

Reinventing the Leader Selection Process

by Everett Spain

From the Magazine (November–December 2020)



Klawe Rzechy

Summary. The U.S. Army has long struggled with toxic and inept leaders, and no wonder: It has historically chosen battalion commanders, a linchpin position, on the basis of 90-second file reviews. Last year it undertook an ambitious revamping of that selection process, which... [more](#)

Addressing a class of West Point cadets in 2011, Secretary of Defense Robert M. Gates asked bluntly, “How can the army break up the institutional concrete—its bureaucratic rigidity in its assignments and promotion processes—in order to retain, challenge, and inspire its best, brightest, and most-battle-tested

young officers to lead the service in the future?” The question was, he said, “the greatest challenge facing your army—and frankly, my main worry.”



Listen to this article

To hear more, download [Noa app](#) for iPhone or Android.

The secretary’s concern was not ill founded. In a 2009–2010 survey of 22,000 soldiers, 20% said they were serving under a toxic leader. Another survey showed that fewer than 50% of army majors believe the service promotes its best members. (The picture in the corporate world is similarly bleak. In one study, researchers estimated that half of senior executives were failing in their leadership duties. Another found that 16% of managers were toxic and 20% were incompetent.)

In response to such feedback, the army designed an entirely new process for selecting battalion commanders—its first executive-level position, typically attained 17 to 20 years after an officer has joined the service. It chooses approximately 450 a year, each of whom is responsible for the training and development of 500 or so soldiers. Battalion commanders thus have an outsize influence on combat readiness and junior-leader talent retention; they are also the primary source of generals. That’s why Army Chief of Staff James McConville put the overhaul of their selection process at the core of his talent reform efforts.

Over the coming year the first class of officers appointed under the new system will assume their commands. The selection process, which capitalizes on recent and emerging talent-management ideas from both the public and the private sector, includes physical fitness, cognitive, communication, and

psychological tests; peer and subordinate feedback; and interviews rigorously designed to reduce bias. While specifically aimed at improving the validity, reliability, and developmental impact of the army's executive leader choices, it offers important lessons for any organization seeking to bolster its talent assessment and promotion practices.

Transforming an Industrial-Era Process

It's little wonder that the army suffered a crisis of competence in its leadership ranks. Ever since centralizing its officer selection process, in the 1980s, it had chosen battalion commanders by having multiple senior officers simply score each eligible lieutenant colonel's file, which contained subjective performance evaluations, an assignment history, and an official photo. On average, some 1,900 officers would be eligible for consideration each year. Each file review took about 90 seconds; the key text examined in each performance evaluation was shorter than a typical tweet.

Changing course in any large bureaucracy is never easy, of course, and the army faced all the usual obstacles and then some. The dominant laws governing its personnel practices had been written in 1947 and 1980. They directed that several thousand second lieutenants a year be commissioned, brought up to a minimum level of competence, and assigned and developed on the basis of seniority, specialty, and performance. People were managed largely as if they were interchangeable parts—and the system was more or less frozen in place because of its codification in law. But in 2018 Congress passed the John McCain National Defense Authorization Act, which granted the army the flexible personnel authority it had lacked. McConville—then the vice chief of staff—began making plans to improve the quality of the officer corps.

The new process included cognitive, communication, and psychological assessments.

McConville arguably has more HR experience than any previous army chief of staff. Having spent three years as deputy chief of staff for personnel—the service’s lead human resources officer—he has insight into the diverse talent needed in the thousands of army jobs. As a former commander of the 101st Airborne Division, he has learned that every soldier possesses unique skills and that the army’s diversity is increasing. And as the parent of three young army officers, he knows firsthand that generational norms are changing and that Millennials and Gen Zers want more control over their careers.

