Designation: E2853 – 12

Standard Test Method for Evaluating Emergency Response Robot Capabilities: Human-System Interaction (HSI): Search Tasks: Random Mazes with Complex Terrain[1]

This standard is issued under the fixed designation E2853; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 *Purpose:*

1.1.1 The purpose of this test method, as a part of a suite of human-system interactions (HSI) test methods, is to quantitatively evaluate a teleoperated ground robot's (see Terminology E2521) capability of searching in a maze.

1.1.2 Teleoperated robots shall possess a certain set of HSI capabilities to suit critical operations such as emergency responses, including enabling the operators to search for required targets. A passage that forms on complex terrains and possesses complex and visually similar branches is a type of environments that exists in emergency response and other robotically applicable situations. The complexity often poses challenges for the operators to teleoperate the robots to conduct searches. This test method is based on a standard maze and specifies metrics and a procedure for testing the search capability.

1.1.3 Emergency response robots shall enable the operator to handle many types of tasks. The required HSI capabilities include search and navigation on different types of terrains, passages, and confined spaces. Standard test methods are required to evaluate whether candidate robots meet these requirements.

1.1.4 ASTM E54.08.01 Task Group on Robotics specifies a HSI test suite, which consists of a set of test methods for evaluating these HSI capability requirements. This random maze searching test method is a part of the HSI test suite. The apparatuses associated with the test methods challenge specific robot capabilities in repeatable ways to facilitate comparison of different robot models as well as particular configurations of similar robot models. (See Fig. 1.)

1.1.5 The test methods quantify elemental HSI capabilities necessary for ground robots intended for emergency response applications. As such, based on their particular capability requirements, users of this test suite can select only the applicable test methods and can individually weight particular test methods or particular metrics within a test method. The testing results should collectively represent a ground robot's overall HSI capability. The test results can be used to guide procurement specifications and acceptance testing for robots intended for emergency response applications.

NOTE 1—The teleoperation performance is affected by the robot's as well as the operator's capabilities. Among all the standard test methods that ASTM E54.08.01 Task Group on Robotics has specified, some depend more on the former while the others on the latter, but it would be extremely hard to totally isolate the two factors. This HSI test suite is specified to focus on evaluating the operator's capabilities of interacting with the robotic system, whereas a separately specified sensor test suite, including Test Method E2566, focuses on the robots' sensing capabilities.

NOTE 2—As robotic systems are more widely applied, emergency responders might identify additional or advanced HSI capability requirements to help them respond to emergency situations. They might also desire to use robots with higher levels of autonomy, beyond teleoperation to help reduce their workload—see NIST Special Publication 1011-II-1.0. Further, emergency responders in expanded emergency response domains might also desire to apply robotic technologies to their situations, a source for new sets of requirements. As a result, additional standards within the suite would be developed. This standard is, nevertheless, standalone and complete.

1.2 *Performing Location*—This test method shall be performed in a testing laboratory or the field where the specified apparatus and environmental conditions are implemented.

1.3 *Units*—The values stated in SI units are to be regarded as the standard. The values given in parentheses are not precise mathematical conversions to inch-pound units. They are close approximate equivalents for the purpose of specifying material dimensions that are readily available to avoid excessive fabrication costs of test apparatuses while maintaining repeatability and reproducibility of the test method results. These values given in parentheses are provided for information only and are not considered standard.

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

[1] This test method is under the jurisdiction of ASTM Committee E54 on Homeland Security Applications and is the direct responsibility of Subcommittee E54.08 on Operational Equipment.

Current edition approved Feb. 1, 2012. Published April 2012. DOI: 10.1520/E2853-12.

