TS

NOT MEASUREMENT SENSITIVE

DOE STD-1219-2016
May 2016

DOE STANDARD

Analysis and Evaluation of the Operability and Reliability of the Intrusion Detection and Assessment Systems

U.S. Department of Energy
Washington, D.C. 20585

AREA SANS

FOREWORD

1. This Department of Energy (DOE) technical standard is approved by all DOE components and their contractors

2. DOE technical standards, such as this technical standard, do not establish requirements. However, all or part of the provisions in a DOE technical standard can become requirements under the following circumstances:

    a) they are explicitly stated to be requirements in a DOE requirements document, or

    b) the organization makes a commitment to meet a technical standard in a contract or in an implementation plan or a program plan required by a DOE requirements document.

Throughout this technical standard, the word “shall” is used to denote actions which must be performed if the objectives of this technical standard are to be met. If the provision of this technical standard are made requirements through one of the two ways discussed above, then the “shall” statements would be requirements. It is not appropriate to consider that “should” statements would automatically be converted to “shall” statements as this action would violate the consensus process used to approve this technical standard.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

Analysis and Evaluation of the Operability and Reliability of the Intrusion Detection and Assessment Systems

1. SCOPE

A primary intent of alarm management is to ensure that the system and its components are configured (including balanced sensitivity and specificity) to minimize unnecessary alarm activation. This allows for a presentation of alarms at rates that the human operator can assimilate in order to maximize the alarm management system. The responsible security authority must select, install, configure, implement, operate, and maintain the system such that the system directs the human operator's attention to the most important alarm and the operator responds as required. Analyzing and reacting to system and operator performance data should enable optimization of the system and related human factors/operator interface.

Figure 1.1, Characteristics of Alarm Management, depicts a high-level outline of alarm installation and operations. The region circled by the dashed line displays what areas of alarm management this technical standard addresses.

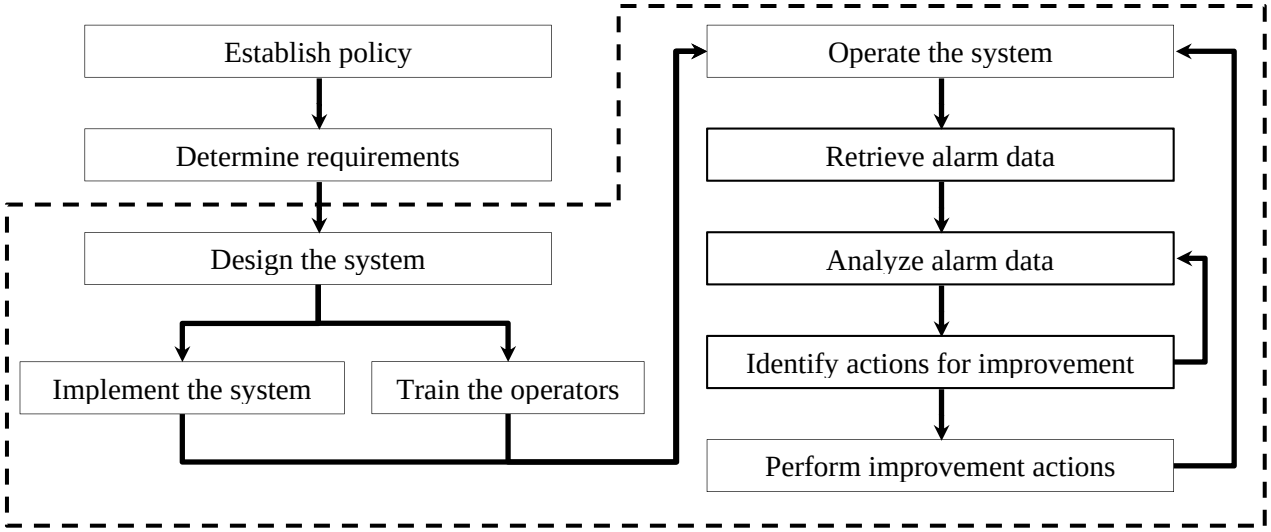

Figure 1.1 Characteristics of Alarm Management

2. PURPOSE

This standard provides an understandable process and methodology of the analysis and evaluation sites should consider conducting as required by Department of Energy (DOE) Order (O) 473.3A, Attachment 3, Section A, Chapter IX, 1.i.(1). In addition, the intent of this standard addresses the protective force equipment needs as outlined by DOE O 473.3A Appendix A, Section C, 2. Equipment and Attachment 2, Section C.3. Equipment. The analysis allows sites to identify and improve the intrusion detection and assessment systems (IDS) performance and define vital characteristics of the alarm system and station operators that influence operation and response. This standard establishes terminology and a methodology for identifying, analyzing, and communicating characteristics to support planning, operation, evaluation, and improvement of security system strategy and performance. Characteristics

listed may not be valid for all Department locations. This analysis process does not assess the effectiveness of the site-specific programs when compared to the rest of the department complex.

The analysis in this standard relies on a combination of measurable system and operator characteristics and subject matter expert evaluation. Analysts and responsible security authorities' use ratings to identify and examine weaker performing characteristics. A single poorly performing characteristic may not indicate the entire IDS is performing below requirements. However, responsible security authorities should not overlook a poorly performing characteristic. Identifying numerous system and operator characteristics allows trending, analysis, and correction of the root causes of system or operator weaknesses. Organizations can use this process to increase their confidence in the operability and reliability of the IDS. Sites may also elect to employ the processes and results as part of the vulnerability analysis and risk assessment.

Employing the process outlined in this technical standard does not provide a guarantee that security will detect, address, and classify all alarms effectively. However, the process provides responsible security authorities with performance data of the IDS and the status of supporting characteristics.

3. APPLICABILITY

This standard applies to all DOE locations employing IDS with system operators to protect assets and personnel. The Officially Designated Federal Security Authority (ODFSA) and Officially Designated Security Authority (ODSA) are responsible for analyzing the system.

4. REFERENCES

DOE O 473.3A, *Protection Program Operations.*

5. ACRONYMS AND DEFINITIONS

5.1. Acronyms:

CAS	Central Alarm Station
DOE	Department of Energy
FAR	False Alarm Rate
FoF	Force on Force
IDS	Intrusion Detection and Assessment System
LSPT	Limited Scope Performance Test
NAR	Nuisance Alarm Rate
ODSA	Officially Designated Security Authority
ODFSA	Officially Designated Federal Security Authority
OJT	On-the-Job Training
UAR	Unresolved Alarm Rate

5.2. Definitions

Alarm Frequency: Alarm frequency is the characteristic that measures the number of occurrences of an alarm per unit of time. These are false, nuisance, or unresolved alarms. The environment, system installation and maintenance, component health, or actual security events can affect the number of alarms. A system with numerous false, nuisance, unresolved, or authorized alarms may result in ineffective alarm station operator's response performance. Alarm frequency is a characteristic of the system scale.

Alarm Rate: Alarm rate is a quantity of alarms with respect to periods of time.

Analysis Matrix: A four-quadrant matrix used to display the system scale and the operator scale. The matrix provides a visual representation of how the IDS was functioning. The matrix gives an overall picture of the system health and identifies the weaker scale, as applicable. Identifying the lower performing scale allows the responsible security authority the opportunity to isolate the problematic characteristics.

Analyst: The Analyst is the person or group conducting the data gathering, rating of the characteristic, and plotting the resulting scores on the matrix. This person or group can be responsible for the entire analysis process or defining the rating of a single characteristic. It is vital that the analyst remains free of influence and that rating scores remain an honest and accurate assessment of the characteristic.

Ancillary Duties: Ancillary duties are any additional duties for the operator beyond their primary duty. As the number of other responsibilities increases for the system operator, the result is less attention to the system. Ancillary duties are a characteristic of the operator scale.

Authorized Alarm: An authorized alarm is an alarm for which stimulus activated a sensor, but the sensor was either:

- Placed into a state where access is permitted but sensors are still capable of generating an alarm (for example, the sensor was taken out of service for maintenance or repair and replaced by compensatory measures), or
- Activated by a known, expected, and operationally acceptable cause, such as performance testing of the IDS. The frequency of performance testing or planned alarm activities may still be a detriment to the health of the IDS as the operator will become complacent in response if they are constantly reacting to authorized alarms.

Average Alarm Acknowledgement Time: The average alarm acknowledgement time is a measure of time for the system to receive and display the alarm and the operator to acknowledge the receipt of the alarm and initiate assessment of the alarm point. Systems personnel generally gather this information through software analysis of data pulled from the system. This characteristic may also highlight the importance of a standard operator response to alarms. Average alarm acknowledgement time is a characteristic of the operator scale.

Characteristics: Characteristics are elements of the system that define how well the system is functioning and form the basis for rating the level of operator proficiency. The analysis process employs these characteristics to define the system health and identify weaknesses in system operation, maintenance, or design. For this Technical Standard, there are 13 total characteristics identified, eight system focused, and five operator based.

Ergonomics: The scale of human engineering applied to system design and maintenance. A well-arranged system results in rapid alarm assessment and system operation. A poorly engineered system requires an operator to move about a room and look in various directions usually resulting in slow response and possibly missed events. Ergonomics is a characteristic of the system scale.

False Alarm: An alarm for which the specific cause is unknown. Responsible security authorities may view a false alarm as an unwarranted alarm activated for reasons other than those for which the alarming device was designed. False alarms can be an indication of electronic malfunction such as component failure, communications failure, loose connections, power faults, or many other issues.

Similar to nuisance alarms, false alarms may occur in any system, but the frequency of these events is meaningful because they affect operator confidence in the system, and thus, overall security system performance. The False Alarm Rate (FAR) is the measurement of the frequency of this event.

Intrusion Detection System: A security system consisting of sensors capable of detecting one or more types of phenomena, signal media, annunciators, energy sources, alarm assessment systems, and alarm reporting elements including alarm communications and information display equipment. For the purposes of this technical standard, the operator(s) are considered part of the IDS.

Maintenance Rate and Repairability: A characteristic rate of the frequency and type of maintenance and the availability of repair and replacement parts required to keep the system functioning as required. The maintenance rate and repairability scale is a characteristic of the system scale.

