STAT COE-Report-09-2021

Best Practice for Effective Requirements and Measures for Autonomous Defense Systems

Authored by: Charlie Middleton

30 June 2021

The goal of the STAT COE is to assist in developing rigorous, defensible test strategies to more effectively quantify and characterize system performance and provide information that reduces risk. COE products are available at www.afit.edu/STAT.

STAT Center of Excellence
2950 Hobson Way – Wright-Patterson AFB, OH 45433

Table of Contents

Executive Summary

The development of autonomous defense systems is increasing throughout the Services in the Department of Defense. These systems can be stand-alone or integrated subsystems that can respond to situations not fully anticipated or tested, with a degree of self-directed behavior, and may or may not have a human operator involved during mission operations.

Effective requirements for autonomous defense systems are desired, to aid in development and fielding of quality systems that meet the user needs, on schedule and at an acceptable cost. Having inadequately defined requirements can result in scope creep, program delays, cost overruns, and poor product quality that does not meet customer needs and safety requirements.

For both capability requirements such as in the Capability Development Document, and performance specifications written into a developer's contract, effective system-specific requirements should be specific, verifiable, clear, concise, accurate, feasible, realistic, necessary, consistent, not redundant, and explicit. Challenges of autonomy requirements arise from the use of autonomy technology to replace the historical role of a human operator, removing the operator's experience from the control of the system, causing the need for judgment and tactics for system employment to become part of the system itself. Examples of effective autonomous defense system requirements that exhibit the effective characteristics are discussed to illustrate how these challenges can be overcome.

Measures are parameters or metrics of quantitative assessment used for measurement, comparison or to track performance or production. Measures are used to assess system characteristics and capabilities such as speed, range, power, lethality, survivability, availability, reliability, and more. Categories of measures used for DoD systems are measures of performance, measures of effectiveness, and measures of suitability. Effective measures allow the program manager and other stakeholders to know whether the autonomous system is meeting performance and mission effectiveness requirements, whether cost and schedule objectives can be met, and whether the system requires design or integration changes in its development. Measures should be relevant, specific, easily obtained, clear, precise, repeatable, standardized, and traceable. Challenges of autonomy measures arise because for requirements like perception, reasoning, deciding, planning, and coordinating, no historical measures exist for system capability measurement. Examples of effective autonomous defense system measures that exhibit these effective characteristics are discussed to illuminate ways to overcome these challenges.

Future research will expand upon specific test and evaluation methods, techniques, best practices and lessons learned for applying effective requirements and measures to address the challenges of the development of autonomous defense systems.

Definition of Autonomous Defense System

There are various types of systems that come to mind when one hears "autonomous system." One may consider an automatically operating system as autonomous, although their operations are deterministic, based on very predicable inputs and simple programming, such as:

- A coffee pot that automatically turns on and makes coffee according to a schedule
- A military radar that scans for targets according to a schedule

More complicated systems (or subsystems) may use some sensor or other perception of conditions to dictate their operations, such as:

- A home thermostat that manages the air conditioner and heater based on room temperature
- A military defensive countermeasure suite that detects electronic threat signals and tracking, to manage an active electronic jammer and chaff dispensing

As system complexity increases, the system may itself have to "decide" what is or is not valid criteria for its operations, such as:

- A self-driving car deciding whether there is too much snow on the road to continue safely driving down it
- An unmanned military submarine deciding whether an approaching object is an enemy submarine or a pod of whales

Autonomous systems may have dangers or costs that result from incorrect or faulty operations:

- A package-delivery drone could crash into a bus, killing dozens of people
- A military counter-mortar fire system could rapidly deplete all its ammunition erroneously firing at flocks of birds

Autonomous systems may have a human operator or monitor able to intervene if automatic operations are incorrect or faulty:

- A pilot can disable the electronic jammer if notified that it is responding to spurious threats
- A remote technician can command a vehicle stop and park when the weather gets too rough

Ultimately, the most autonomous of systems may be able to someday perform human-like functions such as perceiving, reasoning, planning, deciding, learning, and coordinating among a larger group:

- Fleet of military air vehicles that determines its own routes to penetrate enemy air defenses using real-time threat assessment, chooses its own valid targets based on real-time sensors, employs munitions according to a priority scheme, conducts battle damage assessments, and executes re-attacks if targets are not sufficiently destroyed, all without any human interaction due to long-range communications jamming (notional 22nd-century system?)

