A cannibalized high mobility multipurpose wheeled vehicle (HMMWV) sits next to other HMMWVs awaiting repair. Such cannibalized vehicles or other pieces of equipment are sometimes used to skew unit readiness reporting requirements and mask unit logistical and materiel deficiencies. (Photo courtesy of J. VanDomelen, https://blogs.mentor.com/jvandomelen/blog/2011/12/28/power-problem-what-now/)

Unit Status Reports and the Gaming of Readiness

Capt. Theo Lipsky, U.S. Army
From March 2018 to November 2019, the Department of Defense Office of the Inspector General conducted an audit of the U.S. Army’s active component readiness. The audit’s resultant report was, on the whole, positive. Yes, the Army could use more low bed semitrailers, towed-howitzer telescopes, and electromagnetic spectrum managers. But overall, the Army had “met or exceeded” the goal of 66 percent of its brigade combat teams (BCT) reporting the “highest readiness levels for seven consecutive quarterly reporting periods.”

The unfortunate truth of the report, and others like it, is that it substantiates its findings with data from the Department of Defense Readiness Reporting System-Army (DRRS-A). The DRRS-A readiness data in turn comes from unit status reports (USR) provided by BCTs’ constituent battalions. This reporting labyrinth obscures what anyone who has compiled a USR knows: unit status reports are deeply flawed. The effects of those flaws are twofold: USRs not only fail to capture the readiness of reporting units, but they also actually harm the readiness of reporting units. The reports do so because they demand inflexible quantitative measurements unfaithful to the outcome they purport to depict—how ready a unit is to accomplish its mission. The commanders and staff chase readiness as the USR measures it, often at the cost of actual readiness.

This paradox, wherein organizational obsession with quantifying results corrupts them, is what historian Jerry Z. Muller has called “metric fixation.” The corruption in the case of readiness reporting takes many forms: the displacement of actual readiness with empty numbers, short termism among commanders and their staff, the collapse of innovation, the burning of endless man hours, and the hemorrhaging of job satisfaction. But to understand the scope of the harm, one must first understand the desired end (in this case, readiness) and the metrics used to measure it—the USR and its components.

The impetus to explore the USR’s shortcomings comes from my experience working twenty-four months as a troop executive officer. The following argument represents that single, tactical perspective on the problem, but I derive confidence in it from lengthy discussions and review with tactical and operational leaders across every type of BCT in multiple combatant commands. With uncanny unanimity and precision, leaders have echoed these concerns. This signals strongly to me that these issues are unfortunately not limited to a single formation.

The Anatomy of the Unit Status Report

In 2011, Congress established the readiness reporting requirement and defined readiness in the first paragraph of 10 U.S.C. § 117. Readiness, it says, is the ability of the Armed Forces to carry out the president’s National Security Strategy, the secretary of defense’s defense planning guidance, and the chairman of the Joint Chief’s National Military Strategy. Simply put, readiness is the capacity for the armed forces to fulfill assigned missions.

The U.S. code, having defined readiness, outlines how it ought to be reported. The language unambiguously requires discrete, quantitative metrics. Any system “shall measure in an objective, accurate, and timely manner.” The verb of choice in this sliver of code is “measure,” trotted out no fewer than seven times over two paragraphs.

The imperative to quantify readiness does not find a mandate in code alone. It also enjoys a vociferous booster in the Government Accountability Office.
(GAO). A 2016 GAO report typifies its argument for hard numbers and the tongue-clicking that ensues when results are insufficiently quantified: “The services have not fully established metrics that the department can use to oversee readiness rebuilding efforts and evaluate progress toward achieving the identified goals.” Testimony from the GAO in February 2020 sustains this tone, lauding the Department of Defense’s progress as it develops “metrics to assess progress toward readiness recovery goals that include quantifiable deliverables at specific milestones [emphasis added].”