Nuisance Alarm: An alarm for which a condition activated the sensor within the intended envelope and parameters of activation. Nuisance alarms are not related to any intent or activity for which the alarming device has been designed and operated. Personnel may view a nuisance alarm as a superfluous alarm in that the alarm has been activated by conditions for which the alarming device has been designed but not intended. The underlying intent and activity of that which has caused the activation do not inherently require action to counter a threat. Three of many examples of events that may activate alarms, but are, in fact, nuisance alarms include presence of an animal in a monitored location, authorized personnel failing to operate the system and generating an alarm, and weather-induced alarms. Excessive nuisance alarms may indicate a poor sensor selection for the environmental conditions. Nuisance alarms may occur in any system, but the frequency of these events is meaningful because they affect operator confidence in the system, and thus, overall security system performance. The Nuisance Alarm Rate (NAR) is the measurement of the frequency of this event.

Operability and Reliability: The IDS will detect and dispatch with a level of assurance necessary to meet security response requirements as identified in supporting security plans and documentation approved by the ODFSA.

Operability and Reliability Matrix: A four-quadrant matrix used to display the connection between system performance and expected operational capability. This matrix provides a visual representation of how the IDS meet detection and assessment requirements. Like the analysis matrix, this matrix gives an overall picture of the IDS and identifies the weaker scale as applicable.

Operator Performance Test: The operator performance test characteristic is a measure of operator testing success. Performance tests have pass/fail results and may be performance based or written. A team separate from the training section should conduct performance test to support this characteristic rating, though results should be shared with both training and performance assurance groups as applicable. The operator performance test is a characteristic of the operator scale.

Operator Scale: An average of the five operator proficiency characteristics rated from one to five in tenths, combined and plotted with the system scale to define the overall operability and reliability of the IDS. Analysis of the characteristics comprising this scale should encompass not only the operator's proficiency but also those external influences such as procedures, policies, and management decisions that may affect operations.

Oversight: Experienced operators tend to require less oversight than less experienced operators. This rating addresses how much oversight the operators require beyond initial qualification requirements. Oversight is a characteristic of the operator scale.

Performance Testing Alarm Rate: This alarm rate is a measure of the increase in frequency of alarms caused by a simulated threat and responding forces. Responsible security personnel should consider what alarms may be activated and the rate of those alarms during Force on Force (FoF) exercise and training activities against an appropriate adversary force as defined by the site's Risk Assessment, Vulnerability Analysis, as required by the Graded Security Protection Policy. Analysts may assess the alarm rate during a simulated event such as a tabletop or determine it directly during an actual test event. This characteristic simulates how the system responds during an assault and the effects on the operator response. Performance testing alarm rate is a characteristic of the system scale.

Primary Duties: The alarm station operator responsible for the primary alarm system focusing on the most critical targets or material is primarily responsible for alarm acknowledgement and dispatch. For example, other duties include documenting and classifying alarms, managing non-alarm response communications, or controlling access are the responsibilities of other operators or personnel or are ancillary duties of the alarm station operator (See Ancillary Duties).

Responsible Security Authority: Person or organization responsible for addressing the characteristic rated. As facilities employ different contractors or organizations to address protective force personnel and systems, the technical standard employs this neutral term.

Scale: A range of numbers from one to five defining the capability of the characteristics analyzed, a score of one indicates a poorly performing characteristic and a five reflects an almost perfect operating characteristic.

System Scale: An average of the eight system characteristics rated from one to five in tenths, combined and plotted with the operator scale to define the overall health of the IDS.

System Interface: This characteristic measures the ease of computer interface and system software operation. A system requiring a single button push to acknowledge and review alarm data promotes a rapid and accurate response by security forces. A system with a lengthy interface process using multiple systems and keyboards to acknowledge and review an alarm generally yields a slower, less accurate response by security personnel. The system interface is a characteristic of the system scale.

Tamper Alarm: An indication that unauthorized access to a security alarm management and control system enclosure or device is being attempted.

Technical Knowledge: Technical knowledge is a measure of the skill and training required to maintain and repair the system. Technical knowledge is a characteristic of the system scale.

Testing Frequency: The number of different personnel and organizations completing system and operator testing provides a level of assurance that the system is performing as designed. Fewer different testing entities might indicate that responsible security authorities rarely confirm the system abilities. More resources looking at the system results in an increased opportunity of finding and addressing weaknesses in a timely manner. Testing frequency is a characteristic of the system scale.

Training Test Scores: Test scores are a rating of operator understanding of system training and re-training. Training test scores is a characteristic of the operator scale.

True Alarm: Also known as, an intrusion alarm is a condition activated the alarm for which the alarming device has been designed and operated (for example, an attempted intrusion by unauthorized personnel).

Unresolved Alarm: An alarm for which the cause has not been determined but investigation of the cause is still ongoing. The Unresolved Alarm Rate (UAR) is the measurement of the frequency of this event.

Z-Score: The Z-Score is a statistical measure of the health of the alarm sensor. This measure identifies if a sensor is operating as expected or if the alarm frequency from the sensor is increasing or decreasing in an uncharacteristic way, usually indicating the sensor requires repair or replacement. The Z-Score is a characteristic of the system scale.

6. BASIC ANALYSIS PROCESS

This technical standard establishes thirteen characteristics that provide indication of the health of the IDS. These generic characteristics are alarm frequency, ancillary duties, average alarm acknowledgement time, ergonomics, operator performance tests, maintenance rate and repairability, oversight, performance testing alarm rate, system interface, technical knowledge, testing frequency, training test scores, and Z-score. These thirteen characteristics are separated into system and operator categories. An analyst, ideally a subject matter expert not responsible for the characteristic, rates each characteristic from one to five supported by documented data and evidence. This standard provides a separate scale for each of the thirteen characteristics in Section 7. Section 7 also provides a narrative of each characteristic and sources of data and information the analyst may consider using to scale the characteristic. Table 6.1, Basic Scale, below provides a generic description of each scale level. As the analysts are rating the characteristic the individual sensor, system, or operator should be rated for better causal analysis and corrective action development. Increased analysis granularity also supports a more accurate rating of the characteristic.

Table 6.1 Basic Scale

Basic Scale		
Scale	Designator	Description
1	Poor	Operating well below levels defined in Section 7
2	Moderate	Operating below levels defined in Section 7
3	Fair	Operating at levels defined in Section 7
4	Good	Operating above levels defined in Section 7
5	Excellent	Operating well above levels defined in Section 7

Once analysts have identified and rated all characteristics, the analyst responsible for the entire report then averages the characteristics as part of either the system scale or the operator scale. The analyst rounds the average result of each scale to the tenths (0.01-0.04 round down to 0.0 and 0.05-0.09 round up to 0.1) and plots the resulting point on the analysis matrix to identify the IDS operability and reliability. The horizontal axis of the matrix is the system scale and the vertical axis is the operator scale, refer to Figure 6.1, Analysis Matrix, below. The analysis matrix is broken into four quadrants emphasizing the strengths and weakness of the IDS, refer to the Analysis Matrix Definition Chart 6.1 below. The goal of this process is to provide a visual representation of the health of the IDS. As resources are generally scarce, any response to characteristics performing poorly should begin with a detailed investigation of the causes. The intent is not to make important or costly decisions on averaged values but understand the health of the IDS as a whole and respond to weaknesses appropriately, as resources allow.

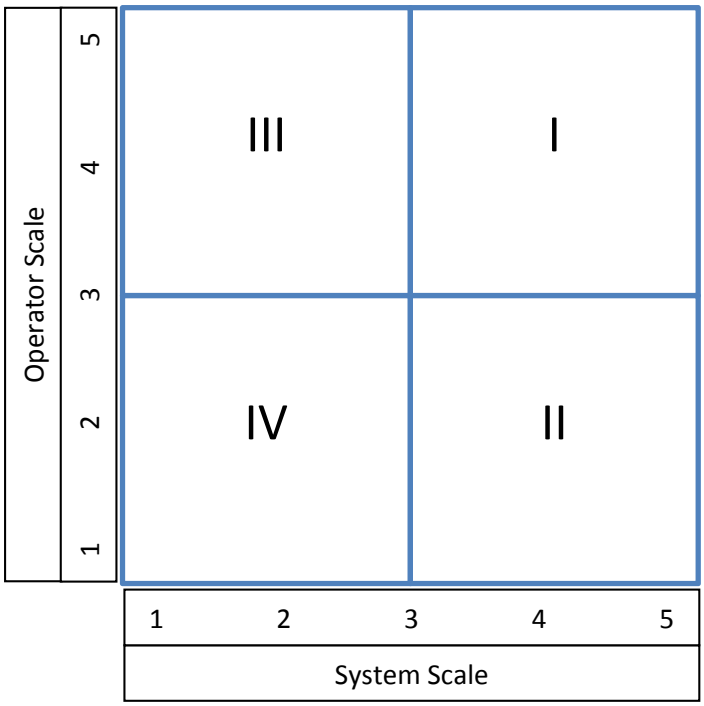

Figure 6.1 Analysis Matrix

I – Optimal level of operator proficiency and system performance
II – Low operator proficiency and good system performance, the system is heavily responsible for the success of operation and detection
III – High operator proficiency and poor system performance, the operator is heavily responsible for the success of operation and detection
IV – Low operator proficiency and poor system performance, indicative of unsatisfactory overall performance.

Chart 6.1: Analysis Matrix Definition

After plotting the system and operator scale on the analysis matrix, the analyst should identify the location of IDS health based on Figure 6.2, Operability and Reliability Matrix, below. This matrix is similar to the analysis matrix. However, this matrix includes an overlay of colors denoting the expected operational area of the IDS, the color the plot should fall within, based on the assets protected defined in the Operability and Reliability Matrix Definition Chart 6.2 below. If the plotted value is near the border of the overlay, the responsible security authority should investigate and begin causal analysis and corrective actions to address the weakness. Again, as resources are generally scarce any response to characteristics performing poorly should begin with a detailed investigation of the causes. The intent is not to make important or costly decisions on averaged values but rather be aware of the health of the system and correct weaker performing characteristics as resources become available. Analysts shall document all decisions in a report reviewed by the responsible security authorities and approved by the ODFSA. The report shall highlight the strengths and weaknesses of the IDS so responsible security authorities can make resource adjustments as necessary. Analyst should present the report to the ODFSA on a periodic (annually, semi-annually, or quarterly) basis for tracking and trending purposes.

Each site has specific training, testing, and detection requirements established specifically for their security needs. Responsible security authorities should not use scores from this process when compared to other locations as the basis for changing a program. For example, a location rating their training poor due to low test scores should not adjust their training lesson plans to mirror another location with a higher rating, unless additional process improvement techniques are employed that support these changes. The analysis completed by responsible security authorities should use the terminology provided in this standard when possible and define any site-specific system characteristics as applicable.