Copyright © ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959. United States

FIG. 1 HSI: Search Tasks: Random Maze Illustration

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E2521 Terminology for Urban Search and Rescue Robotic Operations

E2566 Test Method for Determining Visual Acuity and Field of View of On-Board Video Systems for Teleoperation of Robots for Urban Search and Rescue Applications

E2592 Practice for Evaluating Cache Packaged Weight and Volume of Robots for Urban Search and Rescue

2.2 *Additional Documents:*

National Response Framework U.S. Department of Homeland Security[3]

NIST Special Publication 1011-I-2.0 Autonomy Levels for Unmanned Systems (ALFUS) Framework Volume I: Terminology, Version 2.0[4]

NIST Special Publication 1011-II-1.0 Autonomy Levels for Unmanned Systems (ALFUS) Framework Volume II: Framework Models, Version 1.0[4]

3. Terminology

3.1 *Definitions:*

3.1.1 *abstain, v*—the action of the manufacturer or designated operator of the testing robot choosing not to enter the test. Any decision to take such an action shall be conveyed to the administrator before the test begins. The test form shall be clearly marked as such, indicating that the manufacturer acknowledges the omission of the performance data while the test method was available at the test time.

3.1.1.1 *Discussion*—Abstentions may occur when the robot configuration is neither designed nor equipped to perform the tasks as specified in the test method. Practices within the test apparatus prior to testing should allow for establishing the applicability of the test method for the given robot.

3.1.2 *administrator, n*—person who conducts the test—The administrator shall ensure the readiness of the apparatus, the test form, and any required measuring devices such as stopwatch and light meter; the administrator shall ensure that the specified or required environmental conditions are met; the administrator shall notify the operator when the safety belay is available and ensure that the operator has either decided not to use it or assigned a person to handle; and the administrator shall call the operator to start and end the test and record the performance data and any notable observations during the test.

3.1.3 *emergency response robot, or response robot, n*—a mobile device deployable to perform operational tasks at operational tempos to assist the operators to handle a disaster.

3.1.3.1 *Discussion*—A response robot is designed to serve as an extension of the operator for gaining improved remote situational awareness and for accomplishing the tasks remotely through the equipped capabilities. The use of a robot is designed to reduce risk to the operator while improving effectiveness and efficiency of the mission. The desired features of a response robot include: the ability to be rapidly deployed and remotely operated from an appropriate standoff distance and to be mobile in complex environments, sufficiently hardened against harsh environments, reliable and field serviceable, durable and/or cost effectively disposable, and equipped with operational safeguards.

3.1.4 *fault condition, n*—a certain situation or occurrence during testing whereby the robot either cannot continue without human intervention or has performed some defined rules infraction.

3.1.4.1 *Discussion*—Fault conditions include robotic system malfunction such as de-tracking, task execution problems such as excessive deviation from a specified path, or uncontrolled behaviors and other safety violations which require administrative intervention.

3.1.5 *full-ramp terrain element, n*—1.2 by 1.2 m (4 by 4 ft) surface ramp with 15° slope using solid wood support posts

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] Available from Federal Emergency Management Agency (FEMA), P.O. Box 10055, Hyattsville, MD 20782-8055, http://www.fema.gov/emergency/nrf/.

[4] Available from National Institute of Standards and Technology (NIST), 100 Bureau Dr., Stop 1070, Gaithersburg, MD 20899-1070, http://www.nist.gov/el/isd/ks/autonomy_levels.cfm.

with angle cuts. The material used to build these elements shall be strong enough to allow the participating robots to execute the testing tasks.

3.1.5.1 *Discussion*—The material that is typically used to build these elements, oriented strand board (OSB) is a commonly available construction material. The frictional characteristics of OSB resemble that of dust covered concrete and other human improved flooring surfaces, often encountered in emergency responses. Solid wood posts with 10 by 10 cm (4 by 4 in) cross-section dimensions typically support the ramped surface.

3.1.5.2 *Discussion*—Elements similar to this type are used, sometimes mixed and assembled in different configurations, to create various levels of complexities for such robotic functions as orientation and traction.

