



President Donald Trump leads a strategic assessment discussion with senior military leaders 5 October 2017 at the White House in Washington, D.C. (Photo by Yuri Gripas, Reuters)

# Are We There Yet? Implementing Best Practices in Assessments

Col. Lynette M. B. Arnhart, PhD, U.S. Army, Retired

Lt. Col. Marvin L. King, PhD, U.S. Army

**T**he purpose of a strategic assessment is to determine if an organization is achieving its strategic objectives. This is often a difficult process to

implement, given normal staff aversion to introspective processes and a lack of doctrine specific to assessments. The purpose of this article is to discuss best practices and

common pitfalls in military assessments while outlining steps needed to continue to improve assessments across the Department of Defense (DOD). First, we outline the doctrine and literature guiding the DOD. Second, we provide a review of common assessment methods used across the military. Next, we present the four best practices proven successful in the joint staff, strategic commands, and recent conflicts. Last, we provide recommendations on how to improve the state of assessments in the DOD.

## Doctrine

Joint Publication (JP) 5-0, *Joint Operation Planning*, and JP 3-0, *Joint Operations*, provide doctrine to the joint force on the staff processes and methods from receipt of mission through developing and implementing a vision and strategy.<sup>1</sup> For implementation of assessments, JP 5-0 and Joint Doctrine Note 1-15, *Operation Assessment*, provide general frameworks for implementing an assessment process within a joint staff.<sup>2</sup> While joint doctrine reserves comment on methods and techniques, multiservice doctrine compensates for this shortfall, outlining existing methods, assisted by a number of journal articles describing successful methods used in Iraq and Afghanistan.<sup>3</sup> The process for gap assessment run by the joint staff to collect data from the combatant commands (CCMDs), outlined in various policies and instructions, is conducted through the Annual Joint Assessment (AJA, formerly known as the Comprehensive Joint Assessment, CJA) and tasked in the Guidance for Employment of the Force.<sup>4</sup> The joint staff recently added additional policy providing common joint terminology for risk in its publication of the Chairman of the Joint Chiefs of Staff Manual (CJCSM) 3105.01, *Joint Risk Analysis*, which allows clear communication of the results of an assessment from one echelon to another.<sup>5</sup>

The consistent themes across doctrine include descriptions of common staffing processes such as boards, bureaus, and working groups; discussion of data calls and data collection during the assessment process; and emphasis on commander involvement, while continuing to adhere to legacy terms from effects-based assessment. Literature, mostly from federally funded research and development centers, provides current methods in assessments, while

doctrine only partially assists the joint force in informing assessment methods, as we outline later in this article. While doctrine provides an overview of how to implement a process and a few of the main techniques, neither doctrine nor other supporting military publications

provide clear guidance on best practices. This lack of guidance contributes to a joint environment where

**Col. Lynette M. B. Arnhart, PhD, U.S. Army, retired**, is the former division chief for analysis, assessments, and requirements at U.S. Central Command and was responsible for developing and establishing the command's quarterly assessment of the Coalition Military Campaign Plan to Defeat ISIS and for conducting and continuous improvement of the Annual Theater Campaign Plan and Annual Joint Assessments. Arnhart served as a field artillery officer at the battery, battalion, and brigade level. Later she served as a commander in the Adjutant General's Corps, then as an operations research analyst. She has significant experience at the Headquarters, Army, and combatant command level, and she has conducted and led analysis of all types—strategic assessments, human capital, weapon-system effectiveness, modeling and simulation, programming and budgeting, and decision analysis. Arnhart earned a BS from the United States Military Academy, an MS from the Colorado School of Mines, and a PhD in operations research from George Mason University.

**Lt. Col. Marvin King, PhD, U.S. Army**, is a directorate senior military analyst at the Training and Doctrine Command Analysis Center. He is the former assessments, analysis, and studies branch chief for the Africa Command J-8, responsible for quarterly and annual strategic assessments, analytic support of the Integrated Priority List and Program Budget Review issues, sponsored studies, and wargaming from 2015–2017. Working at the Center for Army Analysis, he deployed as an ORSA in Iraq and Afghanistan, and developed quantitative wargame methods that were used to conduct analysis for those theaters. King possesses a BS in electrical engineering from the United States Military Academy, an MS in engineering management from the University of Missouri–Rolla, an MS in mineral and energy economics from the Colorado School of Mines, and a PhD in operations research from the Colorado School of Mines.



(Figure by authors)

**Figure 1. Example of an Unstandardized Thermograph**

there is no authoritative delineation between good and bad practices, and display techniques for condensing and conveying assessments of data.

### Inadequate but Common Assessment Methods and Display Techniques

To understand best practices, leaders should recognize inadequate assessment methods in use across the DOD and their corresponding narratives in data displays.

