



2nd Place, General Douglas MacArthur Military Leadership Writing Competition, CGSC Class 13-02

(U.S. Air Force photo by Tech. Sgt. Jeremy T. Lock)

Lt. Col. Earl M. Hairston, a commander with the 354th Civil Affairs Brigade, Special Functioning Team, Combined Joint Task Force–Horn of Africa, briefs his team 29 March 2008 at Camp Lemonnier, Djibouti, before moving out for a medical civil action project.

Caution Required

Multirater Feedback in the Army

Maj. Gregory G. Lee, U.S. Army

In response to several highly publicized cases of leader misconduct, the Army has made removing toxic leaders from its ranks a priority. Chief of Staff of the Army (CSA) Gen. Raymond T. Odierno states, “We are relieving people, battalion and brigade commanders, for toxic leadership, and

we will continue to do that. The units know, and to me that’s what it’s about. We’re taking action against commanders who are creating environments that are not acceptable.”¹

The Army should screen for toxic leaders—but how? Several books and news articles address the

issue; some propose solutions. Many, such as Tim Kane's *Bleeding Talent*, suggest the solution lies in 360-degree feedback, or multirater feedback, with proposals to incorporate those types of feedback into officer evaluation reports (OERs), promotion boards, and even central selection list (CSL) boards.² Gen. Martin Dempsey, chairman of the Joint Chiefs of Staff, seems to agree. *Army Times* reporter Andrew Tilghman reports that Dempsey said, "as time passes and the force grows more comfortable with 360-degree reviews, they may ultimately be integrated into the command screening process."³ However, Odierno, in an April 2013 address to students at the U.S. Army Command and General Staff College, rightly said the Army must be careful about how it proceeds with implementing multirater feedback into evaluation and leader development processes.⁴

Multirater feedback has been one of the fastest growing assessment instruments in business for leaders, managers, executives, and employees alike, so why not use it in the Army as well? This paper examines multirater feedback, its validity and reliability, appropriate uses for it, and its pitfalls.

An understanding of this assessment instrument—how it performs effectively under certain conditions and how it can be damaging to an organization under other conditions—supports the conclusion that multirater feedback should be used only for development purposes. Using multirater feedback directly for performance evaluations, promotion boards, and CSL boards would lead to improper selection of future leaders and could needlessly damage Army leader development, while failing to address toxic leadership.

What is Multirater Feedback?

Multirater feedback, 360-degree feedback, multisource assessment and feedback, and similar assessment approaches share the common characteristic of providing individuals (leaders, managers, or anyone that interacts with more than one level of the organization) with feedback on their behaviors from the perspectives of others. For the purposes of this article, 360-degree feedback and multirater feedback will be used interchangeably, with the understanding that 360-degree feedback involves

subordinates, and multirater feedback may or may not involve subordinates. Although the number of reports and levels surrounding the individual varies, the concept remains the same. When used properly, this instrument can benefit the target individual and the organization as a whole. From 1982 to 1992, the number of off-the-shelf 360-degree instruments being sold quadrupled. Companies spent \$152 million on multirater feedback systems in 1992, with 90 percent of Fortune 500 companies using a form of multirater feedback in 2003.⁵ Although powerful as an assessment instrument, multirater feedback requires certain conditions: a safe learning environment, experienced coaches or counselors, and a longitudinal development plan.

Development of Multirater Feedback

Over time, organizations and employees have sought more fair appraisal and assessment systems, other than the traditional top-down formal assessments typically prepared by managers about their subordinates. Well-documented cases of unfair, inaccurate, or low-quality top-down reports have demonstrated the need for bottom-up feedback from subordinates. Feedback from multiple raters was intended to counter the subjective nature of top-down ratings, yielding a "fairer and possibly less biased view than simply relying on superior's ratings."⁶ However, according to Clive Fletcher, Caroline Baldry, and Nicole Cunningham-Snell, unless feedback systems are constructed and evaluated along the lines associated with psychometric tests, they may produce misleading assessments.⁷

