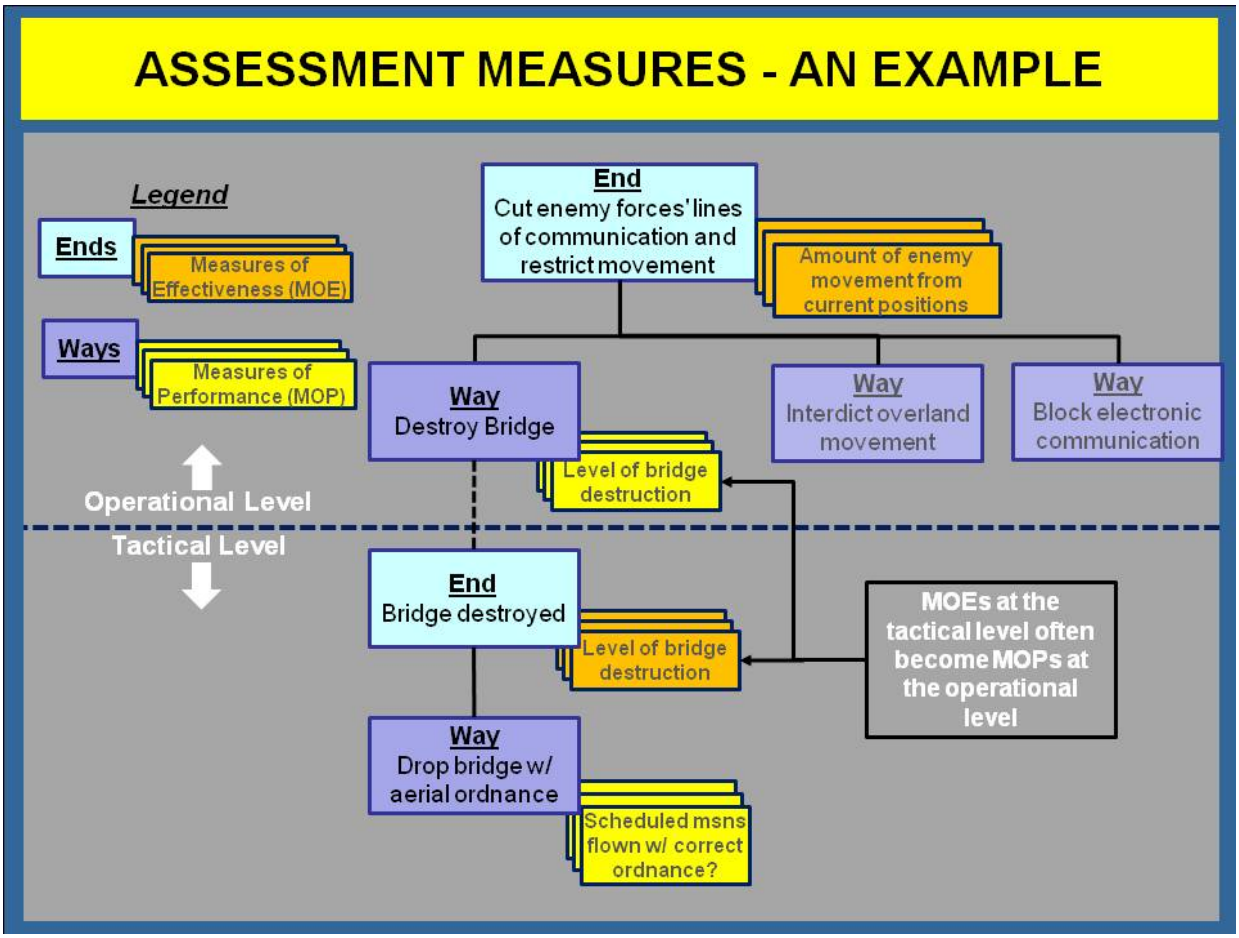
ASSESSMENT MEASURES

Last Updated: 04 November 2016

Assessment measures are simply the data elements that, via the criteria, provide insight into the effectiveness of the commander's strategy. Assessment measures are commonly divided into two types:

- **Measure of Performance (MOP)** — A criterion used to assess friendly actions that are tied to measuring task accomplishment.

- **Measure of Effectiveness (MOE)** — A criterion used to assess changes in system behavior, capability, or operational environment that is tied to measuring the attainment of an end state, achievement of an objective, or creation of an effect.



Assessment Measures — An Example

MOPs address the ways and means employed during execution to help achieve desired effects; they indicate progress toward accomplishing planned tasks or actions. MOEs assess progress toward creating desired effects and thus achieving the objectives and end state (Simply put, MOPs help tell us if we are doing things right; MOEs help tell us if we are doing the right things).