3.1.6 *human-scale, adj*—the environments and structures typically negotiated by humans, although possibly compromised or collapsed enough to limit human access. Also, that the response robots considered in this context are in a volumetric and weight scale appropriate for operation within these environments.

3.1.6.1 *Discussion*—No precise size and weight ranges are specified for this term. The test apparatus specifies the confined areas in which to perform the tasks. Such constraints limit the overall sizes of robots to those considered applicable to emergency response operations.

3.1.7 *maze, n*—a network of passages interconnected without any repetitive order of opening and closing directions and meant to challenge robotic navigation from the designed starting and end points.

3.1.8 *operator, n*—person who controls the robot to perform the tasks as specified in the test method; she/he shall ensure the readiness of all the applicable subsystems of the robot; she/he through a designated second shall be responsible for the use of a safety belay; and she/he shall also determine whether to abstain the test.

3.1.8.1 *Discussion*—An emergency responder would be a typical operator in emergency response situations.

3.1.9 *operator station, n*—apparatus for hosting the operator and her/his operator control unit (OCU, see NIST Special Publication 1011-I-2.0) to teleoperate (see Terminology E2521) the robot. The operator station shall be positioned in such a manner as to insulate the operator from the sights and sounds generated at the test apparatuses.

3.1.10 *repetition, n*—robot's completion of the task as specified in the test method and readiness for repeating the same task when required.

3.1.10.1 *Discussion*—In a traversing task, the entire mobility mechanism shall be behind the START point before the traverse and shall pass the END point to complete a repetition. A test method can specify returning to the START point to complete the task. Multiple repetitions, performed in the same testing condition, may be used to establish the tested capability to a certain degree of statistical significance as specified by the test sponsor.

3.1.11 *test event, or event, n*—a set of testing activities that are planned and organized by the test sponsor and to be held at the designated test site(s).

3.1.12 *test form, n*—a collection of data fields or graphics used to record the testing results along with the associated information. A single test form shall not be used to record the results of multiple trials.

3.1.13 *test sponsor, n*—an organization or individual that commissions a particular test event and receives the corresponding test results.

3.1.14 *test suite, n*—designed collection of test methods that are used, collectively, to evaluate the performance of a robot's particular subsystem or functionality, including HSI, manipulation, sensors, energy/power, communications, logistics, safety and operating environment, and aerial or aquatic maneuvering.

3.1.15 *testing target, or target, n*—a designed physical feature to be used by the testing robotic subsystem for evaluating the subsystem capabilities. The feature may be an operationally relevant object, a notional object, or one designed specifically for exercising the subsystem features to its full extent.

3.1.16 *testing task, or task, n*—a set of activities well defined in a test method for testing robots and the operators to perform in order for the system's capabilities to be evaluated according to the corresponding metric(s). A test method may specify multiple tasks. A task corresponds to the associated metric(s).

3.1.17 *trial, n*—the number of repetitions to be performed for a test to reach required statistical significance. The repetitions may be recorded on a single test form.

3.2 Terminology E2521 lists additional definitions relevant to this test method.

4. Summary of Test Method

4.1 The search task for this test method is for a teleoperated robot to traverse in a specified maze to completely cover and clear specified targets. Standard hazardous materials (HAZMAT) labels shall be used as the targets. Coverage of a target is defined as when the operator correctly detects the existence of the target through the video images displayed on the Operator Control Unit (OCU) and conveys such existence to the administrator. Clearance of a target is defined as when the operator correctly conveys the names of at least three out of the following four features on the label: color, icon, number, and words to the administrator. When the operator correctly conveys one or two of the features, it is categorized as coverage.

4.2 A robot's physical capabilities might affect the test operator's abilities in performing the tasks. The test sponsor can elect to weight the coverage metric higher over clearance to reduce the effects of the cameras and/or the lights when her/his primary concern is the operator's capability. Another way of handling the issue of the operator's versus the robotic physical capabilities is for the test sponsor to assign the

respective reliability and confidence values for the two metrics according to the sponsor's emphases. Section 8 specifies these effects.