Multirater feedback instruments request data from the target individual; the individual's supervisors, peers, and subordinates; and if applicable, customers or others. The organization categorizes the behaviors required for performing the job and usually asks respondents to rate the individual's behaviors as observed along a Likert scale, sometimes also including a qualitative portion that allows short-answer input.⁸ The number of observable behavioral competencies varies between instruments, normally tailored to the job type and the organization. The intent of the survey is to present the target individual with a complete picture of his or her behavioral

competencies from multiple perspectives. The respondent data are collated and compared against each category, to include the self-rating by the individual. The comparison demonstrates to the target individual areas where he or she needs to improve, promotes self-awareness, and, with a trained counselor or coach, can identify unexpected discrepancies between the self-report and the respondent reports.⁹

Multirater feedback assessment instruments come in a variety of styles and formats, ranging from web-based surveys with instant output reports to paper-and-pencil surveys scored by hand.¹⁰ Although not recommended, multirater feedback can be used for a one-time developmental assessment. Most commonly, however, the assessments are longitudinal studies conducted over time to allow the target individual the opportunity to demonstrate improved ratings through a developmental program. Feedback can be formatted graphically or numerically, plotted over time, and even accompanied by coaching from trained professionals. The key is that the feedback reaches the target individual, who is allowed time for reflection and self-improvement based on the results. Although some organizations are shifting toward using multirater feedback in performance appraisals, the literature does not support its use without a developmental plan of action. Performance appraisal can be tied into a developmental plan, but an appraisal should not be based on multirater feedback because it has limitations and inaccuracies that are exacerbated by tying them to the appraisal process.¹¹

Psychometric Support—Validity and Reliability

The psychometric support for the various multirater feedback instruments varies greatly depending upon how the instrument is constructed. It is important to assess multirater feedback from the view of validity and reliability. Validity is the degree to which an instrument measures what it claims to measure. If we are using multirater feedback to measure a leader's toxicity, does it do that? If not, the assessment lacks validity.

Reliability, also called consistency, is the extent to which a measurement gives consistent results over time. An instrument with high reliability gives consistent or comparable results over time and

under similar conditions. An instrument that lacks reliability gives inaccurate results. An instrument that lacks validity does not measure what it claims to measure.

The most successful multirater feedback instruments incorporate Likert scales with enough range (no fewer than five choices) to enable respondents to describe the behavior accurately. Too large a range of choices, on the other hand, may reflect trivial distinctions, resulting in less valid data when collated. This is important because respondents must be able to score the behaviors accurately to produce some type of statistical relevancy. Another best practice of multirater feedback instruments is to have a portion that allows for qualitative comments about the rated individual. These are often the most powerful in changing behaviors when used in a developmental forum with a third-party coach.¹²

The validity of multirater feedback tends to be inconsistent. According to D. Theron and G. Roodt, "Blue-collar and service jobs have a higher consistency, based on the fact that these jobs are relatively routine and performance is well defined, compared with managerial and professional jobs with low consistency due to the fact that these jobs are not as easy to define."¹³ Despite an inconsistent record of validity for these types of assessments, effective feedback instruments can be customized for most job types. This requires a thorough job analysis and understanding of the organization's culture (hierarchical organizations will differ from flatter organizations). An administrator of a multirater feedback instrument can design it to create an internal consistency (alpha coefficient of 0.98) when careful attention is paid to the type of questions, number of questions, wording, and scale.¹⁴

Another weakness that can undermine the validity of multirater feedback, according to Caroline Bailey and Clive Fletcher, is "that direct and indirect subordinate [or out-of-department] personnel not of managerial level (whereas targets and their bosses were) ... may not have clear and/or appropriate schema for managerial behavior."¹⁵ In other words, the respondents might have insufficient experience and knowledge of individuals and the job requirements to rate them effectively, reducing the accuracy and reliability of the instrument.¹⁶