Consider one of the problems he recognized. Let’s say the army needed to appoint an officer to advise an allied army overseas. Under its legacy system, it would identify candidates with the appropriate seniority (company commander) and specialty (logistics), perhaps reviewing their performance evaluations to make sure they ranked in the top 20% of their peers, and then choose from that pool. But whereas succeeding as a company commander mainly involves directly leading people who are similar to oneself, succeeding as an adviser abroad involves indirectly influencing people who may be quite dissimilar—and doing so in an unfamiliar environment. Simply giving the job to the best company commander would be unlikely to yield the best match. Better results could be obtained by identifying individuals with superior cognitive flexibility, cross-cultural fluency, and interpersonal skills. Moreover, if the army knew which officers enjoyed international travel and meeting people from different cultures, it could choose someone whose talents and preferences were suited to the position, most likely ending up with a high performer who would enjoy and remain in the job.

Recognizing the need for adaptation that scenarios like this presented, McConville set out to transform how the army acquires, develops, employs, and retains its people, beginning with the linchpin role of battalion commander.

Starting from the Ground Up

First, the army redefined talent as the intersection of knowledge, skills, behaviors, and preferences, or KSB-Ps. Next, McConville energized and resourced the Army Talent Management Task Force—a small group of officers charged with prototyping innovative talent-management ideas—directing that inclusiveness should lie at the initiative’s core. (Disclosure: I serve as an external adviser to the task force, and I moderated one of the interview panels in the new selection process.)

The task force researched army leadership doctrine and identified best practices from government, corporate, academic, and nonprofit organizations and allied militaries. It then designed the Battalion Commander Assessment Program, or BCAP: a four-day evaluation of more than 20 KSB-Ps, including communication skill, creativity, ethical leadership, and the ability to develop others. During the first three days candidates would undergo a physical fitness test, writing skill and argumentative essay examinations, cognitive and strategic talent assessments, psychometric tests, and a psychological interview. They would demonstrate their leadership and problem-solving abilities in a team-based outdoor obstacle course, and extensive peer and subordinate evaluations would be reviewed.

The process would culminate on the fourth day with 30-minute interviews in which panels would evaluate candidates’ oral communication skills and decide who was ready for command. Those deemed so would be ranked according to a cumulative score informed by their BCAP assessments along with the rating

assigned after a legacy-style review of their performance file (which the army still considers a valuable part of the selection process). The top 450 or so would be designated for command.

Following two successful prototypes in the summer of 2019, McConville directed a full rollout of the program. During January and February 2020, 750 lieutenant colonels—eligible officers who opted to participate after being recommended on the basis of an old-style file review—gathered for the new assessment process at Fort Knox.

Implementing Strategies to Reduce Bias

The human brain is lazy; we are constantly looking for shortcuts when processing information. Interviewers are no exception. Research has shown that unstructured interviews are often the least-informative part of an assessment. Even experienced interviewers may spend the first 30 seconds of a meeting jumping to a conclusion about the candidate and the rest of the time subconsciously seeking information to confirm that conclusion.

To guard against such shortcuts, the task force designed a full day of familiarization, calibration, and training for the BCAP panelists. Handpicked colonels were trained to serve as moderators to maintain a fair and consistent process. The work was guided by the following principles:

Create diverse panels.

The selection process spanned four weeks, with six panels operating simultaneously each week. Each panel had five voting and three nonvoting members and was assembled for diversity in terms of gender, ethnicity, specialty, and previous assignments. According to army tradition, voting privileges are limited to officers one level or more above the position under consideration; the voting members of each BCAP panel included three one- or two-star generals and two senior colonels, all of whom had been

successful battalion- and brigade-level commanders. The nonvoting members, included to provide additional perspectives, were a command sergeant major with extensive experience advising battalion commanders, a senior operational psychologist, and the moderator.

Conduct in-depth antibias training.

Panel members were taught strategies for preventing the attributional errors that occur most often during job interviews, including *primacy* (a tendency to focus on first impressions), *contrast* (rating candidates against one another instead of against a common standard), *halo/horn* (allowing a single positive or negative trait to overshadow all else), *stereotyping*, and *similar-to-me* biases. The training also emphasized the tendency among leaders to exhibit *blind-spot bias*: recognizing that others may be biased but falsely believing that you are not. Each morning the panelists received a brief antibias refresher before beginning their work.