4.3 The testing robot shall return to the starting point at the end of the test. The starting point is specified by the test sponsor and is not notified to the operator until at the beginning of the test.

4.4 Teleoperation shall be used from the operator station specified by the administrator to test the robots using an OCU provided by the operator. The operator station shall be positioned and implemented in such a manner as to insulate the operator from the sights and sounds generated at the test apparatus.

NOTE 3—Separate, autonomous search test methods will be separately specified in the future as per community requirements. This standard is, nevertheless, standalone and complete.

4.5 The operator is allowed to practice before the test. She/he is also allowed to abstain from the test before it is started. Once the test begins, there shall be no verbal communication between the operator and the administrator regarding the performance of a test repetition other than describing the target as seen by the operator and instructions on when to start and notifications of faults and any safety related conditions. The operator shall have the full responsibility to determine whether and when the robot has completed a repetition and notify the administrator accordingly. However, it is the administrator's authority to judge the completeness of the repetition.

NOTE 4—Practice within the test apparatus is allowed to establish the applicability of the robot for the test method. It allows the operator to gain familiarity with the standard apparatus and environmental conditions. It also enables the test administrator to establish the initial apparatus setting for the test.

4.6 The test sponsor has the authority to establish the testing policy, including the robot participation, testing schedules, test site at which this test method is implemented, associated environmental conditions, the apparatus settings, and statistical reliability and confidence levels of the testing results.

5. Significance and Use

5.1 A main purpose of using robots in emergency response operations is to enhance the safety and effectiveness of emergency responders operating in hazardous or inaccessible environments. The testing results of the candidate robot shall describe, in a statistically significant way, how reliably the robot is able to perform the specified types of tasks and thus provide emergency responders sufficiently high levels of confidence to determine the applicability of the robot.

5.2 This test method addresses robot performance requirements expressed by emergency responders and representatives from other interested organizations. Robot performance data captured within this test method are indicative of the robotic system's capabilities. Having available a roster of successfully tested robots with associated performance data to guide procurement and deployment decisions for emergency responders is consistent with the guideline of "Governments at all levels have a responsibility to develop detailed, robust, all-hazards response plans" as stated in National Response Framework.

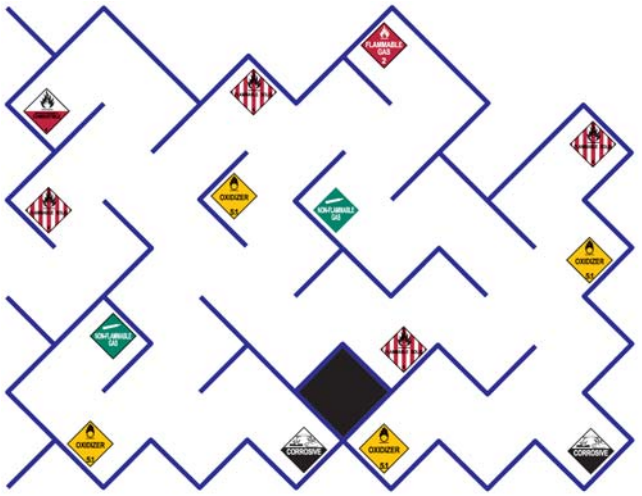

FIG. 2 Random Maze

5.3 This test method is part of a test suite and is intended to provide a capability baseline for the robotic HSI subsystems based on the identified needs of the emergency response community. Adequate performance using this test suite will not ensure successful operation in all emergency response situations due to possible extreme operational difficulties. Rather, this test method is intended to provide a common comparison of technologies against a reasonable simulation of emergency response environments and to provide quantitative performance data to emergency response organizations to aid in choosing appropriate systems. This standard is also intended to encourage development of improved and innovative communications systems for use on emergency response robots.