For the purposes of this paper, an **<u>autonomous defense system</u>** is a stand-alone system, or integrated subsystem of another system, that [1]:

- Has the capability to operate and respond in situations that were not fully anticipated or tested, possibly with degraded performance
- Has some degree of self-governance and self-directed behavior
- Has dangers or costs that result from incorrect or faulty operations
- May or may not have a human involved during its operations in a mission or task

Features that an autonomous defense system have may include perception, reasoning, planning, deciding, learning, and coordinating. These features may be implemented through the use of technologies such as computer vision, speech recognition, sensor fusion, inference engines, machine learning, artificial intelligence, optimization tools, and more. Future papers will expand and explore these particular features and applications, as well as methods to evaluate their effectiveness.

Definition of a Requirement

The term “requirement” is used to mean several different things depending on context – different places in the acquisition cycle, by the Joint Requirements Oversight Council (JROC), the program manager, the vendor, the tester, and the user. Dictionary [2] definitions of the word “requirement” vary, similar to:

- <u>Requirement</u>: Something wanted or needed

The dictionary definitions are too vague to tell us what a requirement for an autonomous defense system needs to include to be a good requirement. Everyone would like for requirements to be thoughtfully designed and well-written, so that a quality product is delivered on time, performs as expected, and can be depended on to perform when needed, at an affordable cost [3]. Throughout the defense acquisition system, requirements are used in different ways depending on the purpose of the document. No single DoD definition exists for the word “requirement.” At the highest level, the Joint Capabilities Integration and Development System (JCIDS) is the process that defines the requirements for future defense programs[4]. DoD strategic guidance provides requirements for the joint force capabilities needed for the future, such as[5]:

- Provide a fire support coordination capability that integrates all fires, including cyber

This level of broad, general military requirement is above the level of this paper. Likewise, the Initial Capabilities Document (ICD) requirements are high-level capabilities needs that may be met by a variety of solutions as opposed to a single system [4].

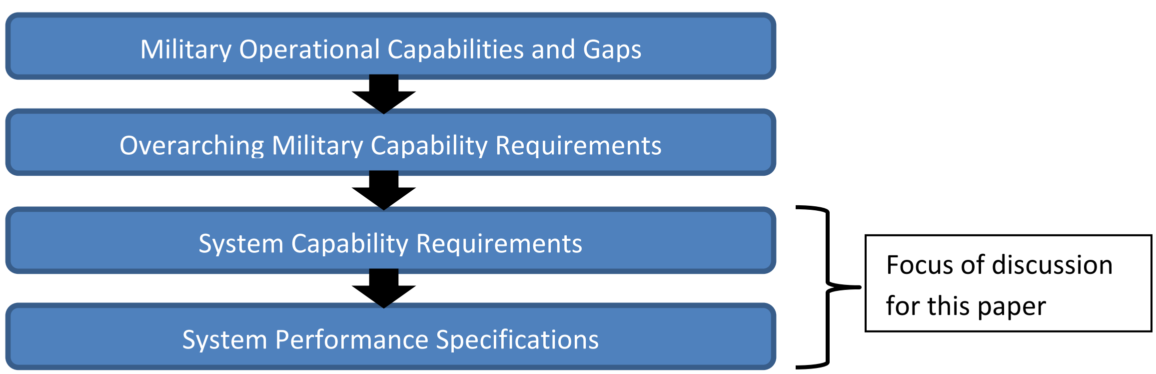
The primary specific system-level requirements documentation occurs in the Capability Development Document (CDD), which proposes development of a specific materiel capability solution intended to satisfy validated capability requirements [4]. The CDD is system specific and applies to a single increment of capability in an evolutionary acquisition program [6].

The CDD is typically used to inform the writing of a specification by the system program office. Specifications are written system requirements documents used that should conform to the definitions, content and format of MIL-STD-961, which defines a specification [7]:

- Specification: A document prepared to support acquisition that describes essential technical requirements for materiel and the criteria for determining whether those requirements are met.

A system specification document is typically used to contract with the system developers and contains various types of specifications. The specification of interest for this paper is the performance specifications; detail specifications are intended primarily as an exact description of the item to be produced, which is not currently very applicable to developing autonomous systems at this time.