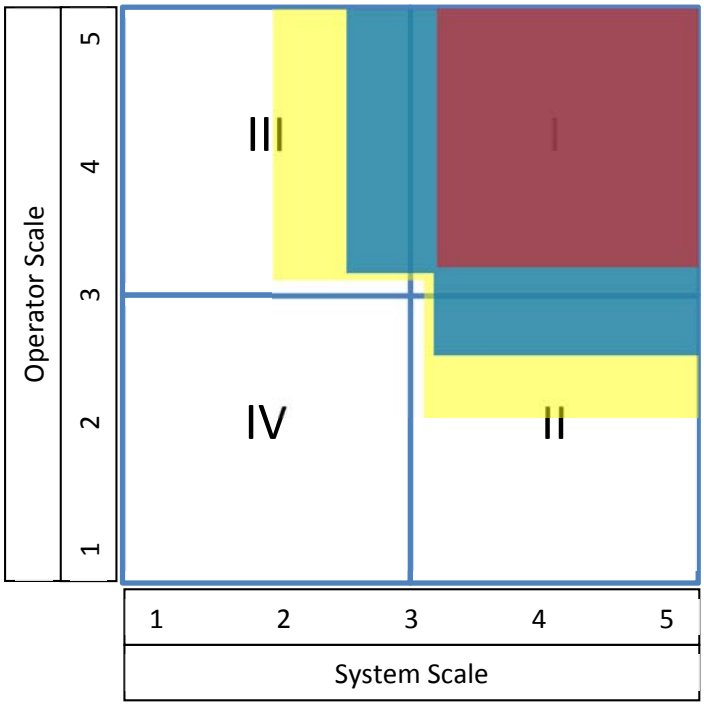

Figure 6.2 Operability and Reliability Matrix

Red – Systems protecting CAT I/II, Top Secret material (Outer edge of scale is 3.2 for both System and Operator) are excellent.

Blue – Systems protecting CAT III/IV, Confidential/Secret material (Outer edge of scale is 2.5 for System and Operator if the opposite scale rates at 3.2) are above average.

Yellow – Systems protecting personnel and other governmental property (Outer edge of scale is 2.0 for System and Operator if the opposite scale rates at 3.1) are average.

White – Systems are operating unsatisfactorily (Outer edge of scale is between 3.1 and 2.0 for System and Operator if the opposite Scale rates below 3.1. Scale may be as high as 5.0 and still unsatisfactorily if the opposite scale is below 2.0).

Chart 6.2: Operability and Reliability Matrix Definition

7. INTRUSION DETECTION SYSTEM CHARACTERISTICS

Outlined below are descriptions for each characteristic, sites should consider including in analysis techniques. Each characteristic reviewed may have multiple conflicting points of data when compared to the chart. Within a characteristic, some data may indicate better performance scoring a four or five and

other data may indicate poor performance scoring a two or three, making analysis challenging. For example, a Central Alarm Station (CAS) may have many stationed operators with few ancillary duties resulting in only one ancillary duty per person with the ancillary duties aligned to their system operations. This example situation should initially score a four on the ancillary duties characteristic scale. However, during an exercise the analyst identifies that the ancillary duty requirements result in a vulnerability indicating a rating of two on the scale. In this case, the analyst developing the score shall document both results. If the vulnerability is such that the responsible security authority cannot quickly address the issue or the vulnerability requires prolonged time and resources to fix, the ancillary duty characteristic may score a two. However, if the vulnerability is of a limited nature and quickly corrected, the analyst could assign a score of four once mitigated. A formal analysis of the issue identified is required along with any possible mitigations and corrections to justify any rating shift. This is a case where the prudent decision is to maintain the score of two until the responsible security authority can correct the concern permanently.

Another aspect of consideration for the analyst is the cross effect of characteristics. For example, if the analyst elects to monitor alarm frequency during a rainstorm and use this as their baseline alarm frequency the results may show a higher level than actually present during a non-rainy day. This poor analysis decision results in a lower score on the alarm frequency scale and possible difficulty in clearly rating other characteristics. Analysts should be aware of the characteristic they are measuring and limit, as much as practicable, the impacts of other characteristics at the time of measurement and analysis.

Characteristics identified below are not applicable to all facilities and sites but are concepts sites should consider when analyzing the health of their IDS. As part of the site-specific analysis, the documentation should include all applicable characteristics and the terminology to define what the characteristic addresses. Site-specific analysis should document the information to support any characteristic weighting (refer to Section 9 of this standard) or adjustments to the matrix. Other sources of information used to define site-specific characteristics include assessments, performance tests, risk analysis, vulnerability analysis, and procedures.

7.1. Alarm Frequency

Alarm data analysis and performance evaluation provides information that is useful for determining current sensor performance, sensor performance trends and patterns, and uncovering potential to improve sensor, and thus system, performance. Analyzing the state of a system with defined evaluation methods is essential because it establishes empirical and quantifiable evidence of performance that the responsible security authority compares to policy and requirements. It also provides feedback on improvement initiatives and motivation/justification for required upgrades or other expenditures.

Analysis and rating of the alarm frequency characteristic does not require separate analysis of FAR/NAR/UAR. The alarm frequency characteristic focuses on the system as a whole. Unplanned alarms that are not actual security threats impact the operator’s faith in the system regardless of the final designation by the operator or system. Identifying the alarm as FAR/NAR/UAR is more vital for corrective actions by system technicians. An analyst should consider breaking down the data for causal analysis and corrective actions. If the granularity identifies one of the rates as a dominant issue, then the responsible security authority should address that aspect of the characteristic. The intent of reviewing alarm frequency is to ensure the alarm rate does not impact the health of the system and the faith of the operator in the system.

Each site should have a documented process in place to retrieve, process, analyze, and identify corrective actions to manage the FARs/NARs/UARs. This process should include performance thresholds used by the responsible security authority in determining whether sensors comply with all

relevant policies and requirements. Many data analysis software options/packages are able to calculate the numbers used in analyzing sensor performance data once the software is configured properly and provided the appropriate data.

Figure 7.1, Alarm Rate Designation Flowchart, shows a flowchart for classifying the alarm type. This flowchart involves processes associated with the “Operate the system” block in Figure 1.1, Characteristics of Alarm Management, because it is the mental process that operators or others go through while assessing each alarm and is illustrative of how alarms are assigned to the various groups. There is some flexibility in how the responsible security authority assigns each alarm to a group. For example, the console operator, a maintenance representative, the system, or another entity that identifies the cause and relevant parameters for making the change should regroup an initially unresolved alarm as a false alarm.

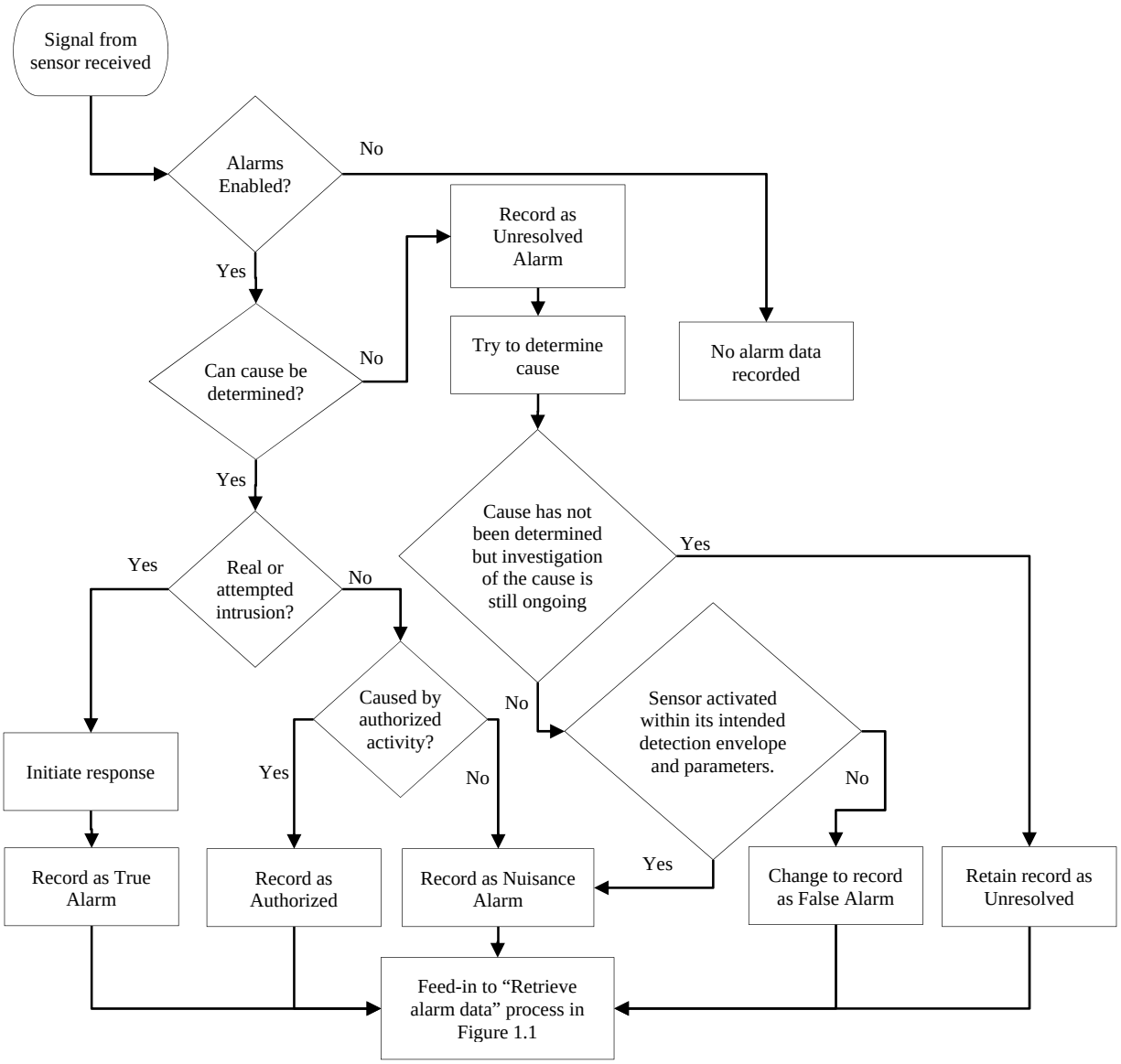

Figure 7.1 Alarm Frequency Designation Flowchart

Once the parameters are established, appropriate management should define thresholds. To calculate the alarm rate, simply divide the number of alarms for a particular sensor by the associated time range. However, when developing a practical approach, several questions arise. These questions include, but are not limited to:

- What is the appropriate time range?
- How should the number of alarms be based? On one sensor, a protection zone, or some other clustering?