In view of the above, Army Regulation (AR) 220-1, Army Unit Status Reporting and Force Registration—Consolidated Policies, endows the USR with an unsurprisingly quantitative structure. It comprises four measured areas: personnel (the P-level), equipment on-hand (the S-level), equipment readiness (the R-level), and the unit training proficiency (the T-level) (see figure 1, page 151; and figure 2, page 152). With the exception of the T-level, the same basic math governs all: divide what the reporting unit has (whether number of medics or number of serviceable grenade launchers) by what that unit ought to have. The ranking in each category is uniform and numeric: a level “1” (such as an R-1) indicates the highest readiness level in that measured area, and a “4” the lowest (such as R-4).

This discussion will focus on the question of equipment on-hand (the S-level) and equipment readiness (the R-level). It is in these measured areas where the USR is most rigid and quantitative, and it is where the metrics chosen least reflect the outcome that the report aspires to measure.

As mentioned, the math at face value is straightforward. The denominator for equipment on-hand is what the Army has decided the reporting unit must have, recorded in what is formally known as the modified table of organization and equipment (MTOE). The numerator is what appears on the unit’s property books; it is a digital record of equipment existent for that unit.

The denominator for equipment readiness is what is on hand, and the numerator is the quantity tracked as “fully mission capable” in the Army’s digital maintenance records. According to regulation, for a piece of equipment to be fully mission capable, it must pass a “preventative maintenance checks and services” inspection without failing a single “not ready if” bullet. The resultant percentage is often called the operational readiness rate, or OR rate.

This sanitized approach obfuscates the manipulation that can and does occur to ensure these basic fractions yield figures between .9 and 1.00. This warping of organizational behavior is the inevitability of Muller’s metric fixation.

The Unit Status Report as Metric Fixation

Muller explains metric fixation as the overreliance on transparent, quantified measurements to capture and incentivize an organization’s performance; it is also the persistence of this overreliance despite myriad negative consequences. Of the negative consequences Muller inventories, the USR most obviously induces the following in reporting battalions across the Army: goal displacement, short-termism, time burdens, innovation aversion, and degradation of work.

Goal displacement. If we take readiness to be the Army’s number one priority (or goal), then goal displacement is the most pernicious consequence of the USR as it definitionally displaces readiness. Robert K. Merton, a founding father of sociology, defined goal displacement as when “an instrumental value becomes a terminal value.” Professors W. Keith Warner and A. Eugene Havens elaborated in a seminal 1968 article that among goal displacement’s chief causes were “records and reports submitted to other echelons of the organization or to the sponsors, the public, or clients. These tend to report concrete ‘statistics,’ or case examples, rather than intangible achievement.” This academy-speak might translate into military-speak by simply saying that the USR makes the Army a self-licking ice cream cone.

Goal displacement in the measured area of equipment readiness (R-level) occurs as battalions grow more concerned with reporting equipment serviceable, such as vehicles, than with fixing equipment. The
perversion of maintenance that results is a familiar story to anyone who has worked in an Army motor pool.

A common illustration is as follows: broken vehicles are not marked as broken in the Army’s digital database (a process known as “deadlining”) until the unit’s maintenance section has diagnosed the issue with the vehicle and identified what parts must be ordered to fix it. Units delay reporting because it reduces the amount of time the vehicle is deadlined, thereby decreasing the likelihood it is deadlined during a USR reporting window. Because maintenance sections are often stretched for time, vehicles that cannot roll or start at all are reported for weeks if not months as serviceable simply because their issues have not yet been diagnosed.

Some units go even further to avoid an unbecoming R-level, displacing maintenance (and therefore readiness) in the process. A maintenance section in an armored formation, for example, might report only a single inoperable tank despite several others being broken. All repair parts for all tanks are then ordered under that single tank’s serial number. Upon receipt of the repair parts, the maintenance leadership divvies them up to the many other inoperable but unreported vehicles. This way, the digital database through which parts are ordered reports only one broken tank, instead of five or six per company. Not only does this produce an inaccurate report, but it also confuses maintenance. Leadership routinely forgets which widget was ordered for which unreported tank, resulting in redundant orders, lost parts, and inevitably, toothless tank companies.