- Performance specification [7]. A specification that states requirements in terms of the required results with criteria for verifying compliance, but without stating the methods for achieving the required results. A performance specification defines the functional requirements for the item, the environment in which it must operate, and interface, interoperability, or compatibility characteristics. It does not present a preconceived solution to a requirement.
- Detail specification [7]. A specification that specifies design requirements, such as materials to be used, how a requirement is to be achieved, or how an item is to be fabricated or constructed.



Generally, DoD prefers performance-based specifications to detail specifications to state requirements in performance terms and give contractors the flexibility to provide innovative, technologically advanced, best-value solutions to meet the customer's requirement [7]. The requirements of concern for the remaining discussion concern the specific requirements developed as capability requirements for the specific system, or as performance specifications for the specific system. For the purposes of this paper, an **autonomous defense system requirement** is a written statement that:

- Describes technical needs for a specific autonomous defense system
- Provides these needs as either a capability requirement, or a performance specification
    - Capability requirement: Written by the government stakeholders as a system capability need, focused on the desired outcomes and effects needed

- Performance specification: Written by either the government or the system developer, derived from a capability requirement, as a system need that includes functional, environment, interface, interoperability, or compatibility needs with criteria for verification

Examples of autonomous defense system requirements discussed in the following sections.

Characteristics of Effective Autonomy Requirements

Effective requirements for autonomous defense systems are desired, to aid in development and fielding of quality systems that meet the user needs, on schedule and at an acceptable cost. The writing of effective requirements is not an easy task, as requirements definition is the most important and most difficult part of some DoD acquisitions [6]. Furthermore, issues in requirements management are often cited as major causes of project failures. Having inadequately defined requirements can result in scope creep, program delays, cost overruns, and poor product quality that does not meet customer needs and safety requirements [8]. To understand the characteristics of an effective requirement for an autonomous defense system, a perspective on the desired characteristics of requirements, in general, can provide useful insights.

Effective system requirements and specifications generally have the following characteristics [8, 9, 10]:

- Specific: The requirement is worded to provide, for one single capability or characteristic only, a definitive basis for acceptance or rejection
- Verifiable: An objective verification can be defined to determine if the requirement is met; can be accomplished by analysis, demonstration, examination, or testing; also known as Testable
- Clear and concise: The requirement is stated in terms that are unambiguous (can only be interpreted one way no matter who reads it), appropriate to the level of abstraction of the system, no more detailed than necessary
- Accurate: The requirement sufficiently describes the needed capability without the need of additional information, including the ranges (tolerances) of acceptable performance in the correct units; using a standardized format aids in this
- Feasible and realistic: The requirement is technically achievable within acceptable cost, schedule, legal, ethical, security, and safety constraints
- Necessary: The requirement defines an essential capability, that if not included causes a deficiency that affects higher-level needs and requirements; allows Traceability of the requirement to higher-level needs
- Consistent and not redundant: The requirement does not conflict with other requirements, the interface and compatibility needs are stated in uniform units, its words have the same meaning when used repeatedly, and it is essential in the presence of all other requirements
- Explicit: The requirement is stated with the operational or environmental conditions that apply; includes the criteria for when the requirement does or does not apply

The definitions of any terms used in the requirements should be documented for consistency and clarity [9]. Published guidance should be referenced for other details about requirements or specifications format (JCIDS Manual, MIL-STD-961). Each capability requirement should stipulate the required operational attributes with appropriate quantitative parameters and metrics, (outcome, time, distance, effect, supportability, or other characteristic), however, capability requirements have slightly more freedom to allow flexibility in potential capability solutions[4]. Capability requirements, for example, may state that a system must detect 95% of expected threats for 10 years in the future; this requirement is not completely specific nor explicit, yet it may be an effective capability requirement, so long as the expected threats can be defined in other documentation. The performance specification, however, should be completely specific and explicit, stating the exact threats expected to be detected with the expected parameters for each; this provides accountability of the contractor in meeting the requirement.

The benefit of having effective requirements is the improved quality of the system development through having clear, traceable, verifiable, consistent, accurate, specific and realistic capability needs by allowing the program office to manage, and the contractor to design and build the system to most precisely meet the user needs at optimum efficiency.