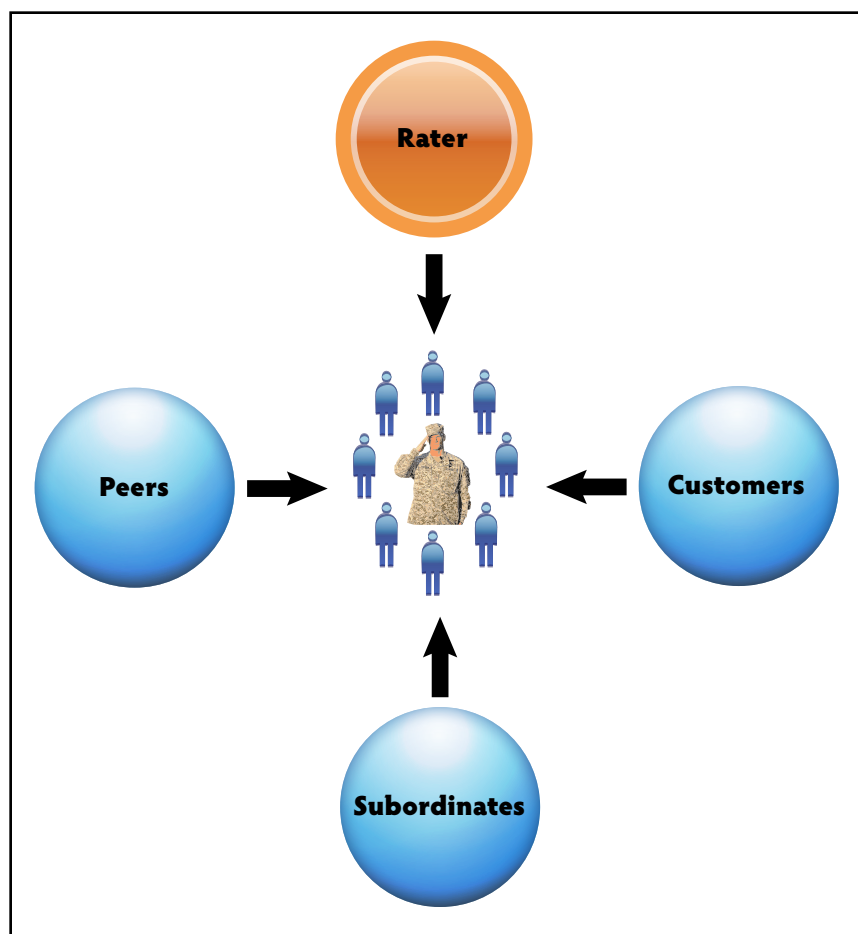
Subjects receiving multirater feedback generally consider the reliability more accurate, fairer, and less biased than traditional assessments due to the numerous sources of the feedback. Statistically, this may not be true because halo effects can skew the data, with raters sometimes scoring the individual as all good or all bad. Increasing reliability in multirater feedback requires reducing rater errors and improving rater agreement. This can be done by clearly defining the rated behaviors with examples in order to establish a common frame of reference for all raters, enabling the comparison of scores across raters. This reduces the ambiguity in the frame of reference and minimizes halo effects.¹⁷

According to Fiona Dent and Judy Curd, the “360[-degree assessment] is completely dependent on collecting data from others about an individual and therefore a highly emotive process that must be handled both in a professional manner and with care.”¹⁸ Some typical emotional reactions by employees receiving feedback are denial, shock, anger, and rejection. Some raters use anonymity or working relationships to score the target employee too high or too low, to prove a point.¹⁹ Again, the emotional aspect of the instrument for raters and target employees can affect its accuracy and reliability, calling into the question the validity of the results. In order to preserve the face validity of the data, organizations should educate, coach, and counsel target individuals before they receive feedback. This is important because as overall ratings became less positive (further from the best), leaders become less likely to accept the feedback as accurate.²⁰ Emotional involvement in the ratings, and how they are reported to the target individual, can affect that person’s perception of the accuracy of the results. According to Theron and Roodt,

egocentric biases are common in perceptions of ratings—people may inflate or deflate them as a defensive measure, or they may attribute good performance to their own behavior and poor performance to environmental factors.²¹ The target may reject the results as invalid and seek external sources to blame for perceived failures. Avoiding this kind of situation depends on integrating third-party coaching and a longitudinal development plan.