Three characteristics prevail among these techniques: lack of standards, subjective data displays, and inadequate source material. These methods and techniques, using monikers defined by their display, include thermographs, standardless stoplights, color averages, simple arrows, indices, one-hundred-point scales, and effects-based assessment. With little literature and no joint doctrine to provide assessment teams the foundation to cite the faults of these methods, it is difficult for commands to leave these techniques behind.<sup>6</sup> This article provides knowledge to inform leadership and empower assessment teams to build their credibility with other staff sections by building their expertise in assessment methods. The paragraphs below describe these inadequate methods and explain why each is a poor assessment technique.

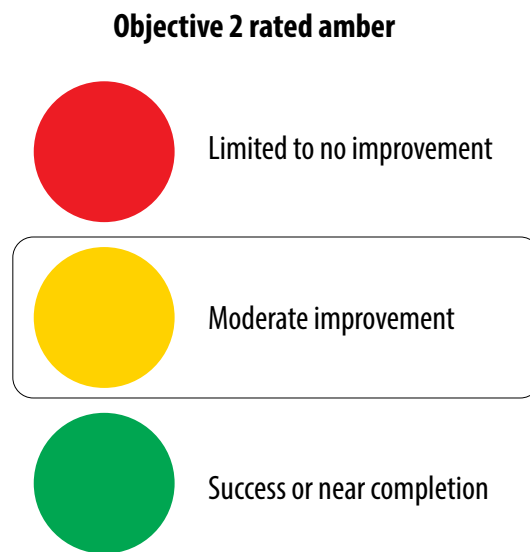
**Thermographs** contain a continuum of rainbow colors, normally red on one side, green on the other, and

yellow between them, with the current status marked with a triangle or tick mark to indicate the current rating (see figure 1). This technique often fails to provide an empirical standard to determine how far to move the progress indicator, leading a staff to move progress indicators subjectively in increments as measures of performance achieved, not as objective measures of verifiable effects achieved. Although they appear to have technical sophistication, “thermographs create the illusion of science,” as there is seldom any quantitative backing for the assessment.<sup>7</sup>

**The standardless stoplight**, consisting of a red-amber-green scale, is the most common form of assessment and is essentially a simplified thermograph (see figure 2). A common practice is to use these colors to create a subjective display, or an evaluation of progress without parameters, absolving the briefing agency of accountability for evaluating progress against a verifiable standard in their assessment. Every stoplight chart should have, at a minimum, a legend providing the short version of what the colors mean on the chart and a written narrative

fully detailing the standards-based bins in reserve.

**Color math, or color averaging**, involves identifying a color for a single indicator, assigning it a number value, using it as part of an index with other indicators, and then translating the index back into a color (see figure 3, page 23). This process treats ordinal variables as continuous variables; the average of ordinal responses is



(Figure by authors)

**Figure 2. Example of a Subjective Stoplight**

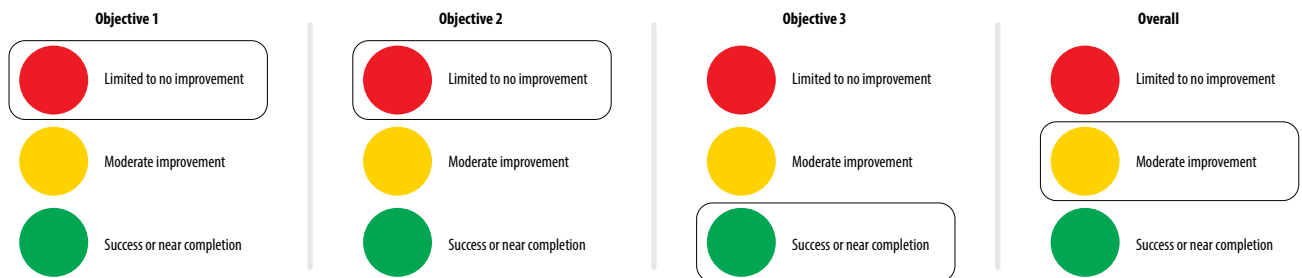


meaningless and in some cases misleading. Consider, for instance, a situation where five of ten provinces are successful and the other five are failures. If one averages the responses, the assessment would be “amber,” or “marginal success.” This provides a clear example of a faulty assessment; it is far more insightful to assess half of the provinces as failures and half as successes.

**Arrows—up, down, and sideways**—provide a single indicator noting only the change from the last report (see figure 4, page 24). Arrows show short-term advances for the sake of demonstrating progress but ignore more im-

Making transparency even more difficult, assessors often leverage proxies for many indicators when substantial data does not exist, thereby degrading the legitimacy of insights analysis may provide.