Challenges of Autonomy Requirements

Autonomy technology presents many new challenges to the development, test, and evaluation of systems[11,12]. These challenges generally arise from the use of autonomy technology to replace the role of a human operator. Historically, defense systems are designed and used with a solder, sailor, airman, or marine fighting the battles, with the systems acting to augment the warfighter's speed, range, firepower, perception, protection, or other ability necessary to fight and win wars. The assumption was that a human warfighter controlled and employed the system, using their experienced judgment into combat tactics required.

Autonomous defense systems flip this entire paradigm on its head, allowing systems to execute combat tactics with no operator actively controlling the system. This paradigm shift removes the operator's experience from the control of the system, causing the need for judgment and tactics for system employment to become part of the system itself.

These challenges apply to the system requirements both directly and indirectly, and arise in several categories. Application of the characteristics of effective requirements aid in uncovering particular concerns regarding autonomy requirements [11]:

- **Perception**: How does the system acquire information about the current state of its position, enemy and friendly forces, environment, neutral or non-participating entities, threats, obstacles, and even degradations of perception or deceptions?
    - How <u>specific</u> and <u>accurate</u> do the autonomous system's perceptions need to be? Infinite possibilities exist for the future environment, threats, and deceptions, so how can requirements possibly provide accurate and specific perception needs?

- **Reasoning**: How does the system conceptualize is perceptions to determine the truth about its reality and draw conclusions? How does the system apply inductive, deductive, analogical, or abductive reasoning to reach its conclusions? What pre-conceived rules, knowledge, or training data is necessary to provide a sufficient model of the world that allows the system to draw correct conclusions?
    - How can the reasoning be accurate and verifiable? In the presence of infinite possible situations, how can the system's reasoning requirements be feasible and realistic? Can the user explicitly state all the conditions that apply to the autonomous system's reasoning and conclusions?
- **Planning / Deciding**: How does the system decide on a response or course of action, with competing priorities and uncertain information about its world? What plan or decision is required for the autonomous system to choose when possible situations and choices are combinatorially intractable?
    - How can the planning and decision tasks be explicitly stated, realistic, and verifiable? How can effective decisions be accurately stated?
- **Learning**: How does the system build on its experience, including successes, failures, and uncertainties, to provide more effective outcomes in the future?
    - How can learning requirements be stated realistically, explicitly, and accurately?
- **Coordinating**: How does the system coordinate and cooperate with other systems to achieve mission objectives? Can the autonomous system adjust its performance to synergize with other friendly systems, assets, and forces?
    - How can coordination requirements be accurately and explicitly stated? How is effective coordination verifiable considering limitless possible scenarios?

In addition to these challenges, another fundamental challenge exists of high interest to defense systems in particular: how should the autonomous system adapt its performance capabilities to counter a **learning, thinking adversary**? Enemy threats, tactics, and behaviors are not static nor predictable, so how can the autonomous defense system perform effectively in an evolving battlefield? How can an effective requirement be defined to ensure the system is flexible and adapts to provide useful capabilities throughout the full range of military operations?

These challenges only begin to uncover the difficulties in defining effective requirements for autonomous defense systems. Management of autonomy requirements will likely prove a critical role enabling the success (or causing the failure) of future autonomous defense systems.

Examples of Effective Autonomy Requirements

An understanding of the characteristics of effective requirements, along with an understanding the challenges of autonomous defense system requirements, provides the basis for beginning to define effective requirements for autonomous defense systems. Some examples of the application of the desired characteristics to autonomous systems may provide useful insight into future requirements definitions.

Consider a (notional) autonomous ground combat vehicle that is intended to provide screening for a manned land maneuver force. The autonomous vehicle needs to detect threats such as mines, adversary armor and artillery, and adversary air vehicles, to provide threat warning and electronic warfare capabilities to degrade these threats. As part of the autonomous system's concept of operations (CONOPS), it needs to safely move to maintain formation with the manned force and change formation positions based on its perception of threat and friendly force movements.

An example for an ineffective autonomy requirement for the system's coordination capability to maintain correct formation position might be:

- Formation: The autonomous system shall maneuver to and maintain assigned formation position for 95% of maneuver operations.