Don't let panelists evaluate candidates they know.

At the outset, panelists were given the names of the candidates and asked if they had any knowledge of them. This allowed organizers to create panels whose members had no preconceived notions about the people they were evaluating. Panelists were told to recuse themselves if they realized during an interview that they knew the candidate, which happened five times.

Level the playing field.

Interviews can unfairly advantage candidates who have extensive interview experience. During the BCAP prototypes, the task force noted that whereas some lieutenant colonels were excellent interviewees, most were not. So candidates were instructed in the

STAR method, which teaches people to answer questions by describing the situation, the task, the action taken, and the result. Although they were not required to use it, a majority did.

Calibrate grading.

To ensure a single grading standard, panel members were given a rubric for each quality to be assessed that described what was needed to attain each score. Before the panels began their assessments, they met together in practice sessions. First, each panelist independently assessed three mock candidates, and the entire group discussed the results. Members then regrouped in their panels to assess three new mock candidates and go over those results. Each group of mock candidates included one who was strong in the KSB-Ps, one who was moderately strong, and one who was weak.

Use double-blind interviews.

Borrowing a best practice devised by the Boston Symphony Orchestra in 1952, BCAP conducted double-blind interviews, with a black curtain separating the candidates from the panel at all times. This allowed panelists to focus on the content of answers and the KSB-Ps they were assessing rather than form judgments on the basis of ethnicity, attractiveness, or physical symbols such as wings on their uniforms. It minimized attribution biases that might be sparked by candidates' physical presence. And it meant that deep issues could be discussed without fear of repercussions should candidates and panel members work together in the future. The task force also directed candidates not to disclose, and panel members not to ask about, specific jobs they had held or locations where they had worked.



Klawe Rzeczy

Although double-blind panels reduce bias (a test showed that the sergeants major incorrectly identified 50% of BCAP's minority candidates as white), they don't eliminate it. It's usually easy to determine gender, and panelists may consciously or subconsciously try to link pitch, accent, speaking style, or content with a certain demographic. Candidates who learned English as a second language or hailed from the deep South, for example, might have readily discernible accents. So the bias-prevention work stressed the need not to penalize or reward speaking styles or accents.

Tap psychological expertise.

Applying a best practice long used by special operations units, BCAP brought operational psychologists into the process. Each of six senior psychologists supervised several junior colleagues conducting one-on-one interviews with candidates before their day-four interviews with panels. The senior psychologists collected summaries from the junior ones on the candidates seen that day and presented the results to the relevant panels in a standardized format. Because they did not interact with candidates themselves, they could be much more objective in conveying information about them. They also synthesized each candidate's BCAP assessments into a summary of strengths and weaknesses and suggested follow-up questions for the panel to pose.

Design questions for clarity and fairness.

The task force developed a bank of behavior-based questions for each KSB-P being assessed, rotating them in and out to reduce the chances of their being leaked. For instance, a candidate might be instructed to "describe a situation when you advised a subordinate about a significant challenge he or she was having."

In the first segment of each interview, the moderator asked questions from the bank in a set order, thus ensuring that all candidates had the same core experience. He or she then posed any questions the panelists had after reviewing the candidate's performance in the first three days of events and hearing the senior psychologist's summary. Panelists could themselves follow up with questions intended to further illuminate strengths or risks.

Panel members were directed to elicit descriptions of specific situations and the actions taken in response and to avoid hypotheticals such as "would," "could," and "should." For example, instead of asking, "How would you deal with an

underperforming subordinate?” they might say, “Please tell us about a recent time when you developed a subordinate who was underperforming.”

Candidates were required to wait 30 seconds before answering each question—an instruction driven by what psychologists know about certain personality traits. Because extroverts are typically comfortable thinking out loud, whereas introverts tend to process information silently, the waiting period was meant to ensure that the former did not have an unfair advantage.