**One-hundred-point scales** source data through a survey, with multiple subordinate commands and/or directorates voting on the status of an objective using a scale of 1 to 100 with the overall score being the average of the votes. While there are general rules on the scoring for these surveys, our ability to measure the difference between natural states is not refined enough for the assessor



(Figure by authors)

**Figure 3. Example of Faulty Color Averaging**

portant long-term trends based on mission accomplishment. The end result of these assessments are uncannily predictable, with approximately one-third to one-half of the objectives assessed with up arrows to demonstrate some success, regardless of the actual scale or progress towards mission accomplishment.

**Indices** comprise a weighted average of normalized data. The purpose of an index is to have a single indicator summarizing an aspect of a problem (see figure 5, page 24). Indices are useful when experts agree on the weights applied to the input data, and the data is used to compare like items, such as state fragility indices. (They combine scores measuring two essential qualities of state performance: effectiveness and legitimacy; these two quality indices combine scores on distinct measures of the key performance dimensions of security, governance, economics, and social development.) Most indices for assessments are not transparent enough to provide value, such as when multiple indicators contribute to the increase or decrease of an index, hiding the key indicators. Further, weighted averages assume a consistent linear relationship and quality data collection, rarely found in the complex problems the military attempts to measure.

to discern the difference between, for instance, 67 and 68, rendering measurement to this fidelity, and the corresponding assessment conclusions, meaningless.

**Effects-based assessment.** Despite being purged from joint professional military education, effects-based operations and the associated assessment process persist throughout doctrine and application in the joint force.<sup>8</sup> There are two distinct problems with effects-based assessment. First, it assumes a deconstructionist mentality, that is, effects “roll up” into intermediate military objectives (IMO). Multiple authors, military and civilian, warn against such a mindset.<sup>9</sup> Second, the structure of lines of effort (LOEs), IMOs, and multiple contributing effects tend to bloat staff requirements for data collection without corresponding benefit to the staff.<sup>10</sup> Because of the prominence of effects-based assessment, assessment sections are expected to collect vast amounts of quantitative data; efficient assessment sections use a streamlined assessment framework to process only the essential data required to measure the progress of their IMOs.






So we might ask ourselves why we continue to use these methods? Quite simply, assessment team members are very often assigned without sufficient education,

training, or prior experience in assessments. Even if assessments personnel have experience, there is little documentation for them to use as references for their methods when meeting organizational resistance within their own staff. The next section provides alternative, proven methods that are manageable in their implementation.

## Better Means for Strategic Assessments

Effective assessment practices clearly articulate progress, gaps, and the risk associated in accomplishing the unit's mission. Gap assessment, strategic questions, standards-based assessments, and written products best provide the tools required to assist operational and strategic commands.

**Gap assessment.** One outcome of an assessment process is to determine progress against a mission. When it becomes apparent we will not accomplish an objective by the target date, it raises the question of what to do next. A structured method to align assessments to answer this question is gap assessment, which defines the gaps in the critical path to obtain a given objective along a timeline. These gaps generally fall into the categories of capacity (insufficient forces allocated or assigned to the command, lack of authorities and/or permissions granted by the U.S. government); capability (shortfalls in any of the doctrine, organization, training, materiel, leadership and education, personnel, or facilities); or shortcomings in the willingness, capability, or capacity of partner nations. Identifying these gaps and attempting to close them provide the staff with a method to take action leading to the accomplishment of their strategic objectives. In the joint staff, the gap assessment is initiated by CCMDs in the AJA and summarized in the Chairman's Risk Assessment through the Capability Gap Assessment.<sup>11</sup> Similar, less formal structures exist in a few of the CCMDs, while other commands focus on recommendations.

Mission accomplishment	
Objective 1	
Objective 2	
Objective 3	
Objective 4	
Overall	

(Figure by authors)

**Figure 4. Example of Unpredictable Arrows**

Mission accomplishment index (1-10)		
	Score	Weight
Objective 1	3	0.4
Objective 2	5	0.3
Objective 3	10	0.2
Objective 4	5	0.1
<b>Overall</b>	5.2	-

(Figure by authors)

**Figure 5. Example of an Oversimplified Index**

**Strategic questions.** In determining progress and gaps for a given LOE or IMO, several common questions arise. Recording these questions is a practice in many assessment

programs because it allows those responsible for the assessment a method to record, in detail, the assumptions and the logical lines followed by working groups to determine why they believe they are progressing or retrogressing. In reviewing these questions on a periodic basis, the working groups revisit their assumptions and their progress, considering changes in the operational environment. While strategic questions are sometimes informed by indicators, indicators are not required if the question is qualitative in nature. Some example questions are shown in figure 6 (on page 25).<sup>12</sup>

**Standards-based assessments.** The method providing the most accurate and successful summation of progress through operational and strategic commands is standards-based assessments. There are four reasons why we advocate for the use of standards-based assessments. First, it is important to display data at the resolution we can effectively measure. For a military objective, this

means dividing the possible states of the operational environment into mutually exclusive bins—that is, described in sufficient detail so it is clear the progress of the objective resides in only one bin. Second, standards-based assessments relate only to the military objective's progress.