Appropriate Uses of Multirater Feedback

Before any multirater feedback assessment instrument can be used, the organization must create a psychologically safe learning environment that values individual development and feedback for development.²² This environment leads to managers and employees accepting the feedback as valid and using it to improve their behaviors—which



Multirater Feedback Diagram

increases the organization's effectiveness.²³ According to Manuel London and James W. Smither, by valuing feedback for improving performance, the organization allows the target individual to associate feedback with critical events, make meaning, and assimilate the feedback into developmental goals.²⁴ According to Mark D. Cannon and Robert Witherspoon, without a conducive and safe environment, the multirater feedback has the potential to negatively affect performance by reducing morale, increasing suspicion, increasing negative competitiveness, and reducing organizational citizenship behaviors.²⁵

The most heralded benefit of multirater feedback comes from the analysis of the difference between the self-ratings and the ratings from the other respondents. Allan H. Church defines self-awareness as the congruence between how managers view themselves and how others view them.²⁶ This delta reflects the level of self-awareness that the target individual possesses; it is a crucial element in leader development and leadership. Research shows that the more self-aware individuals are about their actions and their effect on others, the more leadership potential they have, and self-aware leaders tend to outperform others.²⁷ There is much argument over how to measure self-awareness statistically and track its change over time in a longitudinal study. Caroline Bailey and Clive Fletcher assert that a theory relating the effects of multirater feedback on self-awareness and performance is needed to establish the validity of the instrument.²⁸

It is important to use multirater feedback for developmental purposes, rather than for summative appraisals. Making employees more aware of the behavior competencies the organization rewards will enable managers to align an individual's "performance schema and the performance criteria of the organization."²⁹ When employees' behaviors move closer to the organization's values through self-awareness, their job performance improves. In addition, Fecteau et al. report that "leaders reacted more favorably to evaluations from subordinates if those evaluations were used for developmental and not for administrative purposes."³⁰ This is because the leaders felt psychologically safe from results that could have caused poor performance

evaluations—shielding their income and promotion potential. Even if the target individual believes the feedback is inaccurate, a coach or counselor can use the results to help increase the individual's perception of his or her performance.³¹ Leaders do not necessarily have to change their behaviors to please any respondent group (except, possibly, the superiors that rate them administratively). Instead, when they understand that others rate some of their behaviors as needing improvement, they can set personal goals for better performance in those areas. For an organization to be effective at fostering individual development, management should hold individuals accountable for creating development plans and provide the resources people need for improvement.³²

Additionally, research shows that ratings by subordinates, peers, supervisors, self, and others overlap only modestly; self-ratings correlate weakly with other rater perspectives, with greater convergence between peer and supervisor ratings.³³ This leaves the coach or mentor with varying perspectives on the person rated that must be interpreted for development. The divergence of the perspectives can make it difficult to evaluate behavior for performance assessment. This evidence reinforces the use of multirater feedback for development only—for individuals to improve their behaviors—but not for performance incentives.

Inappropriate Uses of Multirater Feedback

Much multirater feedback research cautions against using it as part of the appraisal process because it may lead to the target individual becoming too focused on pleasing others, especially subordinates, and not performing leader or managerial behaviors necessary for the job. Theron and Roodt state, "ratings used to determine employee reward and promotability are more prone to leniency bias," meaning that others will rate the target individual higher than the person truly performs in order to enable that person to be rewarded.³⁴ In 2003, IBM (International Business Machines Corporation) used multirater feedback as part of employee annual performance reviews, but the practice was halted due to the reviews becoming politically charged and thus unreliable.³⁵



(Photo by Sgt. Daniel Cole, 173rd Infantry Brigade Combat Team PAO)

A noncommissioned officer with Company A, 1st Battalion, 503rd Infantry Regiment, 173rd Infantry Brigade Combat Team (Airborne), provides feedback during an after action review following training on traffic-control point procedures 10 July 2013 during Exercise Rapid Trident 2013 in Yarivov, Ukraine.

It is also possible that using the multirater feedback in the performance appraisal process can create confusion as to a developmental program's objective and thus hinder its effectiveness.³⁶ Despite this caution, many organizations are developing systems to incorporate multirater feedback into performance appraisals. The negative effects on the appraisal process could be mitigated somewhat if the organization has established a learning environment that values development, learning, and feedback.³⁷ However, the risk of negative outcomes outweighs the potential rewards.

Effective multirater feedback has certain characteristics. Employees must receive specific and targeted feedback or they will not develop the insight needed to improve their self-awareness and performance, which results in a waste of the organization's money, time, and effort. To facilitate effective feedback to the target individual, the instrument should include a debrief by a third-party coach, counselor, or trained administrator.³⁸ Although costly, organizations cannot expect consistent and valid results without developing a coaching or counseling plan during the careers of the target individuals. Finally, if the multirater feedback is not used for

the development of the target individual, then the instruments will have little validity or usefulness—whether tied to performance appraisals or administrative action.