To further ensure fairness, panelists were instructed not to give feedback or discuss candidates’ answers and to refrain from any body language, such as a thumbs-up or an eye roll, that could signal approval or disapproval to fellow panelists.

Hear from those the candidates would lead.

Borrowing a best practice from Google, which involves an applicant’s potential team members in the interview process, each panel included a command sergeant major—roughly equivalent to a general manager’s senior operations foreman. Those asked to participate had served as advisers to battalion- and brigade-level commanders and general officers and had keen insights about what the job of battalion commander requires. After each interview they shared their insights about the candidate’s strengths and weaknesses in each KSB-P. To minimize recency bias, they were directed not to indicate their overall assessment of the candidate.

Identify and head off aberrant votes.

After the sergeant majors’ comments, panels held nonbinding votes on each KSB-P, with results visible to the moderator alone. If two panelists differed significantly on an assessment, the

moderator asked them to give the reasons for their rating without sharing the actual scores. To avoid having the senior officer in the pair exert undue influence, the junior officer went first.

Make voting confidential.

Next, panels held their official vote. The moderator reminded members to base their ratings on the rubrics and not to identify their votes or discuss the candidates. With their votes panelists submitted comments about candidates' developmental strengths and weaknesses in each KSB-P; those were relayed to the junior psychologists, who conducted a short "out briefing" with each candidate.

Monitor panels in real time.

To ensure consistency and fairness across panels, the general directing the BCAP initiative held daily meetings with the moderators, giving guidance and asking for input on issues, voting trends, and needs. Each day he observed at least one interview per panel via a live closed-circuit camera system. He would occasionally drop into panel rooms where members were wrestling with procedural issues and offer advice. The six moderators, the director, and a panel coordinator communicated regularly on a closed channel, sharing issues, concerns, and best practices in real time. Panelists could ask that the director observe their panel or visit it before or after an interview to clarify procedural concerns; such requests were accommodated rapidly, often within seconds.

Bringing Key Stakeholders into the Process

The organizational change expert John Kotter holds that a crucial step in leading change is building a guiding coalition. BCAP asked for input or participation from several key stakeholder groups: peers and subordinates of the candidates, including the sergeants major, and general officers.

Gather opinions from peers and subordinates.

Prior to the assessments at Fort Knox, BCAP leaders emailed 10-minute surveys to candidates' peer and subordinate officers. The pivotal question: Should the individual be given a battalion command? More than 65% of recipients responded (response rates for army surveys typically fall below 15%). In reviewing the survey results, panelists were reminded that leaders sometimes have to be stern and that they should consider negative feedback in context: If a clear majority of answers about a candidate were positive, negative responses to one or two items should be deemed outliers.

A vast majority of the candidates were recommended for command by a vast majority of their peers and subordinates—suggesting that most lieutenant colonels are leading well, although some are not. Candidates completed the BCAP process regardless of the survey responses, since those were just one of several factors considered.

Bring strategic leaders on board.

The army's current generals rose through the ranks via the old selection process, so careful thought had to be given to obtaining their buy-in. McConville asked the service's 12 four-star generals to weight the assessments used to generate candidates' final scores, thus signaling that senior leadership was behind the program and that everyone else was expected to be too.

As mentioned, three one- or two-star generals sat on each panel. Because the selection process involved 24 panels in all—six panels in each of the four weeks—72 of the army's one- and two-star generals, or more than 20%, took part.

Gauging Early Outcomes and Looking to the Future

The BCAP assessments cost \$2.5 million in travel fees, supplies, equipment, and so on, along with the opportunity cost of participants' time. What did the army gain in return? BCAP's most immediate impact will be on the soldiers led by the 436 newly selected battalion commanders. Remarkably, 150 of the new commanders, or 34%, would not have been chosen on the basis of legacy-style file reviews alone; although their file scores did not place them among the top candidates, their strengths in the BCAP assessments lifted them into that group. Moreover, 25 candidates whose file reviews would have earned them a posting under the old system were deemed "not ready for command" by their interview panels, many because they exhibited strong and consistent evidence of toxicity. Since future generals will be drawn mainly from today's battalion commanders, these results mean that tens of thousands of soldiers (and their families) ultimately stand to benefit from commanders who are more fit, more capable, better communicators, and more thoughtful. (The army generally doesn't publish demographic information about those selected for command.)