The significance of the appropriate time range of each individual alarm to the rate are reduced if the associated time range is increased. However, increasing the time range may also increase the number of subject alarms. The units used for the alarm rate characteristic are "alarm" and "hour," with the ratio being "alarms per hour."

It is meaningful to view alarm rate distributions according to very short or very long time frames, although this standard cannot address all possible scenarios, it does address the rates on a daily and weekly cycle. An example of a short time frame consideration is a situation where the sun being at a particular position affects a sensor or group of sensors at a certain time of day during a specific time of year. This scenario may result in an exterior sensor activating many times in a one-hour period for those days, but the rest of the day not activating at all. In this case and for this particular sensor, it is meaningful to analyze the collected sensor performance data on an hourly basis. An example of longer periods is seasonal changes resulting in sensor performance that repeats over multiple years. The basic topics for analysis performed according to this standard include:

Daily Individual Alarm Counts
Daily Individual Alarm Rates
Daily Individual Alarm Outliers

Weekly Individual Alarm Counts
Weekly Individual Alarm Rates
Weekly Individual Alarm Rate Means
Weekly Individual Alarm Rate Standard Deviations
Individual Alarm Statistical Thresholds

Daily Individual Alarm Counts: Record the number of True, Authorized, Nuisance, False and Unresolved Alarms that occurred for each individual sensor during the 24-hour day, beginning and ending at midnight of consecutive days.

Daily Individual Alarm Rates: Record the True, Authorized, Nuisance, False and Unresolved Alarm Rates for each sensor by dividing its Daily Alarm Count by 24, resulting in rate units of (true, authorized, nuisance, false and unresolved) alarms per hour.

Daily Individual Alarm Outliers: Based on the Individual Alarm Statistical Thresholds, identify any daily rate that has exceeded the threshold established for that sensor. Security or site managers should also set thresholds based on relevant information regarding the sensor, its environment, the security system configuration or operation, or other pertinent factors. The decision to use statistically developed thresholds use the assumption that the sensors and security system are adequately designed, implemented and operated. These assumptions support the concept that statistically high rates indicate that action is required to identify and rectify any current or emerging problems.

Weekly Individual Alarm Counts: Sum and record the number of True, Authorized, Nuisance, False and Unresolved Alarms that occurred for each individual sensor during the seven consecutive day period.

Weekly Individual Alarm Rates: Sum and record the number of True, Authorized, Nuisance, False and Unresolved Alarms that occurred for each individual sensor during the seven day week, then, divide by 168 hours to result in rate units of (true, authorized, nuisance, false and unresolved) alarms per hour. The approach that an average of the daily rates and some potentially available variability and distribution information is not applied should be sufficient for determining the center and basic variability of the sensor performance data, which analysts should use to develop the Individual Alarm Statistical Thresholds.

Weekly Individual Alarm Rate Means: Analysts should consolidate each weekly individual alarm rate with all other available weekly individual alarm rates for the same sensor and its alarm groups so that the center of this data collection is determined. The center of data used by this standard is the

mean. One method for determining the mean is to sum all of the Weekly Individual Alarm Rates for a particular sensor, by group (a collection of sensors in a defined vicinity (e.g., vault or zones/sectors of intrusion detection, etc.) not necessarily all of the same type of sensor), and then divide by the number of Weekly Individual Alarm Rates. For example, if seven weeks' of Weekly Individual Alarm Rates are recorded and those rates for unresolved alarms are: 27, 34, 12, 41, 29, 27 and 31 unresolved alarms per hour, then the Weekly Individual Alarm Rate Mean for that sensor, regarding unresolved alarms, is:

(27+34+12+41+29+27+31) / 7 = 201 / 7 = 28.7 unresolved alarms per hour (rounded).

Weekly Individual Alarm Rate Standard Deviations: The analyst derives the standard deviation from the data in the frequency distributions assembled from the Weekly Individual Alarm Rate Means.

The standard deviation is the square root of the average of the squared differences from the mean, thus:

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2}$$

Where N is the same sample size, x_i is an individual score, and xbar is the mean.

This formula uses N-1 instead of N in front of the summation sign because it is the standard deviation formula for a sample rather than a population.

Calculators, spreadsheets, statistical programs, and other software have functions for calculating standard deviation. As an example of the actual calculation, for the standard deviation for the numbers, 27, 34, 12, 41, 29, 27 and 31 unresolved alarms per hour, the following approach may be used:

x_i	x_i - mean	$(x_i - \text{mean})^2$
27	-1.71429	2.939
34	5.28571	27.939
12	-16.7143	279.367
41	12.28571	150.939
29	0.28571	0.0816
27	-1.71429	2.939
31	2.28571	5.224
	SUM:	469.429
	SUM/6:	78.238
	Square Root:	8.845

The standard deviation from this data set is s = 8.9 unresolved alarms per hour (rounded).

Individual Alarm Statistical Thresholds: The method for determining the thresholds for individual sensor performance data is to compare the Daily Individual Alarm Rates directly to the Weekly Individual Alarm Rate Means and the Weekly Individual Alarm Rate Standard Deviations.

Using the statistical techniques above allows the analyst to rate the alarm frequency characteristic using Table 7.1, Alarm Frequency Scale, below and prepares to develop data to support rating the Z-Score characteristic.

Table 7.1 Alarm Frequency Scale

Alarm Frequency	
Scale	Description
1	Interior Alarm 1 every 8 hours or Exterior Alarm 1 every hour
2	Interior Alarm 1 every 24 hours or Exterior Alarm 1 every 4 hours
3	Interior Alarm 1 every 240 hours or Exterior Alarm 1 every 8 hours
4	Interior Alarm 1 every 2400 hours or Exterior Alarm 1 every 24 hours
5	Interior Alarm 1 every 4800 hours or Exterior Alarm 1 every 48 hours

7.2. Ancillary Duties

The primary focus of alarm system operators should be the response to alarm systems protecting department personnel and material. However, the system operators have a unique position of responsibility and capability, as they are stationed in one location, usually central to the facility. Due to this central location, additional tasks are often assigned to these individuals. These additional tasks include radio communications, equipment assignment, access control, long-range detection system operation, and video assessment systems operations, among others. These tasks may range in difficulty from very simple data logging to very complex system operations and analysis.

Additional tasks require system operators to multitask, especially during alarm situations. Studies have shown that task shifting or multitasking results in negative effects on accuracy[1]. Furthermore, forced multitasking produces significantly negative results[2], forced shifting common in alarm response activities. Although alarm systems have visual and audible alarm notification, the ancillary duty requirements may still delay or hinder alarm response actions.

Ancillary duties are part of the operator scale and should be a ratio based on the number of operators assigned to the alarm station per work shift. Station orders, response plans, and security plans should outline ancillary duties of the alarm station operators. Additionally, the analyst assigning the rating should get insight by interviewing and observing the station operators. Analyst observation is vital to identify the number of task shifts required during daily activities and alarm response demands and the negative impact of the ancillary duties.

Most sites assign more system operators during working hours to address the volume of duty demands and requirements. Off-shift times tend to have fewer assigned operators and fewer responsibilities. Since the analyst should quantify alarm station staffing and the number of duties, the data should be easily recorded and analyzed. If there are differences in shift assignment, the analyst should elect to either perform an analysis on both shifts separately or combine the two in an average scale. No matter the decision, analysts should assign each separate station or post a scale rating to identify possible weaknesses or enhancements, for future causal analysis.

Another aspect of the analysis of ancillary duties is whether the additional duty aligns with the alarm station primary responsibilities. For example, the CAS has a primary focus of a single facility with a Material Access Area (MAA). As part of the CAS ancillary duties, the operator provides access control to certain areas within the MAA. This is an ancillary responsibility but aligns with their

[1] Adler, R.F. & Benbunan-Fich, R. (2011). Juggling on a high wire: Multitasking effects on performance. *International Journal of Human-Computer Studies* 70 (2012) 156-168

[2] Buser, T. & Peter, N. (2011). Multitasking. *Experimental Economics* 15.4 (2012) 641-655

primary responsibilities since it keeps attention of the operator on the facility of interest. If on the other hand, the CAS operator is responsible for alarms at the MAA and at a limited area outside of and not directly associated with the facility this duty does not align with their primary responsibility, even if the limited area uses the same alarm monitoring and access control system as the MAA facility.

Analysts should identify additional data supporting the rating of this topic while reviewing how the alarm station operator(s) respond during exercise activities. If the ancillary duty requirements are numerous but the operator(s) are experienced, general daily activities may not highlight the impact of these additional duties. This analysis can be an objective for the exercise or test. The analyst should use caution to ensure the impact of ancillary duties are observed and not the characteristic of Performance Testing Alarm Rate or that data for both characteristics remains separate if analyzed concurrently.

Referring to Table 7.2, Ancillary Duty Scale, below the analyst should qualitatively rate the quantitative data.

Table 7.2 Ancillary Duty Scale

Ancillary Duty	
Scale	Description
1	Each operator is responsible for three or more other duties not aligned to system operations. The operator cannot complete all assigned duties related to the primary alarm system. During an exercise or actual alarm response, operators fail to address many alarms and a clear vulnerability exists. During alarm response, task shifting may occur more than 15 times within a half hour.
2	Each operator is responsible for one or two other duties not aligned to system operations. The operator cannot complete all assigned duties related to the primary alarm system. During an exercise or actual alarm response, operators fail to address some alarms and a possible vulnerability exists. During alarm response, task shifting may occur between 11 and 15 times within a half hour.
3	Each operator is responsible for three or more other duties aligned to system operations. The operator cannot complete all assigned duties related to the primary alarm system. During an exercise or actual alarm response, operators fail to address a few alarms but no apparent vulnerability exists. During alarm response, task shifting may occur between 5 and 10 times within a half hour.
4	Each operator is responsible for one or two other duties aligned to system operations. The operator can complete all assigned duties related to the primary alarm system. No vulnerability exists. During alarm response, task shifting may occur less than 5 times within a half hour.
5	The operator is only responsible for the primary alarm system focusing on the most critical targets or material. Alarm acknowledgement and dispatch is the operators' only responsibility. No vulnerability exists. During alarm response, no task shifting occurs.

