due to their makeup. The Service should decide exactly what it is looking for in its commanders and senior officers, and it should design a selection process that maximizes that every time. Outcome predictability is both positive for the individual and the organization.

Implement checks to mitigate human error.

Marines are human and humans are fallible, yet there are few checks to reduce the consequences of that fallibility. For example, absent adverse keywords, the RS cannot be sure that a report’s text corresponds to the intended RV, and as stated above, the text and RV often do not match. For the MRO, without years of interpreting the nuance of fitness reports, one has no way of interpreting the quality of written comments. Who would guess that words like “bold” and “eager” on a captain’s report are correlated with a bottom-tier evaluation?¹⁰ Or that “groom for command” implied a less enthusiastic recommendation?¹¹ In the best case, bright prospects are impacted while raters learn the coded, stylized nuances of the organization’s writing. In the worst case, some raters never master it, and their reports continuously present unhelpful noise to

the boardroom, which they will never be informed of.

Merit should be dominant among the selection criteria and promotion timing.

Currently, it is not. The primary criterion for selection to the next rank and for the most premier billets is seniority, followed by merit. Promotion timelines follow the same pattern. The best performing captain in the entire Marine Corps will be promoted on the exact same timeline, give or take a few months, as the captain in the 50th percentile. That trend will remain consistent throughout the entirety of their respective careers, a woefully deficient incentive structure to energize performance. It also bears the logic of an industrial-era personnel system and is a poor fit for a ruthlessly enterprising 21st-century military outfit.

Recommendations

With the above principles in mind, the process should evolve to meet the talent requirements of today's operational environment. Start with report writing. The use of currently available large language models can evaluate a report's effectiveness before the rater submits. With large data sets, it can identify trends in the textual effectiveness of a report, check it against the calculated RV, and inform the writer if there are inconsistencies that require remedy. This ensures the report being generated is the one the rater intended. It also levels the playing field so those MROs with talented raters are no longer at an advantage relative to others who were unfortunate to be written on by someone who struggles rhetorically.

If the purpose of the RO's evaluation is to contextualize an MRO among a larger sample size, there is no reason why that data is not shaped appropriately. There is currently no oversight or penalty for an RO to place 95 percent of their evaluations in one of the eight blocks available, nullifying the data's use to the board and obviating the purpose of an RO submission in the first place. Simple statistics can help. The standard deviation of an RO's report should be

tracked, and prominently displayed to the rater. The RO can then be guided toward established benchmarks in appropriate profile dispersion. Over time, a model can take the standard deviation of a profile, the writer's skill (based on an AI-driven scoring model), and timeliness in submission (thereby resolving another intractable problem) to calculate the first objective fitness report marking for the RS/RO: evaluations.¹² To deal with the absentee RO problem, ensure the RO verifies they know the MRO. There is still a requirement to take a photo annually. Before being allowed to submit a report, the RO should be required to select the correct Marine from a photo lineup. If the RO cannot, they probably should not be writing that report. Banks have "know your customer" rules. Time to embrace, *know your MRO* rules.

On evaluations, reports should be weighted on a variety of metrics, like length of report, depth of the RS/RO profile, and the caliber of the evaluators.¹³ Smaller RO profiles should be weighted to ensure that the statistical data they contain is not buried by one huge profile. Beyond weighting, a model based on years of historical data and board outcomes can be built to rank the competitiveness of every Marine using factors like at-processing and cumulative RVs, physical fitness scores, awards, and education, among others. The guidance can finally evolve beyond that which currently advises a Marine to "look for an hourglass."¹⁴ When presented to the MRO, the Master Brief Sheet data should contain probabilistic chances of selection and promotion based on historical outcomes. Obviously, the analysis of board members will ultimately carry the determination, and Marines should be taught that probabilities are not certainties but merely aids for decision making.

On boards, far more data should be collated and presented. Board information should be transparent as far as statute permits, which may mean that those on non-statutory boards can learn what determining factors predominantly precluded their selection. When someone's career fails to progress, it is a grave in-

justice to reward a lifetime of dedicated service with silence. The postmortem ritual of boardroom divination has no place in a modern talent management program. No one should have to guess why they failed to muster enough support to clear the selection threshold. The popular argument around opacity is that more transparency would allow the population to game the game, or mold itself to the prioritized criteria. It is unclear why that would be a bad outcome. An institution that clearly articulated what it valued in its senior leaders and a population highly responsive to those standards would be much more efficient. A more cynical reason to withhold information is that it may reveal board outcomes that reflect highly variable criteria, clearly at odds with a well-functioning organization's inclination for predictability and consistency. While this mostly applies to non-statutory boards, an even more ambitious proposal would be for senior leadership to urge congressional action during the next National Defense Authorization Act passage to unwind the strictures on promotion board proceedings. Call it radical to insist that successful organizations thrive on transparency and openness.