5.4 The standard apparatus is specified to be easily fabricated to facilitate self-evaluation by robot developers and provide practice tasks for emergency responders that exercise robot actuators, sensors, and operator interfaces. The standard apparatus can also be used to support operator training and to establish operator proficiency.

5.5 Although the test method was developed first for emergency response robots, it may be applicable to other operational domains, such as law enforcement and armed services.

6. Apparatus

6.1 The test apparatus is composed of 1.2 by 1.2 m (4 by 4 ft) floor units. Each of the four sides of such a unit is designed to be, randomly, either covered with 2.4 m (8 ft) high wall or open. The apparatus is fabricated from 1.2 by 2.4 m (4 by 8 ft) OSB or plywood (Fig. 2). The maze is formed by 48 of such a unit. Ramp terrain elements are used on the floor. Fig. 3 provides an illustration. Gray shades indicate the elevations from the ground—the darker, the higher. The white lines indicate the boundaries of the floor units that are without wall, whereas the blue lines indicate the wall. Additional, thin white lines indicate the ridges formed by the ramps.

6.2 Standard HAZMAT labels shall be posted randomly on the wall of the maze to serve as the search targets. The testing sponsor has the authority to determine the number of the labels to be used and the locations to post them, including the heights.

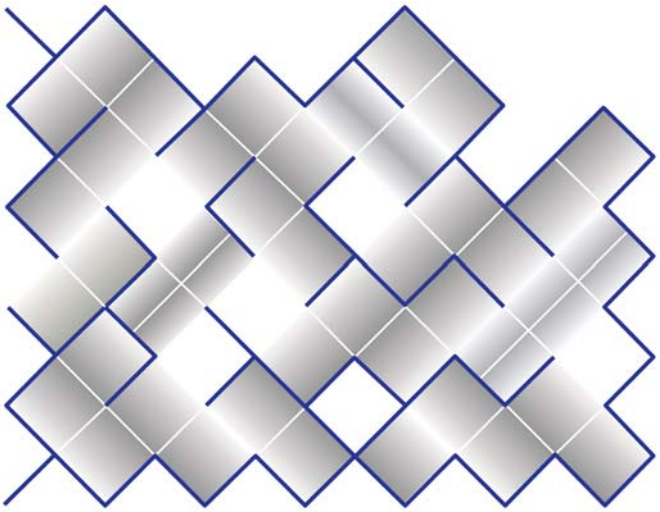

FIG. 3 Maze Floor Filled with Full-Ramp Terrain Elements

6.3 The test sponsor has the authority to determine the starting point for the search, which can be a position inside or at the entrance of the maze.

6.4 Various test conditions such as apparatus surface types and conditions, including moisture and friction levels of the floor, temperature, types of lighting, smoke, humidity, and rain shall be facilitated when the test sponsor requires. For example, for a test run in the dark environment, a light meter shall be used to read 0.1 lux or less. The darkness shall be re-measured when the lighting condition might have changed. The actual readings of these conditions should be recorded on the test form.

NOTE 5—The testing apparatus can be scaled down and implemented in a standard International Standards Organization (ISO) shipping container, in which some of the testing conditions can be furnished. For example, for testing in the dark condition, turn off all the lighting sources inside and entirely cover the entrance with light-blocking drapes. The darkness is specified as 0.1 lux due to the implementation cost concerns for the apparatuses and due to the fact that robotic cameras are less sensitive than human, such that any darkness below 0.1 lux would not make a difference in the cameras' functioning. It is recognized that the environments in real applications may be darker than the specified test condition.

6.5 A stopwatch shall be provided to measure the timing performance.

7. Hazards

7.1 Besides section 1.4 that addresses the human safety and health concerns, users of the standard shall also address the equipment preservation concerns and human-robot coexistence concerns.

NOTE 6—Adverse environmental conditions, such as high or low temperatures, excessive moisture, and rough terrains can be stressful to the humans, can damage the robotic components, or can cause unexpected robotic motions.