7.3. Average Alarm Acknowledgement Time

The average time for alarm acknowledgement is a characteristic reflecting the speed in which operators address alarms displayed by the system. Trained operators with defined responsibilities and clear procedures should be capable of operating the system efficiently ensuring that alarms are acknowledged, initially assessed, and security forces are dispatched as required. Longer acknowledgement times could indicate that operators are not sure of responsibilities, the system is not functioning properly, or the number of sensors the operator is expected to monitor is too high

resulting in an excessive volume of alarms. Analysts should rate all personnel and work shifts to highlight any weaknesses. Analysts may gather and analyze acknowledgement times using data mining software (e.g.Tableau), running performance tests, or both. Table 7.3, Average Alarm Acknowledgement Time Scale, identifies a purely quantitative rating of the amount of time required for the operator to complete initial alarm response actions. Analysts should remove authorized alarms generated during performance testing or maintenance from the analysis as this can skew the data.

Analyst should develop a separate scale reflecting the acknowledgement time assessed during performance testing. The analyst should use the data to support the rating of the Performance Testing Alarm Rate characteristic and any causal analysis conducted to identify the root weakness affecting the IDS. Analysts should also adjust this scale in the final report to reflect the site-specific vulnerability analysis and response timeline requirements.

Although this characteristic is part of the operator scale, if problems emerge, a causal analysis may identify other aspects affecting this characteristic such as procedures, communications, system functionality, etc. In addition to pulling time data from software, analysts should observe the operators and system to understand how these aspects influence the time. Although this characteristic is an average, analysts should be aware of and investigate extreme outliers for causal analysis and corrective action.

Table 7.3 Average Alarm Acknowledgement Time Scale

Average Alarm Acknowledgement Time	
Scale	Description
1	Greater than 60 seconds
2	Greater than 45 seconds but less than or equal to 60 seconds
3	Greater than 30 seconds but less than or equal to 45 seconds
4	Greater than 20 seconds but less than or equal to 30 seconds
5	Less than or equal to 20 seconds

7.4. Ergonomics

This characteristic is a rating of how well engineers designed the system to support the operator(s). This is an important characteristic, as a well-engineered system makes operations simple, pain free, and convenient. Systems with everything within reach without straining by the operator are likely to result in quicker, more accurate alarm response. Facilities that require one operator to move across the room to operate numerous systems or scan a wall of monitors to assess an alarm have slower response times and are burdensome on the operator.

The Department of Labor through the Occupational Safety and Health Administration (OSHA) provides an appropriate tool for assisting in rating the ergonomic characteristic. The Computer Workstations eTool is a fillable form used to create a safe and comfortable computer workstation. The form is online at https://www.osha.gov/SLTC/etools/computerworkstations/checklist_evaluation.html. A ‘No’ selection on the form could indicate a problem.

When a number of system operators are assigned to the alarm station, this scale should analyze each operator’s workstation and rate the whole layout. As work schedules influence the number of personnel assigned to the alarm station, analysts have to review conditions during working and non-working hours. Many Department locations employ more than one alarm station, which may also require a separate review and rating. This characteristic is part of the system scale and is more of a qualitative characteristic rating scale. For this reason, analysts sometimes elect to employ a health

professional, specially trained to review the ergonomic arrangement of workstations and recommend a rating using the scale presented in the Table 7.4, Ergonomics Scale, below.

Table 7.4 Ergonomics Scale

Ergonomics	
Scale	Description
1	Duty station fails to conform to the operator, lighting levels are lacking (Less than or equal to five footcandles at the workstation), ambient noise is too excessive (greater than or equal to 75 decibels or greater), the operator is required to stand and move to address alarms and dispatch security, workstations are more than an arm length apart. Analysts using the OSHA Computer Workstation eTool selected more than six 'No's on the form.
2	Duty station conforms to the operator, lighting levels are limited (greater than five footcandles but less than or equal to ten footcandles at the workstation), ambient noise is high (70 decibels or greater), the operator is required to stand and move to address alarms and dispatch security, workstations are about an arm length apart. Analysts using the OSHA Computer Workstation eTool selected between five and six 'No's on the form.
3	Duty station conforms to the operator, lighting levels are acceptable (greater than ten footcandles but less than or equal to twenty footcandles at the workstation), ambient noise is moderate (60 decibels or greater); the operator is required to stretch and stand to address alarms and dispatch security. Analysts using the OSHA Computer Workstation eTool selected between three and four 'No's on the form.
4	Duty station strongly conforms to the operator, lighting levels are good (greater than twenty footcandles but less than or equal to fifty footcandles at the workstation), ambient noise limited (55 decibels or greater), multiple computer stations are required to conduct alarm activities abut all are within arm's reach with only limited movement require to observe and dispatch security. Analysts using the OSHA Computer Workstation eTool selected between one and two 'No's on the form.
5	Duty station seems to be engineered for the operator individually, lighting is excellent (greater than fifty footcandles at the workstation), ambient noise levels good (less than 50 decibels), all operations are within immediate reach and require no additional movement to observe and dispatch security. Analysts using the OSHA Computer Workstation eTool selected no 'No's on the form.

7.5. Maintenance Rate and Repairability

A system requiring more frequent or intensive maintenance may lead to a weakness in system operability and reliability, potentially affecting the operator's faith in the system. Systems that fail to operate properly require more compensatory measures or may fail to alarm as required. Table 7.5, Maintenance Rate and Repairability Scale, below is part of the system scale and is quantitative in nature. Analysts should track corrective maintenance records, including corrections made due to performance testing results. Characteristics that have analysis that is more granular produce better utility in supporting a causal analysis. The analysis granularity should focus on maintenance frequency down to the individual sensor but should score based on an average for the entire IDS for rating purposes.

Analysts should identify sources of data by reviewing maintenance records and system operator records. The frequency of maintenance may also be reflective of the age of the system, resulting in a low score for both repairability and maintenance. If the site employs system technicians, analysts should note the frequency of assistance by external technicians or manufacturing sources that also result in an impact to the Technical Knowledge characteristic. The size of the system and the number of technicians may also be a data point analysts should review. A system with numerous sensors

protecting multiple facilities is likely to have more frequent repair maintenance than a small office with limited sensors. Analyst's data should reflect the size of the system and number of technicians as part of the analysis process. The technical standard does not intend this scale to address the single catastrophic event such as lightning strikes or flooding requiring large system replacements but rather the average efforts required to keep the system running on a daily basis.

Additionally, a newer system with many options for repair and replacement parts increases the operator trust in the system and likelihood that the system is performing at optimum levels. By contrast, a system that is mature with limited repair options and requires technicians to generate replacement parts or cannibalize other systems to maintain operations is likely to be operating at less than ideal levels. A well-operating system works in conjunction with talented operators to ensure protection of assets. A mature system operating at the end of life requires talented operators to address alarms and respond accordingly. Systems should not remain at this senior stage of their repairability. One point of caution, a senior system may still be very effective if parts are common and inexpensive and repairs are easy. Analysts should document decisions on their scales to ensure sites do not use precious resources to replace established systems simply to purchase the 'latest' device.

Sources of information for rating this characteristic should include system installation manuals, drawings, purchase orders for replacement parts and maintenance records, among other things. Like many other scales, analysts should apply this rating differently to the different supporting systems. For example, an alarm system over ten years old that had the access control system replaced within the past year requires the analyst to average the results or complete the analysis separately for each system and average the results for an overall picture.

Table 7.5 Maintenance Rate and Repairability Scale

Maintenance Rate and Repairability	
Scale	Description
1	Corrective maintenance is a daily occurrence and is commonly a costly piece of equipment or system replacement. On average compensatory measures are in place for months. Replacement parts are not available or are expensive (>$5,000), and are difficult to obtain with over three months waiting time, system is greater than ten years old and no longer supported by the manufacturer.
2	Corrective maintenance is almost a daily occurrence and is commonly a major piece of hardware (multiple sensors, extensive wiring, cards within the data gathering panel). On average compensatory measures are in place for weeks. Replacement parts are of limited availability, are expensive (less than $5,000 but greater than or equal to $1,000), and are difficult to obtain with over two months waiting time, system is ten years old.
3	Corrective maintenance is less frequent (once a week) and is commonly a minor hardware replacement (sensor head). On average compensatory measures are in place for a week. Replacement parts are available but difficult to obtain, are moderately expensive (less than $1,000 but greater than or equal to $500), with over one month waiting time to get parts, system is less than eight years old.
4	Corrective maintenance is infrequent (once every couple of weeks) and usually a minor system or setting adjustment. On average compensatory measures are in place for days. Replacement parts are available, are moderately inexpensive (less than $500 but greater or equal to $100), are easier to obtain with less than one week waiting time, system is less than five years old.
5	Limited system corrective maintenance (once a month), on average compensatory measures are in place for hours. Parts are easily obtainable by the technician at a local supplier and cost less than $100, system is less than three years old.

7.6. Operator Performance Test

Analysts use operator performance test to determine how well operators are performing and maintaining system knowledge. As with all performance tests, performance based tests are preferable, written knowledge exams are acceptable, and most sites use a blend of the two methods. An important aspect of an operator performance test is that the responsible security authority conducts the test separately from training. Performance tests should have a clear pass/fail result and provide no re-training. The intent of the performance test is to closely mirror real activities and generate real operational data. It is common for performance tests results to be slightly less favorable than training results as there is little if no preparation provided for a performance tests.

The responsible security authority, performance-testing section, quality assurance section, or an external assessment, provides performance test data. Analysts, or those responsible for maintaining performance test scores and results, should maintain them for the individual operators to allow for better causal analysis should analysts identify a weakness. Analysts should also track performance test data by topic or title for the same reasons.

Table 7.6, Operator Performance Test Scale, is part of the operator scale. Some common performance tests include but are not limited to, alarm response procedures, system operations, record keeping, and notification procedures. Analysts should rate this characteristic based on the percent of passing results. As with most characteristics in this analysis, the performance tests characteristic is a quantitative value. Additionally, the responsibility rests with the analyst to ensure the performance test information used is applicable to the job of interest. For example, the performance-testing section completing a performance test on the operators concerning their knowledge of deadly force is unlikely to apply to alarm systems operation. Analysts should include other styles of testing such as Alarm Response and Assessment Performance Test and FoF exercises focusing on the resulting data for the performance test characteristic.