Effective Multirater Feedback in Army Organizations

Multirater feedback can be an enormously beneficial instrument to individuals and organizations by providing individuals with specific feedback on their behavior competencies as perceived by superiors, peers, subordinates, customers, and others. By comparing a self-assessment to the reports of others, individuals can improve their self-awareness, create a developmental plan of action, and even align their behaviors with those the organizational values. Without feedback from others, individuals would be hard-pressed to measure self-awareness.

Although the surveys have the potential for error, detailed and well-described behavior competencies can be properly framed to reduce latent error.³⁹ Multirater feedback can help the organization develop its employees at all levels. Not all toxic leaders are toxic in every job—staff positions will differ

from command billets. With the Army's job rotation system of development, using multirater feedback in annual evaluations would require significantly more frequent feedback sessions, overloading our already busy superiors, peers, and subordinates with another survey. Additionally, a single multirater feedback session does not enable behavior improvement in and of itself, but it requires multiple sessions over time to demonstrate real change, a process that eventually uncovers a lack of development or toxicity.⁴⁰

Now that we understand multirater feedback's strengths and weaknesses, we can look at how to integrate it with performance evaluations and CSL boards, as noted by Gen. Dempsey and Gen. Odierno.⁴¹ Odierno and other senior leaders acknowledge the weaknesses in the Army's Multi-Source Assessment Feedback (MSAF) system, which includes coaches and which requires an assessment to be started for completion of an OER.⁴² The current MSAF system is for developmental purposes only and offers optional coaching from a third party outside the chain of command—all supported by research as appropriate use and likely to enhance validity and reliability. However, the focus of this paper is not to evaluate the current MSAF system, but to illuminate issues and conditions that the Army should take into account if implementing multirater feedback into evaluations, promotion boards, CSL boards, or other performance-based assessments.

Where these instruments may work in the civilian business world, the Army is different in both its binding regulations and culture. Commissioned officers, warrant officers, noncommissioned officers, and other enlisted soldiers are all subject to the Uniform Code of Military Justice, which states that soldiers must follow lawful orders. "You do not have to like it—just do it," is an oft-quoted phrase. The Uniform Code of Military Justice regulates the military profession to ensure good order and discipline; this is an artifact of coercion and authority within the strong culture of the military.⁴³ This hierarchical and authoritative structure drives a strong culture of following orders, even if the subordinate does not agree with or like the order.

Also within the Army culture, as espoused at Army leadership schools, is the idea of "owning" orders that come down the chain of command, which means to avoid shifting blame to superiors

for orders subordinates may not like. This results in subordinates not always fully understanding the background of an unpopular order, which can skew how subordinates view a leader. If a leader follows this principle, then subordinates may never know how much the leader corrects, fights, or accepts unpopular taskings or orders. This knowledge could change how subordinates view their supervisor and affect subordinate inputs into multirater feedback.⁴⁴

Furthermore, even though the Army attempts to prepare leaders for future jobs with increased responsibility, most do not comprehend the jobs of their supervisors fully until they become supervisors. Bailey and Fletcher found that subordinates and peers may have insufficient experience to rate effectively, thereby reducing the accuracy and reliability of any performance rating or developmental feedback.⁴⁵ Multirater feedback has more consistency, according to Theron and Roodt, in routine and well-defined jobs.⁴⁶ However, the Army expects innovative and adaptive leaders to perform many complex tasks. Army leaders must solve complex problems rapidly and perform tasks that are not routine or well defined—such as individual counseling, unit evaluations, family support, media engagement, combat functions, and many more. Rating these tasks is unlikely to lead to reliable and valid data. Using these data in a performance evaluation or board selection would not only fail to solve the toxic leader problem but also would damage the reports so that the board would be promoting or selecting the wrong leaders based on faulty data.