8. Calibration and Standardization

8.1 The robot configuration as tested shall be recorded in detail on the test form, including all subsystems and components and their respective features and functionalities. The configuration shall be subjected to all the applicable test methods as determined by the test sponsor. Any variation in the configuration shall cause the resulting robot variant to be retested across all the determined test methods to provide a consistent and comprehensive representation of the capabilities. Practice E2592 shall be used to record the robotic configuration.

8.2 Once a robot begins to be teleoperated to execute a specified task, the task shall be performed for the specified number of repetitions through completion without leaving the apparatus. During the process, any human physical intervention, including adjustment, maintenance, or repair constitutes a fault condition.

8.3 The metrics for this test method are the number of covered target(s), the number of cleared target(s), and the average rates of advance—the averaged number of targets that the robot covered or cleared per unit time. The test sponsor also has the option of dividing the maze wall into certain height ranges to either limit the search space or to partition the search capabilities.

8.4 In addition, the test sponsor has the authority of requiring the traversing routes to be marked on the test forms as per how the operators convey to the administrators. This information will show if there are areas and targets that either are repeatedly visited or are missed, further indicating the efficiency and completeness of the searches.

8.5 The test sponsor has the authority to determine the relative weights of these metrics.

NOTE 7—When she/he determines the search criteria to be finding a target or not to miss any target in the maze, an operator might choose to ignore search efficiency and choose to visit certain areas repeatedly, instead to ensure the search results.

8.6 Although the metric is based on teleoperation, autonomous behaviors are allowed as long as the testing procedure is followed, with the associated effects reflected in the testing scores. See NIST Special Publication 1011-I-2.0 for the definition of autonomy.

8.7 The test sponsor has the authority to specify the lighting condition and other environmental variables under which to test with. All environmental settings shall be noted on the test form.

8.8 A robot's reliability (R) of performing the specified task at a particular apparatus setting and the associated confidence (C) shall be established. The required R and C values dictate the required numbers of successful coverage's or clearance's and the allowed numbers of failures during the test. With a given set of the R and C values, more successes will be needed when more failures are allowed. A test sponsor has the authority to specify the R and C values for her/his testing purposes. The factors to be considered in determining the values are mission requirements, consistency with the operating environments, ease of performing the required number of repetitions, and testing costs such as time and personnel. The values can be calculated by referring to general statistical analysis methods. Section 11 provides sets of illustrations.

9. Procedure

9.1 For data traceability and organization purposes, the administrator shall obtain and record the testing conditions and

administrative information first. A set of specified fault conditions shall be followed during the test.

NOTE 8—For example, different robot models could help partially explain the differences in the test results. Different trial numbers could partially tell how much effort an operator has taken to accomplish the results.

9.2 *Testing Conditions and Administrative Information:*

9.2.1 *Date*—Provide testing date; some test methods, when explicitly specified, can allow the tasks or repetitions to be distributed into multiple days; the time-of-the-day information may also be included.

9.2.2 *Facility*—Provide the name of laboratory or field where the test is to be conducted.

9.2.3 *Location*—Provide the names of campus, city, and state in which the facility is located.

9.2.4 *Event/Sponsor*—This field shall be recorded as general when a robot is tested for its performance record purposes independent of any particular event.

9.2.5 *Robot Make*—Provide the name of the manufacturer of the robot.

9.2.6 *Robot Model*—Provide the specific name and model number, including any extension or remark to fully identify the particular configuration of the robot as tested.

9.2.7 *Robot Configuration*—Identifier of the particular configuration of the robot as tested.

9.2.8 *Operator/Org*—Provide the identification of the person who will teleoperate the robot for testing and the name of the organization with which the operator is associated.

9.2.9 *Environment*—Provide the subject conditions under which the test will be conducted; the test sponsor can specify the lighting condition, temperature and humidity.

9.2.10 *Robot Communications*—State whether the robot is using radio, tether, or a combination to run the test.