Analysts should always be aware of disparities in results conducted by different assessment groups. Slight differences in the conduct of the test, test objectives, or testing methodologies may create differences in the results. The testing criteria and results should be reviewed and, in some cases, the analyst could conduct their own testing or observe numerous testing events by multiple sources to ensure data integrity and consistency. Testing scores outlined below are based on criterion-referenced grading systems as defined by the Department of Education.[3]

Table 7.6 Operator Performance Test Scale

Operator Performance Test	
Scale	Description
1	Average test scores less than 65%
2	Average test scores between 65% - 75%
3	Average test scores between 75% - 85%
4	Average test scores between 85% - 95%
5	Average test scores between 95% - 100%

7.7. Oversight

Another method of assessing the expertise of the system operators is to measure the amount of required oversight or leadership. The oversight characteristic is a scale of the amount of supervision required for the operator to respond to alarms appropriately. Oversight includes but is not limited to, uniformed supervisors, training instructors, or senior system operators. The focus of the oversight

[3] Department of Education (2008) Structure of the U.S. Education System: U.S. Grading Systems. Located at www2.ed.gov/about/offices/list/ous/international/usnei/us/grading.doc

characteristic is not the amount of training and supervision required prior to operator qualification, but the amount of additional attention provided to the operator to maintain proficiency. An alarm station is a location for command and control by leadership, simply observing the presence of the supervisor near the operators is not indicative of the oversight characteristic. Accurately assessing this characteristic may require analysts to complete a baseline review of oversight duties to ensure rated activities are above and beyond the directed amount of oversight required by the responsible authority. The oversight characteristic focuses on active oversight actions between an experienced authority and operator. Oversight is not always a senior position over a subordinate position.

A properly trained system operator should be able to acknowledge, assess, and dispatch to alarm stimulus correctly without the assistance or direction of oversight. Analysts should be able to assess the oversight characteristic by rating each operator separately and averaging the score. Rating the individual is vital for this characteristic as operators can be entry level or ten-year veterans with established levels of experience, rating this characteristic based solely on one or the other could result in an inaccurate score. Rating individually is also valuable as a causal analysis tool to identify corrective actions.

The analyst gathers data to support the score for this characteristic by observing daily operations, reviewing operator performance documentation, and interviewing the supervisors and support provided to the operators. Another source of data should be observation of the operators during testing activities. Analysts or responsible security authorities identify weaknesses in operational performance not highlighted during daily routine operations. As with all other assessments, analysts should be sensitive in data collection and ensure they are assessing and rating the correct characteristic. Analysts should use Table 7.7, Oversight Scale, below to rate this characteristic. This characteristic is part of the operator scale.

Table 7.7 Oversight Scale

Oversight	
Scale	Description
1	The operator cannot operate the system unsupervised. During a normal workweek, oversight is present more than 75% of the time to oversee operator activities.
2	The operator can operate the system with little oversight. During a normal workweek, oversight is present approximately 50% of the time to oversee operator activities.
3	The operator can operate the system with little or no oversight. During a normal workweek, oversight is present less than 25% of the time to oversee operator activities.
4	The operator provides experience to others operating the system. No oversight is present and the operator assists others in operating the system or in developing procedures and training documentation on operating the system.
5	Although responsible for system operation, the operator is considered a subject matter expert and frequently conducts On-the-Job-Training (OJT) to train others to operate the system.

7.8. Performance Testing Alarm Rate

Alarm rates and operator proficiency are vital during actual alarm events. No other characteristic reflects this like the Performance Testing Alarm Rate. The primary purpose of this alarm rate is to identify how a full-scale exercise, such as a FoF, increases the alarm rate. This testing or training event should be representative of the actual numbers of aggressive forces that assault the facility as defined by the graded security protection policy and thereby indicate the number of alarms that analysts should reasonably expect during that kind of event. The system and operators should be able to address all alarms and dispatch protective force personnel as required, in addition to providing security awareness to responding forces. This characteristic rates the system's alarm levels separate from the operator performance test or oversight characteristics which are used to rate the operator's

success at addressing the number of alarms generated by these tests. Analysts may identify the alarm rate during a simulated event such as a tabletop or measured directly during an actual test event.

As with the operator performance test characteristic, analysts should rate this characteristic during exercise activities supported by alarm logs and acknowledgement time data. Unlike the operator performance test characteristic, analysts should only rate this characteristic during extensive exercises such as a FoF. Rating only at this time identifies the most alarms that are likely to be present and truly assesses the effects on the system. Stressing the alarm system in this manner should be an objective of the test itself. Analysts should be aware that FoF activities are dynamic and analysis is different based on the scenario. This characteristic is part of the system scale.

As with rating using other scales, the analyst may touch on many points within the scale as presented in Table 7.8, Performance Testing Alarm Rate Scale. Again, analysts are required to rate the various effects and justify their rating to ensure the responsible security authority can identify and address weaknesses. This scale is unique in that it presents analysts with the description of the operator's abilities but focuses on the system's alarms. The scale focuses in this manner to indicate the effect of the numerous alarms and communications required during response activities. Rating in this manner is also necessary as these alarms are intentionally part of the event and analysts cannot rate it as outlined in the Alarm Frequency characteristic. A well-engineered and maintained system should present only a limited number of pathways that operators can easily respond to while possibly conducting other ancillary duties. Analysts should observe numerous testing events with different system operators, as a poorly performing system appears to be effective given talent and experience of operators.

Table 7.8 Performance Testing Alarm Rate Scale

Performance Testing Alarm Rate	
Scale	Description
1	During LSPT/FoF testing, the alarms are so numerous operators cannot support the responding forces and the system crashes.
2	During LSPT/FoF testing, operators cannot respond to communication requirements and some alarms (less than five) are not addressed, alarm acknowledgement time increases by 2 minutes or more and alarm rate increases by 75%.
3	During LSPT/FoF testing, communication by operators is brief, not all information is gathered, and a few alarms (less than three) are not addressed, alarm acknowledgement time increases by 1 minute or more and alarm rate increases by 50%.
4	During LSPT/FoF testing, operators can respond to the system and communicate as required with little loss of information, no alarms are ignored or missed, alarm acknowledgement time increases by 30 seconds or more and alarm rate increases by 25%.
5	During LSPT/FoF testing, operators complete all communication, address all alarms, and make all required notification with no loss of information. There is no noticeable increase in alarm acknowledgement time.

7.9. System Interface

Analysts use the system interface information in Table 7.9, System Interface Scale, to assess the number and different types of operational interfaces of the alarm system and other security systems. A simple system with familiar interface lends to quicker operation with clear alarm display information and limited interface requirements reducing the difficulty in operation and responding to the system. Counterintuitively, the system interface is not a scale of system age, as new systems functionality increases with capability and options, complexity grows.

Analysts should review the alarm system and supporting systems separately and group the results. If an operator is responsible for multiple systems with different interfaces and operating characteristics, more specialization and training is required. In addition, operators will need to be very skilled to identify issues with the system should so many supporting systems be employed. Separate review is necessary for causal analysis and corrective actions as well as the score as a whole. As with all topics affected by multiple sources, analysts are required to document the separate levels of each system analyzed and average the resultant score.

The analyst identify data for this scale by observing the operations, reviewing operator manuals, reviewing training lesson plans, and reviewing operating procedures for the system. Observation of alarm response activities highlight the level of effort required by the operator to respond to system alarms, identify the cause, and dispatch as necessary. This characteristic is also reflected in other characteristics such as oversight, performance testing alarm rate, and operator performance tests. As always, analysts should be very cautious to ensure the characteristic reviewed is specific and not encroaching on other topics. This topic is part of the system scale.

Table 7.9 System Interface Scale

System Interface	
Scale	Description
1	Each operator is responsible for more than five separate alarm and observation systems. Operation requires knowledge of software coding, typing for searches to locate alarms, switching between software, multiple button clicks to perform any function in the system.
2	Each operator is responsible for more than four separate alarm and observation systems. Operation requires knowledge of numerous (four or more) software types, typing for searches to locate alarms, switching between software, multiple button clicks to perform any function in the system.
3	Each operator is responsible for more than three separate alarm and observation systems. Operation requires knowledge of some (three or more) software types, menus are drop down and system has numerous functions that the operator does not use, system requires little keyboard use to operate effectively.
4	Each operator is responsible for more than two separate alarm and observation systems. Operation requires knowledge of a couple software types, menus are drop down and some additional functionality remains but is not employed by the operator, two-button mouse click to operate most functions, some limited keyboard use required.
5	Each operator is responsible for two or fewer alarm and observation systems. Operation requires knowledge of a single software type, menus and system operations have no extraneous features, all options used by the operator, menus are limited or the entire system can be operated from basic button interface, one button mouse click to operate any function of the system. No keyboard use to operate the system after login.

7.10. Technical Knowledge

A system requiring specialized training and extensive focused instruction on maintenance procedures is limited in resources to address problems when they arise. Specialization increases the difficulty of communicating issues and completing repairs in a timely manner. An effective detection and assessment system is simple to maintain and difficult to defeat. This scale outlines the impact of specialized training or tribal knowledge on system operability and reliability. As with most scales in this process, the analyst should assign a qualitative score based on quantitative input.

Analysis of the characteristic is extensive as each system and each technician requires review. The analyst should complete interviews with technicians and operators, conduct a review of system maintenance documents, and assess training and lesson plans. As with all topics, the more granular

the initial score the easier causal analysis and corrective actions are attained. This topic is part of the system scale. The technical knowledge characteristic score is a reflection of the maintenance, system interface, and repairability characteristics. As with other characteristics, analysts should focus their rating activities on the characteristic and limit the effect or influence of other characteristics.

As the primary interest of this scale is the IDS, this should be the focus of the assigned score as presented in Table 7.10, Technical Knowledge Scale, below. However, any system for which the operator is responsible should also be assessed as these systems impact their ancillary duties and ultimately their trust in the system as a whole. Although this topic appears to be very similar to the maintenance characteristic, the difference is not the completed maintenance but the effort and knowledge required to complete it. For example, if a talented and experienced team of technicians with very specialized knowledge keeps the system running smoothly, maintenance scores very well, hiding the impact of the specialized/tribal knowledge necessary to keep the system operational. Scoring this characteristic separately is necessary, especially if the above situation is occurring. By identifying this situation early, the responsible security authority addresses the issue prior to any shift in staffing levels. If the technicians assessed are not employees at the site then the characteristic may not be accurate, it is recommended analysts not use this characteristic to assess the technician until the warranty has expired or the contract has changed. However, the time to repair, length of compensatory measures, and cost of parts should still be tracked as part of the Maintenance Rate and Repairability characteristic.