The most glaring issue with using any multirater feedback data for a performance evaluation or selection board is that the multirater feedback instrument cannot predict performance, and it cannot guarantee that an increase in self-awareness of behavior will lead to improved performance.⁴⁷ The ratings, combined with self-assessment, enable comparison against the self-assessment only, illuminating the leader's self-awareness. The rated leaders can begin to understand how others view their performance. If their self-ratings differ significantly from the feedback, with proper coaching they can determine how to adjust their behaviors toward maximum performance for the organization.

An editorial in the *Army Times* asserts that multirater feedback "should be for leadership

development . . . These reviews can be addressed in counseling sessions . . . and [used] to inform an officer's . . . evaluation. But [360-degree reviews] should not be used as a means of measuring one officer against another for promotion."⁴⁸ Unfortunately, feedback becomes performance-linked as soon as it is viewed by the subject's supervisor, and if used to inform an "officer's evaluation," it is then being used for promotion purposes as the OER will be reviewed by the promotion board to determine if the officer should be promoted to the next rank.⁴⁹ The *Army Times* editorial suggests what the Army needs to understand explicitly: supervisors might—wrongly—use multirater feedback as part of officer evaluations.⁵⁰ The Army should not allow this because it would destroy multirater feedback as a means for professional development. It would directly reduce its validity and force officers into an unsafe learning environment while receiving "development" from their superiors—potentially reducing valued leader behaviors in units.⁵¹

Using quantifiable survey data is hazardous without understanding the data and structure of the survey. For example, on a scale of one to ten, a leader who receives a five from reviewers but overrates himself or herself at a nine could have more issues than a leader who receives a four and self-assesses at four. The latter is much more self-aware than the former—and, is a score of four poor? What is being measured? Moreover, can multirater feedback across year groups or branches be compared for promotion potential? Typical multirater feedback for developmental purposes focuses on the leader and how the



(Photo by Spc. Joshua Edwards, 129th Mobile Public Affairs Detachment)
Sergeant Major of the Army Raymond F. Chandler III speaks with soldiers 22 September 2013 during a visit at Forward Operating Base Walton, Afghanistan. Chandler traveled to bases throughout southern Afghanistan to speak with soldiers about their concerns regarding current and upcoming Army events.

leader's behavior affects others. Self-awareness is a crucial trait for good leaders and should be measured.⁵² If the leaders in the above example both increase the scores they receive from others to 10, have they achieved the organization's goals? Maybe not. Leaders may be getting better ratings over time at the expense of the organization—by managing how others perceive them rather than through true behavior change and job performance.⁵³

As Congress reduces the Army's budget and decreases the size of its force, competition for promotion, or even for retention, will increase. This could result in a zero-defect environment for tolerating failure. A zero-defect environment is not a safe learning environment that supports and encourages

leadership experimentation and behavior change, let alone risk taking and innovation.⁵⁴ Furthermore, London emphasizes that the organization must provide those being rated with the resources needed for change, third-party (outside of the chain of command) coaches for development, and organizational education for performance evaluations.⁵⁵ All of these requirements are expensive to develop, implement, and maintain. Without serious momentum, the parts of a multirater feedback system would be an easy budget-cut target. Canceling third-party coaches would cripple its use for leader development because the obvious, and already proposed, low-cost solution would be to have supervisors conduct leader development activities with the multirater feedback, immediately converting development into performance appraisal.⁵⁶ W. Warner Burke also states a high level of psychological safety is required of the organization to allow leadership experimentation and to build trust among its employees.⁵⁷ Using multirater feedback as a performance measure is very hazardous if done improperly and has the potential to erode organizational trust, arguably the most important component of leadership both in and out of the military.⁵⁸ Future reduced budgets could impact leader development or even the expansion of multirater feedback into the force.

Multirater Feedback Potential

There is a place for multirater feedback in the Army, with the clear choice being a system used only for leader development. If used in the performance evaluation and promotion systems, the multirater feedback instrument would require a completely different survey and considerable educating of the selection boards, raters, and administrative personnel prior to implementation and evaluation. The education of the selection boards and raters would be difficult to sustain since trained coaches, who make meaning of multirater feedback, are expensive and have their own biases. It would be very difficult to take biases and emotive data from respondents and make an objective

measure of performance for evaluations or selection for command. The Army created the CSL system under the Officer Professional Management System in 1971 to remove subjective bias from commanders and to create an objective and fair promotion and selection system.⁵⁹ By instituting multirater feedback as a direct part of performance evaluations, promotion boards, and command selection, the Army would be inputting subjective bias into an objective system. The Army should not link multirater feedback directly to any performance evaluation or selection process.