Even more radical, though, is the idea that seniority should no longer be the first criterion for selection. Boards should be encouraged to use the full extent of the below zone, up to ten percent of the eligible population. The counter-argument that there is a magic number one must hit in their time-in-grade to be considered "ready for promotion" is belied by the fact that the promotion zones fluctuate constantly and, in the last decade, dramatically. Besides, readily available empirical evidence exists to show that everyone progresses at a different pace. Merit reorder was a step in the right direction, but it is sparingly used, and—confusingly—its results are not published. Why the boards would not want to advertise their premier officers is unclear. It also detracts from the motivational benefit it would have otherwise delivered. Many people are intrinsically motivated by public recognition, part of the reason why awards are presented in formation and not be-

hind a closed office door. The results of merit reorder should be public, and boards should make more aggressive use of the practice. According to *Talent Management Update November 2024*, since fiscal year 2021, only 300 officers were reordered, meaning around just 5 percent of field-grade officers promoted since then, hardly moving the needle (If it was used for promotions outside of field grade, the percentage is even lower.)¹⁵

This proposal would inject much-needed dynamism into the promotion process. At present, every individual has one realistic chance of promotion. The chance of being selected from the below zone approaches zero percent, and the probability of selection from the above zone is orders of magnitude lower once passed. While the Service can make more forceful use of the below zone authorization, to truly liven the personnel system, the up-or-out stipulation for officers would need to be amended at the congressional level. The promotion timeline would become more fluid as a result, moving the strongest performers along quicker while giving those with weaker track records time to mature in grade. No longer would a candidate have just one realistic chance at promotion, but several. Like this, most fully qualified would no longer just mean most fully qualified in a particular year group.

Conclusions

The personnel management system needed to cultivate an elite fighting force capable of winning fast-paced, multidomain warfare is within reach. Generative artificial intelligence can level the playing field among report writers while grading their work, incentivizing more timely and helpful reporting. Large, publicly available data sets can provide probabilistically minded insights to board populations and enable officers to make more informed decisions about their careers. Transparency can improve predictability and consistency among selectees. Vigorous use of the below zone and merit reordering will create a fluid talent system that incentivizes better performance and continuously elevates the most talented cadre to po-

sitions of influence. The expansion of opportunities for fast upward mobility will become a retention tool for the most gifted performers, whose services will be needed in the next conflict.

The prognosis for reform is not encouraging. Despite the clarion call to improve Marine Corps manpower in *Talent Management 2030*, three and a half years from its release, very little about the fitness report or promotion process has changed. By most metrics, it still looks and functions much as it did before that publication's release. The changes outlined above are tantamount to a revolutionary overhaul of the Service's evaluation and selection process, and military organizations do not do well with revolutionary change. Persuasiveness alone cannot change that because the number one impediment to a modern talent management system is not logic, but inertia. As many have written before in these pages, there is a stubborn to our stodgy, slow-moving bureaucracy that resists any argument, no matter how compelling. Force Design, despite whatever its shortcomings may have been, was an incredibly unique period in our Service's institutional history where our leaders moved with a sense of urgency and pushed bureaucratic reticence aside. That era now appears behind us, and with it, probably the last great opportunity to create the talent management system we need for the Service we deserve.

Notes

1. Those assumptions are a mid-range profile average, a meaningful range, and a large sample size. Ryan T. Baker, *Stop Using Relative Values: They Don't Work as Advertised* (Arlington: Center for Naval Analysis, 2024).

2. Philipp E.D. Rigaut, "A Text Analysis of the Marine Corps Fitness Report" (Thesis, Naval Postgraduate School, 2017).

3. Ibid.

4. Adam Clemens, et al., *An Evaluation of the Fitness Report System for Marine Officers* (Arlington: Center for Naval Analysis, 2012).

5. Thomas C. Gillespie, "Let's Stop Ghostwriting Reviewing Officer Remarks," *Marine Corps Gazette* 86, No. 10 (2002).

6. Sec 613a of Title 10 of U.S. Code: "§613a. Nondisclosure of board proceedings. The text reads, "Prohibition on Disclosure.—The proceedings of a selection board convened under section 573, 611, 628, or 20211 of this title may not be disclosed to any person not a member of the board, except as authorized or required to process the report of the board. This prohibition is a statutory exemption from disclosure, as described in section 552(b)(3) of title 5."

7. Joel M. Hoffman, "Significant Factors in Predicting Promotion to Major, Lieutenant Colonel, and Colonel in the United States" (Thesis, Naval Postgraduate School, 2008).

8. Ibid.

9. Joel M. Hoffman, "Enhancing the PES," *Marine Corps Gazette* 92, No. 9 (2008).

10. "A Text Analysis of the Marine Corps Fitness Report."

11. Manpower Management Officer Assignment, "FY2025/CY2024 Lieutenant Colonel Command Screening Board Report" (2024).

12. This is not theoretical technology, but in use today. Standardized exams, like the Graduate Record Examination, are already using computer-tabulated scores to evaluate writing.

13. A limited example of this was previously recommended in these pages. Roy L. Miner, "Improving Selection Boards," *Marine Corps Gazette* 105, No. 7 (2021).

14. This guidance relates to the optimal shape of a Marine's Master Brief Sheet statistics, wherein the RS numbers should be high at the top, whereas the RO numbers should be high at the bottom.

15. *Talent Management Update Nov 2024*.



Sponsored by:
Observer
Media Group Inc.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.