Table 7.10 Technical Knowledge Scale

Technical Knowledge	
Scale	Description
1	Specialized school required available from only one source, personnel conducting maintenance and testing still require six months or more of additional OJT prior to maintaining the system, use of external technicians or specialists is commonly employed with almost every repair (50% or greater)
2	Specialized school required, personnel conducting maintenance and testing still require six months of additional OJT prior to maintaining the system, use of external technicians or specialists is commonly employed with almost every repair (25% or greater)
3	Local training required, personnel conducting maintenance and testing still require one month of additional OJT prior to maintaining the system, use of external technicians or specialists is infrequent with repair (10% or greater)
4	Minor localized training or basic OJT is all that is required to get a system technician skilled in completing repair and planned maintenance, use of external technicians or specialists is rare (less than 10%)
5	New employees can readily perform repair and maintenance without additional training

7.11. Testing Frequency

Along with the level of testing stringency to confirm IDS health is the frequency of testing. However this scale, as identified in Table 7.11, Testing Frequency Scale, not only rates the actual amount of testing but the different organizations responsible. The amount of testing is important to allow organizations to detect and address problems. The value of different organizations testing is that each group employs different talent, perspective, and intent. The more review of a topic the more likely the system is to be effective. This testing frequency applies to not only the sensors but also the system and supporting characteristics. As such, the analyst should rate system test frequency separately for a summary score as a total. This characteristic is part of the system scale.

More engagement by responsible personnel increases the faith of the operator in the system. Other characteristics that reflect or influence the testing frequency are maintenance rate and repairability

and operator performance test. The testing frequency should be easy for the analyst to identify in test plans, reports, and security plans. Analysts gather additional information by conducting interviews, assessments, and observations of testing activities.

Table 7.11 Testing Frequency Scale

Testing Frequency	
Scale	Description
1	Performance testing is only completed by an outside agency when the company/site is being audited. Testing is conducted once a year by the agency
2	One local entity conducts performance testing and an outside agency when the company/site is being audited. The groups conduct testing once a quarter
3	Two separate entities from different companies or responsible organizations, conduct testing along with outside agencies during audits. The groups conduct testing once a month
4	Testing is conducted by numerous groups (system technicians, facility representatives, security officers, performance testing) from more than two associations, and by outside agencies during audits. The groups conduct testing weekly
5	Testing is conducted by numerous groups (system technicians, facility representative, security officers, performance testing) from more than two associations, including another group not associated with the facility or site, and outside agencies during audits. The groups conduct testing daily

7.12. Training Test Scores

Test scores are a reflection of the effectiveness of initial and continuous training provided by sites for the system operators. Alone, this characteristic as presented in Table 7.12, Training Test Scores Scale, below, does not indicate a success or failure on the part of the operator, but is a piece of the whole picture concerning IDS operability and reliability. Test scores are a quantitative value derived by the analyst by averaging test score results. As with characterizing operator performance tests, the analyst should ensure testing included in the analysis reflects system operations and ancillary duties. Testing scores outlined below are based on criterion-referenced grading systems as defined by the Department of Education.[4]

The Training Department or Section should have and retain all topic scores for analysis, each operator should be scored separately to identify those individuals in need of additional attention or training. Analysts should review training topics and results separately to ensure granularity of the scale. This score, among others, will fluctuate with experience and test results. As such, analysts should be aware and document their data collection and analysis process. Analysts should always be aware of disparities in results conducted by different assessment groups. Slight differences in the conduct of the test, the objectives, or testing methodologies may create differences in the results. The testing criteria and results should be reviewed and, in some cases, the analyst could conduct their own testing or observe numerous testing events by multiple sources to ensure data integrity and consistency. This characteristic is part of the operator scale.

Table 7.12 Training Test Scores Scale

Training Test Scores	
Scale	Description
1	Average test scores less than 65%
2	Average test scores between 65% - 75%
3	Average test scores between 75% - 85%

[4] Department of Education (2008) Structure of the U.S. Education System: U.S. Grading Systems. Located at www2.ed.gov/about/offices/list/ous/international/usnei/us/grading.doc

Training Test Scores	
Scale	Description
4	Average test scores between 85% - 95%
5	Average test scores between 95% - 100%

7.13. Z-Score

Referring back to the alarm frequency characteristics, another useful statistical tool used to normalize the data for evaluation is the Z-Score. Analysts use the Z-Score to determine the most current sensor performance, and compare these results with the statistically developed "normal" or "typical" performance of that same sensor.

Analyst shall complete calculations for the Z-Scores at the same frequency as calculation of the Alarm Frequency characteristic. The result should be that the sensor's Z-Scores indicate how many standard deviations away from the mean the current day's performance is.

The Z-Score provides the relative position of a score with respect to the mean. The Z-Score is calculated by subtracting the mean from the subject score and dividing by the standard deviation, thus:

$$z = (x - xbar)/(s)$$

The Z-Score is a number that reveals how many standard deviations a score is from the mean of the scores.

For example, using a Daily Individual Alarm Rate of 34 (x) unresolved alarms per hour, Weekly Individual Alarm Rate Mean (xbar) of 28.7 and a Weekly Individual Alarm Rate Standard Deviation (s) of 8.9, the associated Z-Score is:

$$z = (34 - 28.7) / 8.9 = 0.60 \text{ (rounded)},$$

This indicates that the daily rate is within 0.60 standard deviations from the mean for this sensor, based on the sensor's historical performance.

Assuming that a sensor is functioning according to a normal distribution of performance variation (only random variation in its performance), scores that produce ($-2 < \text{Z-Score} < 2$) are considered normal (or, expected) for this sensor. Generally, analysts should expect a larger amount of data to produce a mean and standard deviation representing more accurate performance of the sensor.

Analysts should calculate baseline means and standard deviations each week for each sensor, using all of the previous weekly data scores. Using these values, analysts should calculate a Z-Score for any individual daily or weekly alarm rate and determine whether the sensor is deviating from normal performance with statistical significance. Analysts should evaluate scores with the Z-Score method to determine whether there are an excessive number of alarms that could degrade system operator responsiveness.

Analysts should remove from calculations scores for which extreme outlying results to avoid skewing results if a firm cause for the outliers is determined and remedied. This is a quantitative scale as identified in Table 7.13, Z-Score Scale, below. As discussed above, each sensor should have a separate rating for causal analysis and corrective action. Comparing sensors of the same style may also highlight a poorer performing sensor rather than reviewing the individual sensors score.

However, for the Z-Score, analysts should generate a total value for the system to average the characteristic score. The Z-Score is part of the system scale.

Table 7.13 Z-Score Scale

Z-Score	
Scale	Description
1	-3 < Z-Score < 3
2	-2.5 < Z-Score < 2.5
3	-2 < Z-Score < 2
4	-1.5 < Z-Score < 1.5
5	-1 < Z-Score < 1

8. EXAMPLE ANALYSIS

The example provided below only covers the scale rating and matrix presentation. For a complete analysis report, analysts should submit narrative describing the assessment process for each characteristic rated highlighting the reasoning leading to the characteristic rating.

After much analysis by various subject matter experts, a site protecting Category I SNM has completed their analysis of the IDS characteristics. Their scores along with the average rating of each scale is presented below in Table 8.1, System Scale and 8.2, Operator Scale.

Table 8.1 System Scale

System	
Scale	Description
3	Alarm Frequency
2	Ergonomics
3	Maintenance Rate and Repairability
3	Performance Testing Alarm Rate
3.5	System Interface (Considered a minor characteristic for this example)
3	Technical Knowledge
4	Testing Frequency
3	Z-Score
Average	3.1

Table 8.2 Operator Scale

Operator	
Scale	Description
3	Ancillary Duties
4.5	Average Alarm Acknowledgement Time (Considered a major characteristic for this example)
2	Operator Performance Test
4	Oversight
2	Training Test Scores
Average	3.1

Charting these scale results on Figure 8.1, Example Analysis Matrix, identifies that the system falls slightly above the horizontal centerline between sectors II and I and with slightly above average system health.

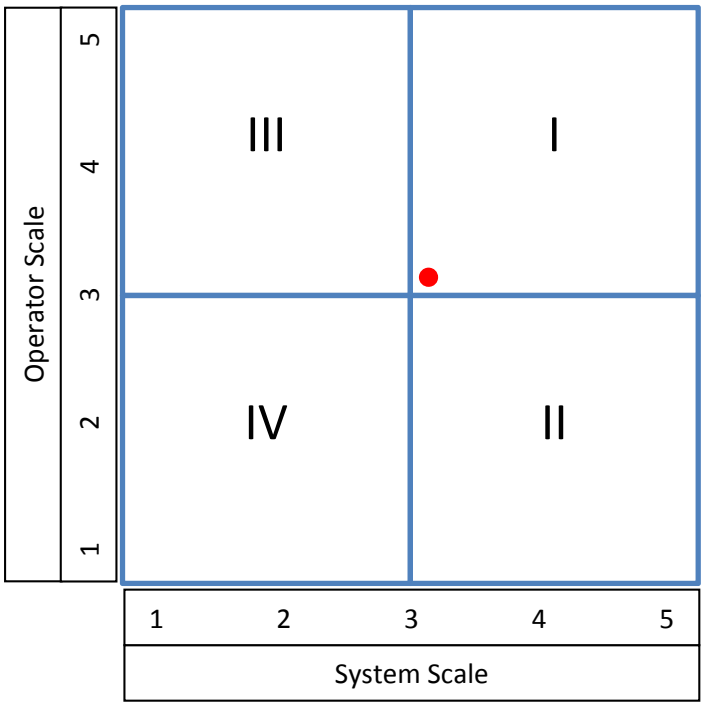

Figure 8.1 Example Analysis Matrix

The analyst plots the matrix point again on the overlay on Figure 8.2, Example Operability and Reliability Matrix. This matrix identifies that although the operator and system are performing adequately for an average system, the IDS is not effective for protecting vital national assets.