Conclusion

While it is possible to use several multirater feedback systems to serve different purposes, any such system tied to performance runs the risks of harming the organization. Kenneth Nowack states, “The potential ... adverse impact or emotional harm from such feedback intervention has often been imprudently overlooked by many coaches, despite a common focus on enhanced insight and self-awareness as major goals of the process.”⁶⁰ Using a highly emotive feedback system for performance data can be damaging to a performance-oriented and objective promotion-and-command selection system by reducing the perceived objectivity of the selection process.⁶¹ If the Army links multirater feedback to performance, its most damaging effects could be reducing multirater feedback’s powerful potential as a developmental tool, denying leaders a safe learning environment and, potentially the most damaging, focusing leaders on pleasing others rather than performing leader behaviors effectively.⁶² These negative effects are particularly likely when feedback is combined with inflated evaluations and a lack of performance and developmental counseling. For any multirater feedback system to be effective, the organization must hold employees accountable for their improvement, give them the resources required, and create a climate supportive of leadership experimentation and behavior change.⁶³ ■

Maj. Gregory G. Lee, U.S. Army, is serving as operations officer for 3rd Squadron, 3rd Cavalry Regiment, at Fort Hood, Texas. He is a graduate of the U.S. Military Academy at West Point and received his master’s degree in organizational psychology and leader development from Columbia University. He has deployed twice to Afghanistan and twice to Iraq.