While this requirement at first glance may seem sufficient, it does not meet most of the characteristics of effective requirements; it lacks specific, accurate, verifiable, unambiguous, and explicit details. This requirement may be acceptable for high-level requirements such as an Initial Capabilities Document; however, for system requirements (CDD and performance specification) this statement is ineffective. Applying the characteristics to the requirement in a disciplined approach enables the definition of a more effective autonomy requirement:

- Formation: The autonomous vehicle shall move to a position between 400 meters and 800 meters lateral spacing, and between 600 meters and 1200 meters longitudinal spacing, from the nearest friendly maneuver vehicle.

Characteristics of Effective Requirements

Specific

Verifiable

Clear and concise

Accurate

Feasible and realistic

Necessary

Consistent and not redundant

Explicit

  - The autonomous vehicle shall begin movement within 10 seconds of movement of the nearest friendly maneuver vehicle. The autonomous vehicle shall stop within 10 seconds of the stopping of the nearest friendly maneuver vehicle.
  - The autonomous vehicle shall move at the same velocity as the nearest friendly maneuver vehicle, within 5 m/sec and within 10 degrees direction.
  - If the autonomous vehicle employs electronic warfare capabilities against a threat, the vehicle shall move to align with the optimum threat suppression axis (*provided elsewhere) while maintaining the position boundaries stated above
  - The autonomous vehicle shall comply with maneuver and obstacle safety requirements (*provided elsewhere) while moving to formation position.
  - If not in formation position, the autonomous vehicle shall begin movement into position within 10 seconds, at greater than 10 m/sec velocity.
  - If not in formation position, the autonomous vehicle shall provide a communications message warning of its formation position, with the reason for being out of position.
  - If the nearest friendly vehicle position is not available, the autonomous system shall provide a communications message warning of its position.

The autonomous system requirements as stated in this example provide specific, verifiable, clear, concise, and explicit direction for what the required system performance is and allows for corrective action and warnings when performance is not met. This provides the system developer with an unambiguous requirement that removes any uncertainty as to what the user really wants the system to do to maintain its formation position.

Another example of an ineffective autonomy requirement for the system's decision capability for prioritizing electronic warfare employment might be:

- Electronic warfare prioritization: The autonomous system shall employ adequate electromagnetic jamming on priority threats while minimizing electromagnetic power transmitted.

This requirement is not accurate, specific, verifiable, clear, or consistent. Words like "adequate" have no objective criteria for evaluation, likewise "priority" and "minimizing" are not specific and verifiable (unless definitions are provided elsewhere). No decision acceptability criteria are given in this statement as to how the tradeoff between adequate jamming and minimized power transmission should be balanced.

Again, applying the characteristics of effective requirements provides for the definition of a more effective autonomy requirement for the autonomous defense system to prioritize electromagnetic jamming employment:

- Electronic warfare prioritization: The autonomous systems shall employ electromagnetic jamming when friendly forces are within 130% of stated threat range, on at least four simultaneous threats according to the threat priority order given in *Table XX*.

- The autonomous system shall calculate the threat range to each friendly asset position in the formation.
- The autonomous system shall prioritize employment on threats within 100% of stated threat range, over threats at greater than 100% of threat range, irrespective of Table XX threat priority order.
- The autonomous system shall update prioritization to comply with threat employment directives received from friendly forces, irrespective of the prior priority order.
- The autonomous system jamming direction shall be centered within 10 degrees of azimuth and elevation of the threat position.
- The autonomous system jamming beamwidth shall be less than 25 degrees and greater than 15 degrees.
- The autonomous system jamming power shall produce greater than 3 to 1, but less than 5 to 1, signal-to-noise ratio against the highest priority threat.
- The autonomous system jamming power shall produce greater than 2 to 1, but less than 4 to 1, signal-to-noise ratio against the second, third, and fourth priority threats.
- The autonomous system transmit power shall be less than 0.5 Watts transmitted when no threats on *Table XX* are within 130% of their stated threat range.

This example provides a more effective autonomy requirement as the specific, accurate, verifiable, unambiguous terms and metrics included allow clarity in the autonomous system's desired performance. The test and evaluation team should have a much easier task of designing tests to verify and validate the autonomous defense system electronic warfare prioritization with the more effective requirement, as opposed to the first, more ambiguous version.