9.2.11 *Trial Number*—Provide the numerical sequence of the test being recorded.

9.2.12 Provide the administrator's name, organization, and the contact information.

9.2.13 Additional information such as the naming convention for the performance-capturing video files is provided at the bottom of the form.

9.2.14 *Administrator*—Name, organization, and the contact information.

9.2.15 See the top and the bottom of the test form in Figs. 4 and 5 for an illustration.

9.3 *Testing Procedure:*

9.3.1 The operator either abstains or proceeds with the test. The abstention shall not be granted after this point.

9.3.2 The administrator sets and verifies the apparatus setting and announces the number of repetitions to be performed. The target number and location information shall not be disclosed to the operator.

9.3.3 The administrator sets and verifies the test environmental conditions.

9.3.4 The operator places the robot at the starting position.

9.3.5 The administrator instructs the operator to begin the search task, starts the timer when the operator begins, and records the total elapsed time.

9.3.6 The operator controls the robot to traverse the maze while she/he watches the video feeds on the OCU to look for the targets. The administrator records the locations and features of the targets as how the operator conveys. If the robot fails to advance anywhere in the maze, this constitutes a fault condition. The administrator shall pause the overall test time and allow the operator to interact with the robot, reset the robot back to the start point, and resume the test when the administrator signals. The administrator shall note, on the test form, the indication of the fault condition and the time at which the pause occurred and shall provide a comprehensive maintenance and repair report if any such actions occur. The operator shall convey the actions that she/he has taken to the administrator to facilitate this reporting.

9.3.7 The test is completed upon the robot's returning to the starting point. The operator can elect to notify the administrator to end the test before the completion.

9.4 *Fault Conditions:*

9.4.1 Robot fails to complete the search task once started.

9.4.2 There is human communication with the operator regarding the status of the robot or the task.

9.4.3 Human intervenes with the robot, such as adjustment, maintenance, or repair.

10. Report

10.1 A test form is required for this test method. The form shall include the following features and allow for recording both the testing information and the test results:

10.1.1 Metrics and corresponding measuring scales and ranges;

10.1.2 Any additional testing features such as those that can reflect performance proficiency;

10.1.3 Important notes to be recorded during the test, including particular fault conditions that occurred, the reason for abstaining, any observations by the administrator that could augment the recorded results in either positive or negative ways, or any comments that the operator requests to be put on the form;

10.1.4 Testing conditions and administrative information.

10.2 Fig. 4 provides an illustration of a blank test form for this test method. The test form shall be filled out to full extent as applicable. Section 10.3 specifies how to fill out a test form.

10.3 The following designations shall be used to indicate the testing results:

10.3.1 *Not Tested*—The scoring section of the test form shall be left blank. The Notes section shall record the reason(s) for not testing, such as:

10.3.1.1 The test method was not available during testing time, including the apparatus could not be properly set up or there were uncontrollable environmental conditions or scheduling difficulties.

10.3.1.2 The robot is not within the scope of the test method, for example, a ground robot test method is not applicable to an aerial robot.

10.3.2 *Abstained*—Typically indicated with a red stamp to the effect to be printed on the lower corner on the right-hand