Often assessment processes confuse rating scales between progress, resource allocation, and risk.

Third, standards-based binning facilitates gap analysis. By listing the current state and the desired state, working groups can determine future operations or activities required to move between bins and associated capability, capacity, and authority gaps bridging

between the two states. Last, binning provides a method to hold subordinate commands and staff accountable for their evaluation; the evaluator must provide evidence that an IMO is in a bin. The process results in a method of clearly rating the progress toward an objective. An example of a standards-based scale, or binning, is shown on the left side of figure 7 (on page 26).

In implementing a standard-based bin, a working group may employ the following steps:

**1. Determine the goal.** The military objective, normally an IMO end state, is defined as the goal condition. If the end state is not clear at any point in the process, it is revised by adding more detail. This becomes the top bin, or goal state of the objective.

**2. Determine the worst case.** We define worst case as the worst possible state of progress, including states the IMO could regress to in the future.

**3. Determine the additional bins.** Determine the main indicators of what you want to discern between additional levels, and define the terms you wish to use to make this determination. Break the possible states of nature into natural breaks based on these terms, normally three to seven bins for a single objective.

If there is a history of the state of the objective, take each year of the prior observations for the conflict, as well as all possible future states of the objective, and a short description of each year, and place them on a continuum between the best and worst cases. This provides

### Stability Operations

- Do perpetrators of political violence find sanctuary and support in neighboring states?
- Do political leaders or elites accept and support the peace settlement?

### Peace Operations

- Are external actors aiding combatants?
- Have the disputing parties resolved their major disagreements?

### Counterinsurgency Operations

- To what extent have military operations deterred the actions of terrorist groups?
- To what extent have group members been effectively removed by counter-network actions?

*Note: Appropriate follow-on questions are required for questions requiring only a "yes" or "no."*

(Figure by authors. Questions are taken from multiple sources; see endnote 12 for sources)

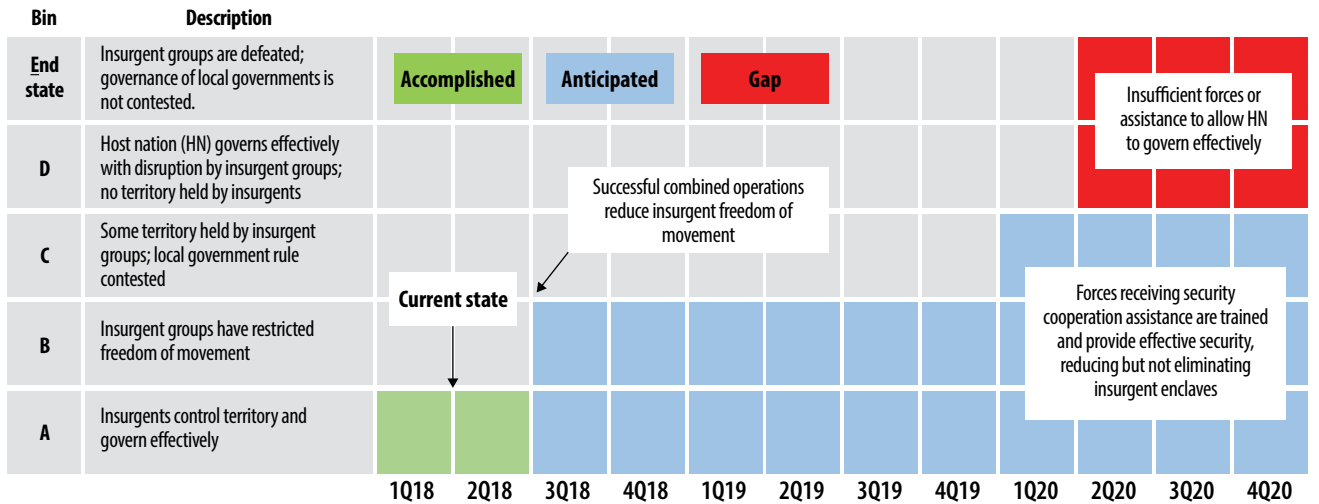
## Figure 6. Example Strategic Questions

a pool of prior and future states the working group can then compile into similar bins.