Notes

1. U.S. Army Gen. Raymond T. Odierno, quoted in Michelle Tan, "Odierno to Top Leaders: Expect Tougher Evals," *Army Times* (8 April 2013).
2. Tim Kane, *Bleeding Talent: How the US Military Mismanages Great Leaders and Why It's Time for a Revolution* (New York: Palgrave Macmillan, 2012). The U.S. Army's tool for 360-degree assessment is called *multi-source assessment and feedback (MSAF)*. Officers are required to use the MSAF tool for self-assessment.
3. Chairman of the Joint Chiefs of Staff U.S. Army Gen. Martin Dempsey, reported in Andrew Tilghman, "Dempsey: 360-degree Reviews Could Help Select Better Leaders," *Army Times* (1 May 2013), <http://archive.armytimes.com/article/20130501/NEWS/305010023/Dempsey-360-degree-reviews-could-help-select-better-leaders>, accessed 21 April 2015.
4. Odierno, "Mission Command" (address, U.S. Army Command and General Staff College Officer Class of 13-02, Fort Leavenworth, Kansas, 9 April 2013).
5. Carolyn L. Fecteau et al., "Reactions of Leaders to 360-Degree Feedback from Subordinates and Peers," *The Leadership Quarterly* 9(4) (Winter 1998): 427-448; Terri Linman, "360-Degree Feedback: Weighing the Pros and Cons," 2004, http://edweb.sdsu.edu/people/arossett/pie/Interventions/360_1.htm, accessed 20 March 2015.
6. Clive Fletcher, Caroline Baldry, and Nicole Cunningham-Snell, "The Psychometric Properties of 360 Degree Feedback: An Empirical Study and a Cautionary Tale," *International Journal of Selection and Assessment* 6(1) (1998): 19-34.
7. Ibid.
8. Manuel London, *Leadership Development: Paths to Self-Insight and Professional Growth* (Mahwah, NJ: Lawrence Erlbaum Associates, 2002).
9. Ibid.
10. Fiona Dent and Judy Curd, "Psychometric Tests: Using 360-Degree Questionnaires and Feedback: The Practicalities," *Training Journal* (March 2004).
11. Caroline Bailey and Clive Fletcher, "The Impact of Multiple Source Feedback on Management Development: Findings from a Longitudinal Study," *Journal of Organizational Behavior* 23(7) (2002): 853-867.
12. D. Theron and G. Roodt, "Variability in Multi-Rater Competency Assessments," *SA Journal of Industrial Psychology* 25(2) (1999): 21-27; Manuel London and James W. Smither, "Feedback Orientation, Feedback Culture, and the Longitudinal Performance Management Process," *Human Resource Management Review* 12(1) (Spring 2002): 81-100.
13. Theron and Roodt, "Variability," 22.
14. Fecteau et al., "Reactions of Leaders."
15. Bailey and Fletcher, "The Impact of Multiple Source Feedback," 865.
16. Ibid.
17. Theron and Roodt "Variability."
18. Dent and Curd, "Psychometric Tests," 1.
19. Ibid.
20. Fecteau et al., "Reactions of Leaders."
21. Theron and Roodt, "Variability."
22. Arthur Shriberg, David Shriberg, and Richa Kumari, *Practicing Leadership: Principles and Applications*, 3rd Ed. (Hoboken, NJ: John Wiley & Sons, 2005).
23. Maxine Dalton, "When the Purpose of Using Multi-rater Feedback is Behavior Change," in *Should 360-Degree Feedback Be Used Only for Developmental Purposes?*, eds. David W. Bracken, et al. (Greensboro, NC: Center for Creative Leadership, 1997), 1-7.
24. London and Smither, "Feedback Orientation."
25. Mark D. Cannon and Robert Witherspoon, "Actionable Feedback: Unlocking the Power of Learning and Performance Improvement," *Academy of Management Executive* 19(2) (2005): 120-134.
26. Allan H. Church, "Managerial Self-Awareness in High-Performing Individuals in Organizations," *Journal of Applied Psychology* 82(2) (1997): 281-292.
27. W. Warner Burke, *Organization Change: Theory and Practice* (Thousand Oaks, CA: Sage Publications, 2002); Church.
28. Bailey and Fletcher, "The Impact of Multiple Source Feedback."
29. Ibid, 854.
30. Fecteau et al., "Reactions of Leaders," 429.
31. Ibid.
32. London, *Leadership Development*.
33. Kenneth Nowack, "Leveraging Multi-Rater Feedback to Facilitate Successful Behavioral Change," *Consulting Psychology Journal: Practice and Research* 61 (4) (December 2009): 280-297.
34. Theron and Roodt, "Variability," 22.
35. Linman, "360-Degree Feedback."
36. London, "Variability."
37. Shriberg, Shriberg, and Kumari, *Practicing Leadership*.
38. Cannon and Witherspoon, "Actionable Feedback."
39. Theron and Roodt, "Variability."
40. Bailey and Fletcher "The Impact of Multiple Source Feedback"
41. Tilghman, "Dempsey: 360-degree Reviews."
42. Odierno in Tan, "Odierno to Top Leaders," and Odierno, "Mission Command."
43. Edgar H. Schein, *Organizational Culture and Leadership*, 3rd ed. (San Francisco, CA: Jossey-Bass, 2004).
44. Bailey and Fletcher, "The Impact of Multiple Source Feedback."
45. Ibid.
46. Theron and Roodt, "Variability."
47. Victoria A. Pollman, "Some Faulty Assumptions That Support Using Multi-Rater Feedback for Performance Appraisal," *Should 360-Degree Feedback Be Used Only for Developmental Purposes?*, eds. David W. Bracken et al. (Greensboro, NC: Center for Creative Leadership, 1997), 7-9.
48. *Army Times* editorial, "Targeting Toxic Leaders," *Army Times.com*, 2 May 2013, <http://archive.armytimes.com/article/20130502/NEWS01/305010027/Targeting-toxic-leaders>, accessed 21 April 2015.
49. Fecteau et al., "Reactions of Leaders."
50. *Army Times* editorial.
51. London, "Leadership Development."
52. Burke, *Organization Change*.
53. S. Bartholomew Craig and Kelly Hannum, "Research Update: 360-Degree Performance Assessment," *Consulting Psychology Journal: Practice and Research* 58(2) (Spring 2006): 117-122.
54. Kane, *Bleeding Talent*.
55. London, "Leadership Development"
56. *Army Times* editorial; London.
57. Burke, *Organization Change*.
58. Ibid.
59. Kane, *Bleeding Talent*.
60. Nowack, "Leveraging Multi-Rater Feedback," 283.
61. Dent and Curd, "Psychometric Tests."
62. Craig and Hannum, "Research Update;" Theron and Roodt, "Variability."
63. London, "Leadership Development"