The distinction between MOEs and MOPs can depend on their context within the commander's strategy. The exact same measure can be an MOP for one commander and an MOE for another, lower echelon commander. The figure, "Assessment Measures—An Example" illustrates a practical application of this delineation.

Developing good measures is an art, though there are some general guidelines that can aid in developing high-quality measures:

- **Measures should be relevant and necessary**. Measures should relate to the effect or task they are intended to describe and should feed directly into the already-established criteria. Collection of irrelevant measures that do not shed light on the effectiveness of the commander's strategy is a misuse of valuable time and resources. Focusing primarily on collecting the data necessary to apply to the developed criteria should help avoid the creation of superfluous measures.

MOPs versus MOEs

A Simplified Example

A joint force air component commander (JFACC) working with the ground component attempting to stop a major enemy ground offensive might assess their forces' *performance* by measuring the number of interdiction sorties successfully flown against a crucial element of enemy follow-on forces. If the forces flew the planned number of sorties or more without loss, the JFACC can assess that forces are "doing things right."

The JFACC might assess *effectiveness* by measuring how many of the targeted enemy forces made contact with friendly forces in coherent platoon-size or larger formations. If that number is small, protecting friendly troops and effectively blunting the enemy offensive, the JFACC may conclude that the forces' efforts were effective—that they "did the right thing."

These are very different types of assessment, requiring different measures, and can lead commanders to very different conclusions. Too often, commanders may focus on MOPs (in part because they are more easily measured and yield empirical answers), and pay inadequate attention to MOEs.

Both are necessary, but conceptually different.

- **Measures should represent a scale, not a goal or objective**. Metrics developers may be tempted to write a goal or criterion as a measure. Instead, the goal should be included in the criteria in accordance with the commander's risk tolerance and thresholds. Operators and planners should establish these goals (objectives) in coordination with the assessors. Examples:

  - *Bad Measure:* no friendly fighter losses.

- ✪✪ *Better Measure:* number of friendly fighters destroyed or damaged by enemy air defenses.[1]

- ✪ **The data satisfying a measure should be observable, or at least inferable**. The measurements can be quantitative (numerical) *or* qualitative (non-numerical). In general, the more objectively measurable the better, but **commanders and planners should avoid "the numbers trap:" blindly using rates, numbers, and other quantitative metrics, especially in assessing effects, since their seemingly "empirical" and quantified elements may be based on wholly subjective assumptions** and the number may be meaningless—thus they may often lack direct linkages to the objectives or ends outlined in the strategy, while sometimes also imparting an illusion of "scientific validity" merely because they are quantified. Examples:

  - ✪✪ *Bad Measure:* civilian populace attitude toward stability forces.

  - ✪✪ *Better Measure (Quantitative):* percentage of surveyed civilian population giving "favorable" rating to stability forces; number of riots and civil disturbances in response to friendly force activities; amount of enemy propaganda, graffiti, and the like discovered; and so on.

  - ✪✪ *Bad Measure:* progress towards opening new air base.

  - ✪✪ *Better Measure (Qualitative):* current phase of air base stand-up (secured land, runway operational, 30-day sustainment capability in place, long-term sustainment capability in place).

- ✪ *Measures should be clear and concise.* They should be written in plain language so that someone with no prior knowledge of the measures can still understand the data requirements. Examples:

  - ✪✪ *Bad Measure:* status of enemy fighters.

  - ✪✪ *Better Measure:* number or percentage of enemy fighters confirmed destroyed or rendered combat-ineffective.

Measures should be drafted during planning so that associated intelligence collection needs may inform surveillance and reconnaissance requirements. Measures may need to be refined or amended during the tasking cycle as the operational situation changes. Selection of assessment measures is an iterative, ongoing effort.

Measure the entire plan, but do not overdo it. All elements of the strategy should be measured, and there may be multiple measures required to fully address the relevant criteria. However, attempting to assess too many measures can paralyze the assessment effort. Consider the value to the end result before adding more measures. Also consider what measures are readily available through immediate analysis of mission reports and planned collection tasking, rather than addressing new collection requirements. After assessors have built the entire set of measures, they should conduct a final review to identify those measures that have less relative

---

[1] Note that this is greatly oversimplifies the process, since measures such as that above would probably include friendly sorties forced to jettison ordnance—hence rendered mission-ineffective—due to enemy action, and similar measures.

importance/contribution or take inordinate effort relative to the insight provided, and remove them from the set. In general, assessment teams should prioritize their efforts to best support the commander's decision-making needs.