**4. Refine the bins.** Given the grouping of prior observations, each bin is described in at least a paragraph using the evaluation terms described in step three. Each bin is described in sufficient detail so there is no question as to which bin a given scenario belongs. Bins are collectively exhaustive (every observation fits somewhere in the bins) and may possess mutual exclusivity (each observation can only fit in one bin) or build upon each other (each observation fits into a bin and all the bins below or above it).

**5. Additional means.** If the division of natural states proves problematic, additional observations are used by taking a similar historic situation and placing the observations by year on a continuum between the best and worst cases, then compiling these into similar bins. Using historical examples is helpful because people relate better to conflicts they have experienced, as long as the working group ensures the historic example is relevant to the current objective.

**6. Plan to achieve the end state.** Using the developed bins, plot a course from the present state until the stated date of the objective, similar to a critical-path



(Figure by authors)

**Figure 7. Abbreviated Standards-based Binning Example, with Projected Progress and Gaps for Three Years**

method. Then, using planned activities and operations, determine remaining gaps. This is best executed with the synchronization matrix developed from a wargame while planning the campaign. A graphical tracking representation is presented in figure 7.

On a recurring basis (generally quarterly), the gaps across the IMOs are collected and prioritized, validated, and acted on by other staff processes. While any working group structure may implement this method, two important disruptions frequently occur. First, working groups must design bins so they are mutually exclusive. Just as standards for training must be “trained” or “untrained,” IMO ratings cannot have a “2 plus” or “low 3,” analogous to an “almost trained” rating. Using amplifications to ratings defeats the purpose of binning, gives constructive credit for task accomplishment rather than effect accomplishment, and does not hold the working group accountable for identifying gaps.<sup>13</sup> Second, accountability for rating the IMO must remain with the working group and the IMO/LOE working group lead, not the assessment team collecting and checking the ratings. This separation of evaluator—responsible for the rating—and assessor—responsible for the process and written document—keeps the working group focused on accomplishing the end state; otherwise, narratives diverge into listing activities accomplished rather than effects. Implementing this requires IMO/LOE lead presence at all senior leader assessment briefings to keep accountability and responsibility affixed to the IMO/LOE leads.

**Written documents.** Possessing a written document detailing the command’s assessment is important for several reasons. First, the level of thought, staff coordination, and detail required to articulate the rating of an assessment in words and sentences is far greater than what is required to fill out a chart template. Many assessment processes suffer from lack of detail without a corresponding written document to further explain the nuances of the assessment. This explanation is vitally important because charts without background information are susceptible to a special form of groupthink.<sup>14</sup> These problems are so pervasive that some leaders and analysts recommend exclusive use of written assessments collated from subordinate assessments.<sup>15</sup>

**Written risk assessments.** A written assessment is often the only way to articulate risk in a meaningful manner. CJCSM 3105.01 provides comprehensive definitions of military and strategic levels of risk. A written document can provide the reason for the evaluation of risk, an audit trail based on a gap relating the failure to meet an IMO, LOE, and theater-campaign-plan end state, determined in the standards-based assessment and amplifying facts and data to shore up the argument for the assessment of risk. An example of a written risk assessment begins with a statement of the objective or end state, describes the current level of progress determined from the standards-based bins, evaluates the risk of meeting strategic and military objectives, and identifies the gaps.

**Table. Application of Best Practices in Context of a Gap Assessment**

Best Practice	Input	Output	Advantages	Disadvantages
Strategic Questions	Q&A seeking to draw out issues for discussion	Gaps, validated assumptions	<ul style="list-style-type: none"> <li>· Scopes the problem</li> <li>· Clarifies and revalidates assumptions</li> <li>· Assists in identifying gaps</li> <li>· Provides and input for higher headquarters to add to the assessment</li> </ul>	Quality of responses can vary
Standards-based Assessment	Intermediate military objectives (IMOs)	Gaps in the plan with an audit trail of IMO, line of effort and endstate whose accomplishment is at risk	<ul style="list-style-type: none"> <li>· Measures progress</li> <li>· Determines present and future gaps</li> <li>· Provides an audit trail for the gap</li> <li>· Validates gaps are traceable to the mission</li> </ul>	IMOs cannot change quickly
Written Risk Assessment	Gaps in the plan	Risk of the gaps impacting the ability to accomplish the mission	<ul style="list-style-type: none"> <li>· States the evaluation of progress</li> <li>· States strategic and military risk</li> <li>· Explains the audit trail to the mission</li> <li>· Lists gaps for higher headquarters</li> <li>· Consolidates various requirements into one document</li> </ul>	Analyzes gaps, does not produce them

(Table by authors)