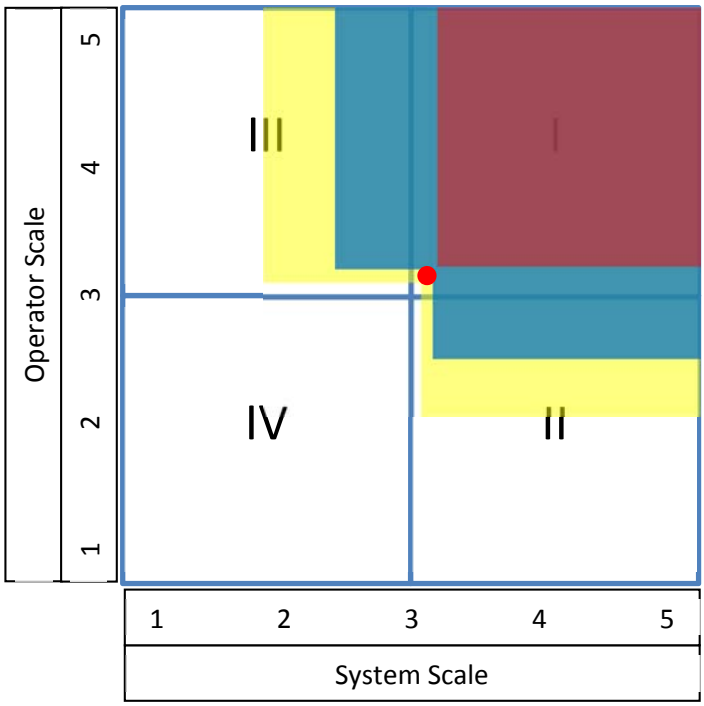

Figure 8.2 Example Operability and Reliability Matrix

The primary analyst should report the results and the responsible security authorities should identify the lowest performing characteristics and address the shortcomings. In this example, both the operator and system have characteristics that the responsible security authorities should address to ensure the IDS is operating effectively.

9. CHARACTERISTIC WEIGHTING

To account for some characteristics having more or less impact on the operability and reliability of the IDS, Table 9.1, Weighted Average Effect, presents weighted scoring scales. Analysts may apply these identifiers to tailor the scoring values of specific characteristics. If weighted characteristics are used, the responsible security authority and analysts should identify the major and minor characteristics prior to final scoring. Analysts shall capture the decision to rate a characteristic as major or minor along with supporting information within the report.

Table 9.1 Weighted Average Effect

Major characteristic scale	Designator	Average characteristic scale	Designator	Minor characteristic scale
0.5	Poor	1	Poor	1.5
1.5	Moderate	2	Moderate	2.5
3	Fair	3	Fair	3
4.5	Good	4	Good	3.5
5.5	Excellent	5	Excellent	4.5

Weighting the characteristic works as the system and operator scales are an average. The larger or smaller the range of the score numbers the more or less impact to the average. Table 9.2, Weighted Average Example One, below provides an example of the effect of weighting on an average.

Table 9.2 Weighted Average Example One

Characteristic	Average scores, not weighted	Weighted Scores
Ergonomics (Minor Characteristic)	5	4.5
Alarm Frequency (Major Characteristic)	1	0.5
Scale Results (Average)	3	2.5

As provided in Table 9.2 above, if a characteristic is identified as minor and is operating well above expected levels rather than receiving a 5 score it should receive a 4.5, reducing the impact of the characteristic when computing the average. By contrast, a major characteristic operating well below expected levels should receive a 0.5 score rather than a 1, increasing the impact to the average. Table 9.3, Weight Average Example Two, below shows the impact of the results in reverse and further highlights how a major characteristic score overrides a minor characteristic score.

Table 9.3 Weighted Average Example Two

Characteristic	Average scores, not weighted	Weighted Scores
Ergonomics (Minor Characteristic)	1	1.5
Alarm Frequency (Major Characteristic)	5	5.5
Scale Results (Average)	3	4.3

Table 9.4, Operator Scale Weighted Average Example, reveals the impact of weighted average over an entire scale.

Table 9.4 Operator Scale Weighted Average Example

Characteristic	Scores not weighted	Weighted Scores
Ancillary Duties	3	3
Average Alarm Acknowledgement Time	3	3
Operator Performance Test (Major Characteristic)	1	0.5
Oversight (Minor characteristic)	5	4.5
Training Test Scores	4	4
Scale Results (Average)	3.2	3

The example in Table 9.4 above highlights how the operator performance test characteristic identified as a major characteristic assigned a low score pulls down the average. This weighted characteristic with a lower score overpowers the excellent scoring oversight characteristic identified as a minor characteristic when calculating the average. Analysts should use the weighted characteristic sparingly and have documentation supporting the decision. The intent of using a weighted characteristic is to ensure less important characteristics performing well, do not hide the impact of an important characteristic performing poorly.

Common major characteristics likely include alarm frequency, performance testing alarm rate, Z-score, and average alarm acknowledgement time. Common minor characteristics likely include ergonomics, operator performance test, and training test scores.

10.ROOT CAUSE IDENTIFICATION

Correcting the immediate cause of a problem may provide a temporary improvement to the issue. However, to ensure the characteristic operates as necessary requires correcting the root cause of the problem. Identifying and correcting this base cause requires a methodical analysis and corrective action process. As part of the quality assurance and assessment programs, sites shall have an established process of causal analysis and corrective action. This technical standard will not define causal analysis and corrective action processes, but rather identifies probable topics influencing the characteristic. As stated above the more granular the initial analysis the easier the cause identification and correction.. .

10.1. Alarm Frequency

Installation and maintenance are the primary cause of unacceptable levels of alarm frequency. Installing the wrong type of sensor will allow environmental or biological activities to have a greater impact. Maintenance adjustments to sensor position and settings may also affect detection and alarm rates. The best methods to identify these issues are through assessment, observation, interviews, and performance testing.

10.2. Ancillary Duties

Resources and requirements affect the number and type of ancillary duties. The amount of staffing resources controls the number of operators and training opportunities. Additionally, the talent and skill level of the operator can hide or highlight the impact of the additional duties. Conducting assessments and interviews identifies the cause and correction of ancillary duties problems.

10.3. Average Alarm Acknowledgement Time

Causes of unacceptable acknowledgement times include training and experience of the operator, difficulty in systems operations, the number and difficulty of ancillary duties, the number of systems requiring operation, and the system engineering (e.g., failed to engineer to address external effects such as the environment or animals). In addition, Analysts should review the categorization time to ensure that the operator is not frequently categorizing the alarm within a second of acknowledging the alarm generally indicating an inattentive operator. Analysts can conduct causal analysis through assessments, observations, and interviews.

10.4. Ergonomics

The impact of ergonomics begins with the initial installation and changes as authorities add systems and make changes. Analysts should identify causes and corrections through observation of the facility and employ trained safety professionals as subject matter experts to conduct reviews as necessary.

10.5. Operator Performance Tests

Analysts should review testing frequency and types of testing when addressing causal analysis for operator performance tests. In addition, the analysts should also review the experience of the operators and the history of testing program as new programs and new operators could affect the outcome of testing. Analysts can collect this data through assessment, observation of testing activities, interviews with performance testing personnel and operators, and document review of testing records.

10.6. Maintenance Rate and Repairability

Planning and installation are initial factors to the causes of a poorly scoring maintenance characteristic. A properly planned and installed system should require limited maintenance beyond planned maintenance of cleaning and inspection. Changing system requirements make problems in this topic worse. Naturally, older systems will commonly have more maintenance issues as they mature. The truly concerning problem is a brand new system that requires maintenance beyond the installation and break in time. Analysts should review documentation, diagrams, and maintenance records as the first steps to identifying cause. Included in this document review should be observation of maintenance efforts and interviews with the technicians completing repairs. Maintenance impact can also be influenced by the experience and training of the technician. Concerning repairability aspects, more than just a measure between the installation date and the current date, the repairability aspect also rates the impact of parts and the difficulties in maintaining the system. In addition to noting the installation date, analysts should review the availability and costs of repair parts and materials. Other sources of information include the technicians who understand the resources and limitations. Interviews and document reviews should identify concerns and corrections in this characteristic

10.7. Oversight

Analyst can identify sources of causal analysis and corrective action within training, talent, experience, and system exposure. Management's initial response to most security issues is to require more attention by leadership. This may fix the problem initially and provide the operator time to increase their talent and experience to mitigate the actual concern. Analysts should be aware that a lower score on the oversight characteristic could be a short-term corrective action of other issues. Analysts will have to review system alarm rates, training records, testing results, and system characteristics to identify whether or not the addition of the supervisor is actually a mitigating factor. Observations, interviews, and document reviews should provide data to identify the cause and correction.

10.8. Performance Testing Alarm Rate

Installation and maintenance are the primary causes of unacceptable testing alarm rate frequencies. As systems expand and provide more multilayered coverage, the frequency of alarms will likely increase. Testing events will provide data for causal analysis and corrective action. Analysts should also review security response plans as responding forces moving through detection zones are also likely to increase alarm rates. As discussed in the characteristic section any review of alarm rates should reference the adversary attack plan as different pathways may produce different levels of alarm.

10.9. System Interface

Analysts should begin causal analysis and corrective action of system interface through review of installation and maintenance documentation. Analysts should review historic documents and complete interviews with operators and technicians to identify the cause and corrective action. Interviews are important to this topic as they provide insight to operations and the limitations based on manufacturing and engineering requirements.

10.10. Technical Knowledge

A system with basic technical requirements purchased off the shelf from a known open source provider can be part of an initial purchase and evolve into a very specialized system. For causal analysis purposes, analysts should review the initial installation and interview operators and technicians to identify if this system is historically difficult to maintain or if an evolution changed the original design. As with all analysis, documentation review is vital to identify if technical

specialization is necessary or a product of resource constraints. Other data for the analyst to review includes the training provided from the system manufacturer.

10.11. Testing Frequency

Interviews and document review concerning testing protocols and company interfaces should define testing levels. Testing can be limited to a few organizations for a number of reasons. This characteristic is an administrative or resource issue and can highlight lack of communication and coordination between organizations.

10.12. Training Test Scores

Analysts should review documentation on training including training records, lesson plans, and instructor's reports to identify any trends. Analysts should be aware, however, that not all weaknesses in training scores point back to the training organization. Analysts can also identify impacts to training in the time and exposure of the operator to the system. As with all topics, influences such as mission, resources, management, administrative controls, or resources may have an effect on the scores of the operator.

10.13. Z-Score

Very similar to the alarm frequency characteristic from which this statistic is derived, the installation and maintenance of the sensor affects sensor functionality and sensitivity. In addition, the age of the sensor may have an effect as the capabilities may degrade with time. Analysts should follow the same path as addressing the alarm frequency and review the same source of information to address the concerns of this characteristic.

CONCLUDING MATERIAL

Review Activity
EM
Policy (AU-51)
EHSS
MA
NE
NNSA
SC
EA

Preparing Activity
Office of Security Policy (AU-51)

Field, Site and Operations Offices
ID
NNSA Service Center
ORO
RL
SRO
INL
LASO
LLSO
NSO
OR
RLSO
SRSO
SSO
NPO

Project Number
P2015-03

External Agency
N/A