Yet another painful example of goal displacement induced by USR involves what regulation calls “pacing items.” AR 220-1 defines pacing items (colloquially called “pacers”) as “major weapon systems, aircraft, and other equipment items that are central to the organization’s ability to perform its designated mission.” A

![Figure 1. Army Methodology for Overall Unit Readiness Assessments](https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/webir/220_1.pdf)
pacer for a medical unit might be a field litter ambulance; for a cavalry squadron, it might be its anti-tank missile systems and the vehicles on which they are mounted. The pacer OR rate is therefore in theory a reliable measurement of a unit’s ability to fulfill its mission, and it enjoys weight in the overall R-level calculus. Battalions, desirous of reporting themselves ready, consequently prioritize pacer maintenance.

The issue is that pacer OR rates are poor indicators of readiness and not just because serviceability rates lend themselves to manipulation. Pacers are also often far from the only equipment essential to fulfill a mission, or they are so numerous that each individual pacer has less impact on the mission than scarcer nonpacer equipment types. For example, a battalion may have twenty anti-tank vehicles, all of which are pacers, but only two command-and-control vehicles, neither of which are pacers. But because pacers enjoy disproportionate weight in the USR, any self-interested battalion prioritizes the maintenance of the twentieth pacer over the first command-and-control truck. Thus, command-and-control vehicles rust in the motor pool while twenty directionless anti-tank trucks roam the battlefield, but as far as the USR is concerned, the unit is combat ready. The goal of reporting a healthy pacer OR rate has displaced the goal of being ready.

Goal displacement abounds in the measured area of equipment on-hand (S-level) as well. Recall that S-level

![Figure 2. Commander's Unit Status Report Metrics](Figure from Army Regulation 220-1, Army Unit Status Reporting and Force Registration—Consolidated Policies, 15 April 2010, https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/webr220_1.pdf)
measures what equipment units have on hand against what the MTOE dictates they should have. In theory, MTOE captures all that a unit needs to fulfill its mission. But inevitably, well-meaning authors of MTOE at Training and Doctrine Command (TRADOC) as well as the approval authority at the Deputy Chief of Staff G-3/5/7 office either include too much or too few of any given item in view of the unit’s assigned mission.

In pursuit of a high S-level, units forsake actual equipment needs for a good report. These monthly campaigns see much-needed equipment transferred off the property books while supply teams bloat books with obsolete or unused equipment in order to meet MTOE quotas. Antiquated encryption tape readers remain while desperately needed high frequency radios or infrared optics disappear. The various forms of appeal, whether an “operational needs statement” or a “reclamation,” prove so cumbersome and lengthy that staffs rarely pursue them except in the direst cases. The goal of a high S-level displaces the goal of a well-equipped unit.

**Short-termism.** Related to goal displacement is short-termism. Muller defines short-termism in *The Tyranny of Metrics* as “diverting resources away from their best long-term uses to achieve measured short-term goals.” And because USR reports recur for battalions monthly, they disrupt long-term strategies for the maintenance, acquisition, and retention of equipment in pursuit of a good monthly read. This adverse effect of metric fixation runs precisely contrary to the stated 2018 National Defense Strategy, which emphasizes a pivot toward long-term readiness.

Examples are ubiquitous in the measured area of equipment readiness. Similar to the tank example above, battalion maintenance sections cannibalize long-suffering vehicles in order to repair newly downed pacers before the reporting windows close, resulting in what the aviation community calls “hangar queens”—sacrificial vehicles used as spare-part trees. This process cuts out the ordering of new parts altogether. Units do so both because of the quick turnaround (one need not wait for a part to arrive from a distant depot if one rips a part off of a neighboring truck), and also because if the maintenance section need not order the part, it need not report the truck as broken, which would spare the USR. Meanwhile, armament sections learn not to order parts for broken machine guns until after USR report-

Armament sections learn not to order parts for broken machine guns until after unit status report reporting windows close, delaying weapons repair by months to avoid flagging them as inoperable.