**Other best practices.** This article focuses on the implementation of a gap assessment given a set of objectives. Other best practices exist in closely related literature, such as logic models, also known as theories of change, or shared diagnosis models, which ensure objectives and measures result from a logical process derived from causal assumptions.<sup>16</sup> While preferred, these methods are difficult to gain consensus to implement, often competing with center of gravity analysis when applied. Additional best practices include using objective development criteria, such as the acronym SMART (specific, measurable, achievable, relevant, and time bound) and the similar initialism RMRR (relevant, measurable, responsive, and resourced).<sup>17</sup> Best practices related to staff organization and functions include assigning senior leaders as line of effort leads and gaining championship by the commander.<sup>18</sup>

## Integrating Best Practices into Assessment

The best practices by themselves do not make a complete assessment; linking them together provides value to the command in the form of insights, gaps, recommendations, and risk. The combination of strategic questions, standards-based assessments, and written assessments—particularly risk assessments—complement each other in

the types of input they accept and the type of output they produce as they relate to the gaps they identify. Successful assessments attempt to leverage all the best practices to best detail progress, identify gaps, make recommendations, and articulate residual risk. An outline of the application of each of the methods in the context of gap assessment is shown in the table.

One example of a successful assessment is the process at NATO's International Security Assistance Force from 2010 to 2013, which leveraged strategic questions, standards-based binning, and written assessments to conduct internal assessments as directed by the National Defense Authorization Act. This shift marked the recent advancement in assessment methodology.<sup>19</sup> The continuing evolution of the joint staff-directed AJA (and former CJA) illustrates the difficulty of moving assessment practitioners and staff to processes that result in a truly useful and informative product. The most recent CJA process for gap assessment used strategic questions; it directs structured written assessments of gaps but struggles with implementation of a consistent standard across the CCMDs for their standards-based assessment. It employs a sliding scale conflating achievement and progress, which confuses commanders and ultimately does not provide the information needed to drive decisions. As



we move to the AJA, this practice should be abandoned. This is especially critical, as in the absence of clear joint doctrine, subordinate commands are replicating this type of conflated scaling or abandoning otherwise solid assessment processes due to the resulting confusion in portrayal.

In the evolution of assessments in the CCMDs, the use of the best practices proves useful for other reasons as well. First, CCMDs are required to provide assessments across multiple operations and plans. With limited staff resources, answering the requests for information for all required assessments individually consumes limited staff resources. Developing a well-managed periodic process based on the joint staff approved AJA assessment can help alleviate the burden of multiple assessments from the staff. Changes from each of the assessments must align in the operations and planning cycle; otherwise, recommendations may be outdated before they can be implemented.

To deal with multiple assessments in the joint environment, CCMDs and the joint staff have seen success in using the language of strategic questions, gaps, and risk as an efficient method. In this process, each level of command (joint staff, coordinating authority, CCMD, service component, and joint task force) produces their own assessment answering strategic questions and articulating gaps with associated risk. Higher headquarters provide strategic questions to lower headquarters that, when answered, inform all levels of assessment. Lower headquarters forward their gaps, along with military and strategic risk as outlined in CJCSM 3105.01. This provides simple methods for incorporating higher and lower assessment processes, which rarely align enough to truly nest. It also avoids multiple different assessments and methodologies converging from both higher and lower headquarters, which leads to confusion, apathy, and unhelpful recommendations.

## Recommendations, Summary, and Conclusion

To promulgate the best practices in assessments, the DOD requires vast improvements in doctrine,

education, and training for assessments, and continues to work to solve these challenges through a community of interest, staffed across the joint force. The latest Military Operations Research Society special meeting on assessments in February 2018 brought together many of the assessment practitioners in the community of interest from the DOD and international partners. The meeting focused on doctrine, education, and training for assessments.<sup>20</sup> We have briefly demonstrated above how shortfalls in these areas impede the adoption of best practices in our collective processes and believe we are at a sufficient stage to endorse the best practices and reject worst practices, as presented in this article. The further improvement of assessments in the DOD can be achieved by paying special attention to doctrine, formal education, and training. We have begun this process by advocating for and obtaining a special emphasis on assessment in joint professional military education, and we will continue to pursue a broader adoption of assessment improvements.

In this article, we have outlined a basic method of implementing strategic assessment techniques, explained why many widely used practices are inadequate, and detailed current best practices, providing references for both. We have offered ideas for proven implementation methods and outlined how the joint force can indoctrinate the best practices to better measure progress against strategic objectives and articulate gaps. We recommend the joint staff better incorporate best practices into doctrine, education, and training. Without improving, the joint force will continue to rely on assessment teams to conduct assessments with varying degrees of quality and utility. Inadequate assessments lead to the command having a lack of clear understanding of their progress against objectives and an inability to clearly articulate refined and tested gaps, which ultimately impacts programming of limited and valuable resources to provide capability to our fighting forces. ■

---

## Notes

1. Joint Publication (JP) 5-0, *Joint Planning* (Washington, DC: U.S. Government Publishing Office [GPO], 16 June 2017); JP 3-0, *Joint Operations* (Washington, DC: U.S. GPO, 17 January 2017).