The process also generated benefits for the candidates, regardless of whether they were tapped for command. The week at Fort Knox reconnected them with old acquaintances and introduced them to new ones. As we know from network theory and social psychology, strong professional networks increase one's ability to get things done, while strong personal networks boost emotional stability and well-being. And all candidates (even those denied the promotion) were offered follow-on leadership development with a civilian executive coach, to work on findings from the process or on self-identified areas for improvement. A majority signed on, including 64% of male officers and 84% of female ones.

Even seasoned interviewers may instantly jump to a conclusion about a candidate.

In exit surveys 96% of the candidates, including 98% of women and 96% of minority officers, said that BCAP was a better way to select commanders. Two months later, after candidates had learned the results, 97% thought the new program should be continued. Some 11% called for major modifications—such as additional feedback, different evaluation criteria and events, and alternative assessment timelines—that will be analyzed and addressed for the future.

Follow-up surveys and an after-action review revealed an unanticipated benefit: panelists' own development. Although some generals initially questioned why they had to spend valuable time improving the process by which they had been chosen, in the end 95% of the panelists said they believed it was a better way to select battalion commanders. Some were grateful to be refreshed on the issues facing younger leaders. Many reflected on their own leadership behaviors, often commenting that the training made them aware of their biases and the need to lead more inclusively.

The process also provided important information about the panelists. In a few years the army will know which new commanders are successful. Because it recorded all the votes on each candidate, it could identify especially effective evaluators and invite them to serve on other selection boards.

And the initiative's effects extend beyond those who went to Fort Knox. BCAP opened the army's eyes to the possibility of creating a broader culture of evaluation and feedback. Some West Point instructors have adapted the writing rubrics for use in teaching

cadets. At least one army unit is organizing a mock BCAP so that future candidates can increase their fitness and their writing and oral skills. The service is also considering using many of the assessments for the development of officers with four or five years of experience. The evaluations could be repeated several years later, allowing officers to see how they had grown (or not). At both points they could help officers and the army alike optimize assignments and development programs. As officers practice the skills spotlighted in the assessments, their abilities will increase, making for stronger leaders even among those who are never chosen for a battalion command.

Finally, the army has used the BCAP template to design a similar program for selecting brigade-level commanders. And building on BCAP's inclusion efforts, the Talent Management Task Force recently established a formal diversity and inclusion initiative that extends across its various programs.

BCAP has given the army the most carefully vetted class of battalion leaders in its history. Candidates say they gained valuable perspectives and learned much about themselves. Soldiers asked to evaluate peer and superior officers were sent a clear message that their opinions matter and that leaders are expected to treat them with respect. Generals and colonels serving on the panels received a powerful refresher in what junior officers experience in their daily jobs and the skills they need to do them well. Many panelists also underwent the most thorough bias-reduction training they have ever received—which should drive more-inclusive treatment of the people they themselves lead.

Editor's Note: The views expressed herein are those of the author and do not represent the U.S. Military Academy, the Department of the Army, or the Department of Defense.

A version of this article appeared in the November–December 2020 issue of *Harvard Business Review*.

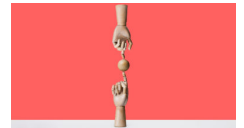
Everett Spain is an active-duty colonel and the head of the department of behavioral sciences and leadership at West Point.

Recommended For You

How to Design a Better Hiring Process



The Psychology Behind Effective Crisis Leadership



Your Approach to Hiring Is All Wrong



PODCAST
Medical Leaves