To aid in the challenging task of writing and defining autonomous defense system requirements that meet the characteristics of effective requirements, requirement templates or patterns should be used as a best practice. Templates, tools, and software solutions are available to aid in defining and managing requirements. Templates are most often used in the field of requirements management to refer to the structure and organization of a requirements document. The use of such templates helps ensure that the authors consider the complete range of concerns when organizing requirement statements [13].

The use of patterns and templates also enables the development of natural language processing tools or digital assistants. These tools, along with aiding writers when writing need and requirement statements, helping ensure the rule and characteristics identified here are met [13]. An added benefit of requirements patterns or templates is the use of requirements parsing software tools, which can examine written requirements and transform them into logical code for easier checking for consistency and completeness. Use of these requirements parsing tools is an expected future topic for best practice research on autonomous defense system T&E.

Once the autonomous defense system requirements are defined, the test team can develop measures for verification and validation that the requirements have been met.

Definition of a Measure

The term "measure" is defined in the dictionary[14] for the purposes of test and evaluation:

- Measure: a way of judging something

Measures are used in DoD test and evaluation to provide the criteria for how performance and characteristics of systems are quantitatively assessed. A synonym for measures used at times is "metrics," which the DAU Glossary [6] defines as "Parameters or measures of quantitative assessment used for measurement, comparison or to track performance or production." Another term used commonly is "technical performance measure," which applies primarily in a systems engineering construct. In the DoD JCIDS process, the Capabilities Development Document (CDD) introduces the term Key Performance Parameter [4].

Key Performance Parameters (KPPs) are performance attributes of a system considered critical or essential to the development of an effective military capability. KPPs are contained in the CDD. KPPs are expressed in term of parameters which reflect Measures of Performance (MOPs) using a threshold/objective format. KPPs must be measurable, testable, and support efficient and effective test and evaluation [6].

While the KPPs provide some measures useful in evaluating autonomous defense systems, other measures are needed during the development, integration, and evaluation of systems. Measures are used to assess characteristics such as [15, 16, 17]:

- Reliability
- Power required
- Weight
- Throughput
- Human factors
- Response time
- Complexity
- Availability
- Accuracy
- Speed
- Integration
- Lethality
- Survivability
- Maintainability
- System security

Well-defined measures can be extremely useful to the program manager and other stakeholders: they can provide predicted values to be achieved in technical performance, can provide visibility of actual versus planned performance, can provide early detection of problems requiring attention, and can support the assessment of the impacts of proposed system changes [17].

In defense acquisitions, three categories of measures are commonly used [6].

- Measure of Performance (MOP): System-particular performance parameters such as speed, payload, range, time-on-station, frequency, or other distinctly quantifiable performance features. Several MOPs may be related to achieving a particular Measure of Effectiveness (MOE).
- Measure of Effectiveness (MOE): The data used to measure the military effect (mission accomplishment) that comes from using the system in its expected environment. That environment includes the system under test and all interrelated systems, that is, the planned or expected environment in terms of weapons, sensors, command and control, and platforms, as appropriate, needed to accomplish an end-to-end mission in combat.
- Measure of Suitability (MOS): Measure of an item's ability to be supported in its intended operational environment. MOSs typically relate to readiness or operational availability and, hence, reliability, maintainability, and the item's support structure.

These three categories of measures are applicable to autonomous defense systems, just as to other military systems. However, defining effective measures can be a difficult task, especially for a new technology like autonomy, similar to the difficulties in defining effective requirements. The following discussions provide characteristics of effective measures, challenges of autonomy measures, and examples of effective autonomous defense system measures.

Characteristics of Effective Measures

Effective measures provide great usefulness to the program manager and other stakeholders during autonomous system development, integration, testing and evaluation. Effective measures allow the program manager and other stakeholders to know whether the autonomous system is meeting performance and mission effectiveness requirements, whether cost and schedule objectives can be met, and whether the system requires design or integration changes in its development. Effective measures also allow testers and users to gain confidence in the autonomous system's abilities to meet capability requirements and provide military utility in an operational context. Without effective measures, testers and managers have a difficult time determining whether the system requirements are being met and whether the system will provide reliable, safe, secure, maintainable operational utility to the warfighter.