2. JP 5-0, *Joint Planning*; Joint Doctrine Note 1-15, *Operation Assessment* (Washington, DC: U.S. GPO, 15 January 2015).

3. Multi-Service Tactics, Techniques, and Procedures (MTTP), *Operation Assessment: Multi-Service Tactics, Techniques, and Procedures for Operation Assessment* (Joint Base Langley-Eustis, VA: Air Land and Sea Application Center, August 2015), accessed 24 January 2018, <http://www.alsa.mil/mttps/assessment/>; Jonathan Schroden et

al., "A New Paradigm for Assessment in Counter-insurgency," *Military Operations Research* 18, no. 3 (2013): 5–20, accessed 24 January 2018, [https://www.researchgate.net/publication/263040323\\_A\\_New\\_Paradigm\\_for\\_Assessment\\_in\\_Counter-insurgency](https://www.researchgate.net/publication/263040323_A_New_Paradigm_for_Assessment_in_Counter-insurgency); Jonathan Schroden, "A Best Practice for Assessment in Counterinsurgency," *Small Wars & Insurgencies* 25, no. 2 (2014): 479–86, accessed 24 January 2018, <http://www.tandfonline.com/doi/abs/10.1080/09592318.2014.904032>; Ben Connable, *Embracing the Fog of War: Assessment and Metrics in Counterinsurgency* (Santa Monica, CA: RAND Corporation, 2012), accessed 24 January 2018, [http://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND\\_MG1086.pdf](http://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1086.pdf).

4. The Comprehensive Joint Assessment (CJA), now referred to as the Annual Joint Assessment (AJA), is intended as the primary data call for the joint staff for campaign assessment purposes, among other data calls. It provides data for many aspects of the PPBE process, outlined in the following documents. "The GEF [Guidance for Employment of the Force] is classified written guidance from the Secretary of Defense to the Chairman for the preparation and review of contingency plans for specific missions. This guidance includes the relative priority of the plans, specific force levels, and supporting resource levels projected to be available for the period of time for which such plans are to be effective." See the U.S. Army War College overview for more details, <https://ssl.armywarcollege.edu/dde/documents/jsps/terms/gef.cfm>; Department of Defense Directive Number 7045.14, *The Planning, Programming, Budgeting, and Execution (PPBE) Process* (Washington, DC: U.S. GPO, 29 August 2013), accessed 24 January 2018, <http://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodd/704514p.pdf?ver=2017-08-29-132032-353>; Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 3100.01C, *Joint Strategic Planning System* (Washington, DC: U.S. GPO, 20 November 2015), accessed 24 January 2018, [http://www.jcs.mil/Portals/36/Documents/Library/Instructions/3100\\_01a.pdf?ver=2016-02-05-175017-890](http://www.jcs.mil/Portals/36/Documents/Library/Instructions/3100_01a.pdf?ver=2016-02-05-175017-890); CJCSI 8501.01B, *Chairman of the Joint Chiefs of Staff, Combatant Commanders, Chief, National Guard Bureau, and Joint Staff Participation in the Planning, Programming, Budgeting and Execution Process* (Washington, DC: U.S. GPO, 21 August 2012), accessed 24 January 2018 [http://www.jcs.mil/Portals/36/Documents/Library/Instructions/8501\\_01.pdf?ver=2016-02-05-175057-247](http://www.jcs.mil/Portals/36/Documents/Library/Instructions/8501_01.pdf?ver=2016-02-05-175057-247).

5. Chairman of the Joint Chiefs of Staff Manual 3105.01, *Joint Risk Analysis* (Washington, DC: U.S. GPO, 14 October 2016).

6. Jonathan Schroden, "Why Operations Assessments Fail," *Naval War College Review* 64, no. 4 (Autumn 2011): 89–102; Stephen Downes-Martin, "Operations Assessment in Afghanistan is Broken: What Is to Be Done?," *Naval War College Review* 64, no. 4 (Autumn 2011): 103–25; Connable, *Embracing the Fog of War*.

7. MTTP, *Operation Assessment*.

8. James N. Mattis, "USJFCOM Commander's Guidance for Effects-based Operations," *Parameters* 38, no. 3 (Autumn 2008): 23, accessed 23 January 2018, <http://strategicstudiesinstitute.army.mil/pubs/parameters/Articles/08autumn/mattis.pdf>.