Examples are ubiquitous in the measured area of equipment readiness. Similar to the tank example above, battalion maintenance sections cannibalize long-suffering vehicles in order to repair newly downed pacers before the reporting windows close, resulting in what the aviation community calls “hangar queens”—sacrificial vehicles used as spare-part trees. This process cuts out the ordering of new parts altogether. Units do so both because of the quick turnaround (one need not wait for a part to arrive from a distant depot if one rips a part off of a neighboring truck), and also because if the maintenance section need not order the part, it need not report the truck as broken, which would spare the USR. Meanwhile, armament sections learn not to order parts for broken machine guns until after USR report-

Armament sections learn not to order parts for broken machine guns until after unit status report reporting windows close, delaying weapons repair by months to avoid flagging them as inoperable. Across all types of equipment, leadership rushes repair jobs or seeks out the easiest fix, undermining long-term serviceability and sometimes further damaging the equipment in the process. The result is an army of highly reactive, chaotic maintenance programs and duct-taped fleets.

Short-termism similarly dominates the measured area of equipment on-hand. Units dedicate time and effort to acquiring items they do not need in order to meet MTOE quotas, even with the knowledge that the obsolete equipment will fall off the MTOE the following fiscal year. Staffs will in turn direct battalions to give away needed equipment that will soon be on their MTOE simply because in that month the item is technically excess. Even worse, units will not turn in irreparably broken equipment (a process known as “coding out”) for fear that the loss will drop them below the MTOE-prescribed quantity, opting to retain unserviceable property and thereby precluding the fielding or even requisition of a functioning replacement.

Amidst all this short-termism, “recovery” becomes something of a four-letter word. To recover from training rotations requires the deliberate deadlining and coding out of equipment, processes that, for a host of good reasons, require time. Soldiers must inspect equipment, mechanics troubleshoot it, and clerks order repairs. Leaders must document catastrophic damage, officers investigate it, logisticians review it, and property book officers direct replacements. When handled
properly, issues identified during recovery take weeks if not months to resolve. It is the work of real readiness. But the price of that due diligence is at times one, if not several, unfavorable USRs, and units are too often unwilling to pay. The purpose of recovery becomes to report it complete, and all the while, units grow weaker.

**Innovation aversion, time burdens, and degradation of work.** In a series of articles this past summer, Gen. Stephen Townsend and his three coauthors called for a reinvigoration of mission command, the Army’s allegedly faltering approach to command and control. To do so, they wrote that leaders must appreciate that “developing competence, establishing mutual trust, and learning to operate from shared understanding does not start in the field. It starts in the unit area.” In doing so, they echoed the chairman of the Joint Chiefs of Staff, Gen. Mark Milley, who in 2017 called for mission command’s practice “even on daily administrative tasks you have to do in a unit area.” Unfortunately, the USR, perhaps the Army’s most quotidian administrative garrison task, plays something of a perfect foil to mission command. Its metric fixation asphyxiates several of mission command’s core tenets: disciplined initiative, risk acceptance, mutual trust, and shared understanding.

Most obviously, the fragility and frequency of the USR discourages innovation, or “disciplined initiative,” and its twin, “risk acceptance,” that might otherwise increase readiness. The dearth of innovation at the top of the Army’s food chain has received due attention, perhaps most famously from former Lt. Col. Paul Yingling in a 2007 article. But metric fixation so deadens innovation at the tactical level that it is no surprise little rises to the level of strategy. Untested methods, whether a change to motor pool management or an alternate approach to equipment distribution, enjoy a slim chance of fruition as they threaten USR calculus month to month.

**Figure 3. Company Leaders’ Estimates of Personal Time Devoted Per Quarter to Job Tasks**

Would-be innovators are told instead to wait until their career’s distant future when, if they perform well enough, they might enjoy influence over the stratospheric decisions that inform doctrinal questions, MTOE, USR, or otherwise. Though there is much to be said for earning one’s place, ideas expire with time, and many exit the profession of arms before entering positions of influence in search of a more enterprising culture.
The unit status report, perhaps the Army’s most quotidian administrative garrison task, plays something of a perfect foil to mission command. Its metric fixation asphyxiates several of mission command’s core tenets: disciplined initiative, risk acceptance, mutual trust, and shared understanding.