Effective measures, in general, should have the following characteristics to provide useful insights into system capabilities:

- Relevant: The measure provides an important factual truth about the capability
- Specific: The measure assesses one single capability or characteristic only
- Easily obtained: The measure is directly observable from test, examination, analysis, or demonstration without complex transformations or calculations
- Clear: The measure is stated in terms that are unambiguous as to how the value is determined
- Precise: The measure provides significant granularity to assess small but meaningful changes in performance or effectiveness

- Repeatable: The measure provides the same information with the same quality every time it is assessed
- Standardized: The measure is usable and effective for similar systems with similar capabilities
- Traceable: The measure assesses definitively the acceptance or failure of an actual system requirement or performance specification

The assessments of capabilities using the application of measures are best communicated using graphs, charts, diagrams, or other figures. These allow visualization of the quantities involved for performance, thresholds, and changes over time. Generally, effective measures provide a high level of information content when applied correctly. Effective measures communicate relevant, specific, precise, standardized and repeatable facts about the system, and empower the tester to draw meaningful conclusions about system performance and effectiveness.

Challenges of Autonomy Measures

Autonomy technology presents new challenges to the development of effective measures, much like the challenges to the definitions of effective requirements [18]. Historically, many measures used in system evaluations were based on physical characteristics, such as weight, speed, range, and explosive force. These measures were easily understood and straightforward to obtain.

For requirements like perception, reasoning, deciding, planning, and coordinating, no historical measures exist for system capability measurement. Or rather, those measures that exist were developed for measuring human performance, not system performance, as autonomous defense systems are intended to perform tasks that human operators historically have done.

Testers therefore have an extremely difficult task ahead, of defining effective measures for autonomous defense systems. This task will be made easier if effective requirements are provided, as discussed earlier. Effective requirements should already be verifiable, relevant, and clear. These qualities will greatly aid in developing effective autonomy measures, as effective requirements should trace directly to unambiguous measures that can verify the capability. The larger challenge for testers will arise when requirements are not effective, but ambiguous, inconsistent, or not explicit. In this event, testers may need to collaborate with stakeholders to derive more effective requirements that can then trace to effective autonomy measures.

Examples of Effective Autonomy Measures

An understanding of the characteristics of effective measures, along with an understanding the challenges of autonomous defense system capability measures, provides the basis for beginning to define effective measures for autonomous defense system performance and effectiveness. Some examples of the application of the desired characteristics to autonomous systems may provide useful insight into the development of autonomy measures.

Characteristics of Effective Measures

Relevant

Specific

Easily obtained

Clear

Precise

Repeatable

Standardized

Traceable

Consider the example of a (notional) autonomous ground combat vehicle that is intended to provide screening for a manned land maneuver force. This vehicle has a requirement to move and maintain a formation position in relation to the other friendly ground forces in the operation. One possible measure of performance (MOP) is:

- Formation position MOP: The percent of total mission time that the autonomous vehicle is in the required formation position.

This measure is ineffective, as it is not precise, as the data from this measure will give no information as to how far out of formation position the autonomous vehicle was. Nor is the measure easily obtained, as it requires computation of where the vehicle should have been throughout the operation, before it can be determined whether it was in that position or not. The measure also lacks relevance, as to why the vehicle was out of position.

More effective measures of formation position performance provide more detailed and actionable information about the position, and aid in uncovering the causes why the requirement was not met.

- Actual formation position MOP: A time-history trace, in meters, of the actual lateral spacing, and of the actual longitudinal spacing, between the autonomous vehicle and the nearest friendly vehicle.
- Perceived formation position MOP: A time-history trace, in meters, of the autonomous vehicle's perceived lateral and longitudinal spacing from the nearest friendly vehicle.
- Maneuver and obstacle safety warnings MOP: A time-history of the autonomous vehicle maneuver safety and obstacle safety warning status.

By defining effective measures of performance for the formation position capability, the test team gains more detailed information not just about whether the required performance was met, but also about what the cause of any performance degradations were. These measures provide actionable information that can be used for decisions about development, not just pass/fail grading. The use of runtime assurance methods for autonomous systems is a best practice that aids in providing detailed time-history data for measures and analysis, and will be the topic of future research papers.

- Was the vehicle out of position due to a problem perceiving its actual position? The developers can focus on improving its positional perception capability.
- Was the vehicle out of position due to actual safety limitations or obstacles that it correctly avoided? The autonomous system may be working as intended.
- Was the vehicle out of position due to a safety constraint that should not have applied? The developers can focus on the reasoning capability for when the constraints apply or not.
- Was the vehicle only a very small amount out of formation position? The user might not find the outcome as truly deficient among the larger military context, and accept a relaxation of the formation position requirements.