9. *Commander's Handbook for Assessment Planning and Execution* (Suffolk, VA: Joint Staff, Joint and Coalition Warfighting, 9 September 2011), accessed 23 January 2018, [http://www.dtic.mil/doctrine/doctrine/jwfc/assessment\\_hbk.pdf](http://www.dtic.mil/doctrine/doctrine/jwfc/assessment_hbk.pdf).

10. Mattis, "USJFCOM Commander's Guidance for Effects-based Operations," 22.

11. CJCSI 3170.01I, *Joint Capabilities Integration and Development System (JCIDS)* (Washington, DC: U.S. GPO, 23 January 2015), accessed 6 March 2018, [http://www.jcs.mil/Portals/36/Documents/Library/Instructions/3170\\_01a.pdf?ver=2016-02-05-175022-720](http://www.jcs.mil/Portals/36/Documents/Library/Instructions/3170_01a.pdf?ver=2016-02-05-175022-720).

12. The stability operations examples came from Michael Dziedzic, Barbara Sotirin, and John Agoglia, eds., *Measuring Progress in Conflict Environments (MPICE): A Metrics Framework for Assessing Conflict Transformation and Stabilization*, version 1.0 (report, Washington, DC: Army Engineer Corps, August 2008); the peace operations examples came from Paul F. Diehl and Daniel Druckman, *Evaluating Peace Operations* (Boulder, CO: Lynne Rienner, 2010). The counterinsurgency operations examples came from Jonathan Schroden, William Rosenau, and Emily Warner, *Asking the Right Questions: A Framework for Assessing Counterterrorism Actions* (Arlington, VA: CNA, February 2016)

13. Downes-Martin, "Operations Assessment in Afghanistan is Broken."

14. Edward R. Tufte, *Beautiful Evidence* (Cheshire, CT: Graphics Press, 2006), 157–85.

15. Michael T. Flynn, Matt Pottinger, and Paul D. Batchelor, *Fixing Intel: A Blueprint for Making Intelligence Relevant in Afghanistan* (Washington, DC: Center for New American Security, January 2010), accessed 23 January 2018, [https://s3.amazonaws.com/files.cnas.org/documents/AfghanIntel\\_Flynn\\_Jan2010\\_code507\\_voices.pdf?mtime=20160906080416](https://s3.amazonaws.com/files.cnas.org/documents/AfghanIntel_Flynn_Jan2010_code507_voices.pdf?mtime=20160906080416).

16. Christopher Paul et al., *Assessing and Evaluating Department of Defense Efforts to Inform, Influence, and Persuade: Desk Reference* (Santa Monica, CA: RAND Corporation, 2015), 32–33, 88–102, accessed 23 January 2018, [http://www.rand.org/content/dam/rand/pubs/research\\_reports/RR800/RR809z1/RAND\\_RR809z1.pdf](http://www.rand.org/content/dam/rand/pubs/research_reports/RR800/RR809z1/RAND_RR809z1.pdf); Susan Allen Nan and Mary Mulvihill, "Theories of Change and Indicator Development in Conflict Management and Mitigation" (Washington, DC: United States Agency for International Development, June 2010), Appendix A, accessed 8 February 2018, [http://pdf.usaid.gov/pdf\\_docs/Pnads460.pdf](http://pdf.usaid.gov/pdf_docs/Pnads460.pdf); David Kilcullen, *Counterinsurgency* (New York: Oxford University Press, 2010), 52–53; David Kilcullen, *The Accidental Guerrilla: Fighting Small Wars in the Midst of a Big One* (New York: Oxford University Press, 2011), 35.

17. Paul et al., *Assessing and Evaluating Department of Defense Efforts*, 73–78; JP 5-0, *Joint Planning*, D9–10.

18. This insight came from the group discussion from the Assessment Community of Interest Meeting, Tampa, Florida, 5–7 April 2017.

19. Schroden et al., "A New Paradigm for Assessment in Counter-insurgency"; Schroden, "A Best Practice for Assessment in Counterinsurgency," 479–86; Andrew Williams, James Bexfield et al., eds., *Innovation in Operations Assessment: Recent Developments in Measuring Results in Conflict Environments* (The Hague, Netherlands: NATO Communications and Information Agency, 2013), accessed 24 January 2018, [http://www.betterevaluation.org/sites/default/files/capdev\\_01.pdf](http://www.betterevaluation.org/sites/default/files/capdev_01.pdf).

20. The Assessments Community of Interest maintains a web presence on milbook at [https://www.milsuite.mil/wiki/Operation\\_Assessment](https://www.milsuite.mil/wiki/Operation_Assessment).