RAND study that they devoted a full 15 percent of their time to “tracking readiness,” second only to USR-adjacent “equipment maintenance and accountability.” Both outstripped the 13 percent of each quarter commanders professed dedicating to “unit-specific training.” Ironically, soldiers shared that a common means of coping with the time burden was to report readiness metrics inaccurately (see figure 3, page 154). This spells doom for mission command’s “shared understanding,” as staffs and commanders dedicate to data’s collection and grooming the attention that mission orders desperately need.

Lastly, least measurable (and therefore, from a metric-intensive perspective, least credible) but just as tragic is what Muller calls the degradation of the work. Implied in the hyperquantification and rigidity of the USR is an organizational distrust of the reporting unit, and therefore the soldiers who constitute it. This distrust is not lost on those soldiers, and it invites them to respond in kind. “Mutual trust” fails. The bedrock of Army morale—the nobility of its mission—crumbles as the mission is reduced to a series of reported fractions.

The Blame Game

Self-righteous blame invites obstinate defense, and both are obstacles to productive discussion. As Leonard Wong and Stephen Gerras wrote in the 2015 report Lying to Ourselves: Dishonesty in the Army Profession (from which this piece draws much), “with such a strong self-image and the reinforcing perspective of a mostly adoring American society,” Army leaders often “respond with indignation at any whiff of deceit.” Discussions thus falter before they begin as all retreat to their respective corners.

Fortunately, organizational theory bypasses these obstacles convincingly. It is not because of the individual but rather because of the devaluation of the individual that such perversions of organizational behavior occur. The system of readiness reporting dismisses individual judgment in favor of metrics so much that all agency, informed by integrity or any other Army value, dissipates. As George Kennan wrote in 1958 when discussing the expanding administrative state and its managerial malaise,

> The premium of the individual employee will continue to lie not in boldness, not in individuality, not in imagination, but rather in the cultivation of that nice mixture of noncontroversialness and colorless semicompetence that corresponds most aptly to the various banal distinctions of which, alone, the business machine is capable.

So, instead of stacking structural incentives impossibly high, diametrically opposing the integrity of the individual, and then blaming the individual for systemic failure, the resolution lies in structural reform. This approach enjoys the dual advantages of preempting the defensiveness Wong and Gerras encountered and more credibly promising results.

**Recommendations**

To critique metric fixation is not, as Muller repeatedly disclaims, to protest the use of metrics altogether. Similarly, to decry the pernicious effects of the USR is not to deny the need for readiness reporting and the use of metrics toward that end. The reform, not the scrapping, of reporting metrics and structure, promises

The USR and its pruning voraciously consumes another resource that serves a battalion’s mission: leaders’ and soldiers’ time. The frequency and high stakes of USRs demand of battalion and brigade staffs days of data compilation that might otherwise be spent planning training. Company commanders reported in a 2019
a reduction in goal displacement, short-termism, innovation aversion, time burdens, and degradation of work. This author acknowledges that the below recommendations are not equally feasible, and if executed improperly, fail to resolve the excesses of metric fixation.

To reduce goal displacement, one must close the gap between the stated goal of readiness and the metrics used to measure it. As the metrics employed by USR gravitate closer to actual drivers of readiness, the risk of the former displacing the latter would necessarily decrease. The massive effort units expend to reach the highest levels of readiness on USRs would therefore more efficiently ready them. A first step toward this end would be to better incorporate the judgment of reporting leadership, those closest to the capabilities of their formations. Rather than empowering them to “subjectively upgrade” overall readiness ratings (as AR 220-1 does now), which obscures rather than resolves metric fixation, reporting units ought have a larger role in the selection of what metrics capture readiness on the ground.