The use of runtime assurance methods for autonomous systems is a best practice that aids in providing detailed time-history data for measures and analysis, and will be the topic of future research and best practice papers.

A measure of effectiveness that assesses the military utility of the autonomous vehicle's formation position may require additional information about the scenario including threat positions and friendly force positions:

- Formation position MOE: A time-history analysis of the lateral and longitudinal difference between the actual autonomous vehicle position and the optimum electronic warfare position (*defined elsewhere) given friendly and threat system positions

This measure of effectiveness shows with precision the formation position relative to the actual military utility provided, with the limitation of needing optimum position calculations for assessment. However, this measure will then provide useful insight into whether the vehicle's autonomy capabilities in moving and maintaining formation position had a significant effect on its electronic warfare employment. MOE's commonly involve interrelated systems, and autonomous defense system MOE's will likely show more linkages to friendly and adversary scenarios than other military systems.

The implications of using effective measures should be clear: the stakeholders, program managers, testers, and ultimately the users can obtain more valuable, understandable, and actionable information about the autonomous defense system's perception, reasoning, planning, deciding, and coordination capabilities by defining effective measures that are relevant, specific, easily obtained, clear, precise, repeatable, standardized, and traceable.

Conclusion

This best practice examined the definitions, characteristics, challenges, and examples of effective requirements and effective measures for autonomous defense systems. The aspiration was to communicate the test and evaluation expectations for what specific characteristics make autonomy requirements and measures effective, or not effective. Simple examples related to autonomous systems were provided to help illustrate the impacts of having effective requirements and effective measures.

The STAT COE is engaged with the DoD autonomy development communities to continue to capture best practices and lessons learned, as autonomy technology matures, and as autonomy test and evaluation methods, practices, and techniques mature along with the technology. Future papers will document best practices for autonomy verification and validation strategies, test strategies, test planning, instrumentation, training, data management, trust, and other challenging autonomy topics. The T&E process builds upon the defined mission and system requirements, and depends utterly on the measures used to evaluate them --- effective autonomy requirements and measures provide the cornerstone for constructing effective evaluations of autonomy technologies.

References

[1] Autonomy Community of Interest Test and Evaluation, Verification and Validation Working Group, "Technology Investment Strategy", Department of Defense Research & Engineering, 2015.

[2] Mirriam-Webster Dictionary, 2021.

[3] General Accounting Office, "Best Practices Focus on Requirements", 2008.

[4] Joint Staff J-8, Joint Capabilities Integration and Development System Manual, 2018.

[5] Joint Chiefs of Staff, "Capstone Concept for Joint Operations: Joint Force 2020", 2012.

[6] Defense Acquisition University, "Glossary", DAU.edu website, 2021.

[7] Defense Standardization Program Office, "Defense and Program-Unique Specifications Format and Content", MIL-STD-961E Change 4, 2020.

[8] International Business Machines, "What is Requirements Management?", IBM.com website, 2021.

[9] Defense Standardization Program Office, "Guide for Performance Specifications", SD-15, 2009.

[10] National Aeronautics and Space Administration, "How to Write a Good Requirement", Systems Engineering Handbook Revision 1, 2007.

[11] Ahner, Thorsen, Divis, Jones, Oimoen, Rowell, Shelton, and Truett, "Advancements in Test and Evaluation of Autonomous Systems", OSD STAT Center of Excellence, 2020.

[12] Deputy Secretary of Defense, "Autonomy in Weapon Systems", DODD 3000.09, 2017.

[13] Requirements Working Group, "Guide for Writing Requirements, Revision 3", International Council on Systems Engineering, 2019.

[14] Cambridge Dictionary, 2021.

[15] Jim Oakes, Rick Botta and Terry Bahill, "Technical Performance Measures", BAE Systems, 2005.

[16] Defense Acquisition University, "Chapter 8 Test and Evaluation", Defense Acquisition Guidebook, 2021.

[17] Defense Acquisition University, "Risk, Issue, and Opportunity Management", DAU.edu website, 2021.

[18] Unmanned Campaign Framework, Department of the Navy, 2021.