Pacer designation is an example. Currently, TRADOC, in coordination with the Deputy Chief of Staff G-3/5/7 office, identifies pacers and accounts little for the nuanced relationship between equipment and units’ assigned missions. This is understandable given the size of the force and the degrees of separation between everyday training and TRADOC. To close this gap, a regular (perhaps biennial) reassessment that solicits division or even brigade input regarding what ought to be considered a pacer would make pacer OR rates more meaningful. The same practice might be employed to adjust the MTOE. Just as reporting units have unique insight into what equipment most contributes to their mission in the case of pacers, so too do they have a strong understanding of what type and quantity of equipment they use to fulfill their missions. Permitting divisions or brigades some role in the authorship of their MTOEs would better marry MTOE materiel with the needs of the unit.

One risk of such a practice would be mission creep. As units and their commanders acquire more influence over what the Army deems essential, they may functionally invent mission essential tasks to warrant desired widgets, bringing at times anomalous personal experience in contest with doctrine. The Army would thus have to maintain a high but passable bar for what equipment supports only *existing* mission essential tasks.

A second obvious objection to unit partial authorship of either MTOE or pacer designation might read as follows: every unit setting its own standard reduces the term “ready” to something just shy of meaningless as each unit proffers its own (perhaps self-serving) definition. The resultant amalgam of definitions cripples the military bureaucracy’s ability to manage. Only strict standardization renders the force legible, whether to the Pentagon or to Congress.

How to negotiate a balance between the dual risks of harmful standardization and unmanageable chaos is explored deeply in the book *Seeing Like a State* by James C. Scott. In it, Scott relays among many examples the challenge Napoleonic France faced as it sought to standardize myriad local measurement codes: “Either the state risked making large and potentially damaging miscalculations about local conditions, or it relied heavily on the advice of local trackers—the nobles and clergy in the Crown’s confidence—who, in turn, were not slow to take full advantage of their power.” Scott notes attempts to strike the balance, such as those by Deputé Claude-Joseph Lalouette, failed to win requisite support for fear of too empowering the landowners. This concern does not apply to the question of readiness reform, for instead of thousands of landowners with ulterior motives, the Army needs to only solicit input of several dozen BCTs supportive of its mission.

Decreasing the frequency of USRs to a biannual or even annual iterations would also assuage many of its ill effects. There is no great advantage to monthly reports but many costs, only some of which have been discussed. Muller has summarized the damage done by quarterly earnings “hysteria” to long-term strategy in the financial sector, and the same basic critique applies to the Army. Less frequent reports would permit units some actual recovery periods between training events without the disincentive of ugly USR reports. Less frequent reports would reduce the pressure on leaders to prioritize readiness metrics over deliberate training progressions. Those leaders would certainly tolerate more programmatic maintenance.

Lastly, lengthening the periods that commanders command to thirty-six months or longer has the potential to preempt the short-termism USRs engender. Often, under pressure to produce short-term results, commanders undermine or outright dismantle systems designed to sustain readiness in the long view because
those systems do not move at the speed of the USR. This practice survives because few commanders command long enough to reckon with the fall out of this behavior. Extending command timelines would force a consideration of long-term effects that are otherwise a problem for the anonymous successor.

The above temporal fixes reduce short-termism. A less frequent USR disrupts long-term planning less frequently. A leader with more time in the driver's seat similarly plans for the longer term. And all of the above empower leaders and soldiers within reporting units. The time burden shrinks as reporting grows less frequent. The risk of innovation lessens, and innovation's long-term benefits assert themselves. As the organization solicits input and metrics of performance acquires meaning, work regains its esteem and morale increases. Put another way, it promises a reinvigoration of mission command.

The need for quantifying readiness will never go away, nor should it. The accessibility of hard numbers and their simplicity render the military's sprawling bureaucracy manageable. It importantly also reduces the opacity of the military to oversight entities like the House Armed Services Committee's Subcommittee on Readiness. But unless the USR undergoes reform, it will neither ready us nor convey how ready we are, to the public or ourselves.

Notes


4. Ibid., § 117(a).


8. Ibid., para. 9-3.

9. Ibid., para. 9-4.

10. Muller, The Tyranny of Metrics, 175.


13. AR 220-1, Army Unit Status Reporting, 99.

14. Muller, The Tyranny of Metrics, 149.


22. AR 220-1, Army Unit Status Reporting, para. 4-5.


24. Ibid.