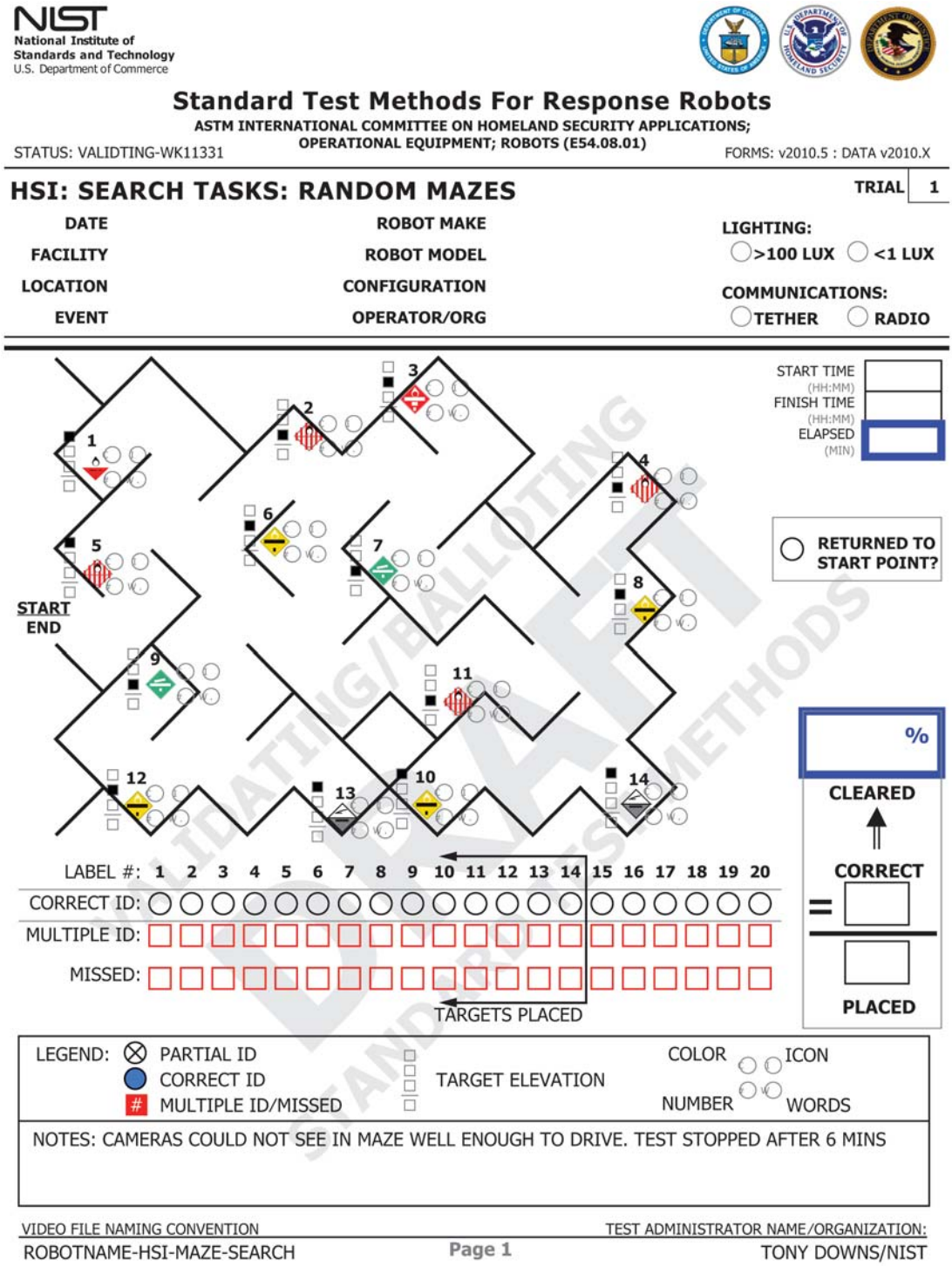

FIG. 4 Test Form Implementation

side. The testing conditions and administrative information shall be filled out. The scoring section of the test form shall be left blank.

10.3.3 *Success*—Typically indicated with a blue checked box or recorded with the achieved numerical value. All the successful testing results shall be explicitly marked.

10.3.4 *Tested but Failed*—Typically indicated with an unchecked mark or an "X" in the red color. When a robot has failed a particular apparatus setting, all the more difficult apparatus settings shall be considered insurmountable.

10.3.5 *Test Result Accepted but Administrative Pause is Necessary*—An orange checked box is typically used to indicate the effect, along with a timestamp and note describing the reason for the administrative intervention. This designation is used when the test apparatus is in need of repair or maintenance neither the fault of the operator nor the robot under test.

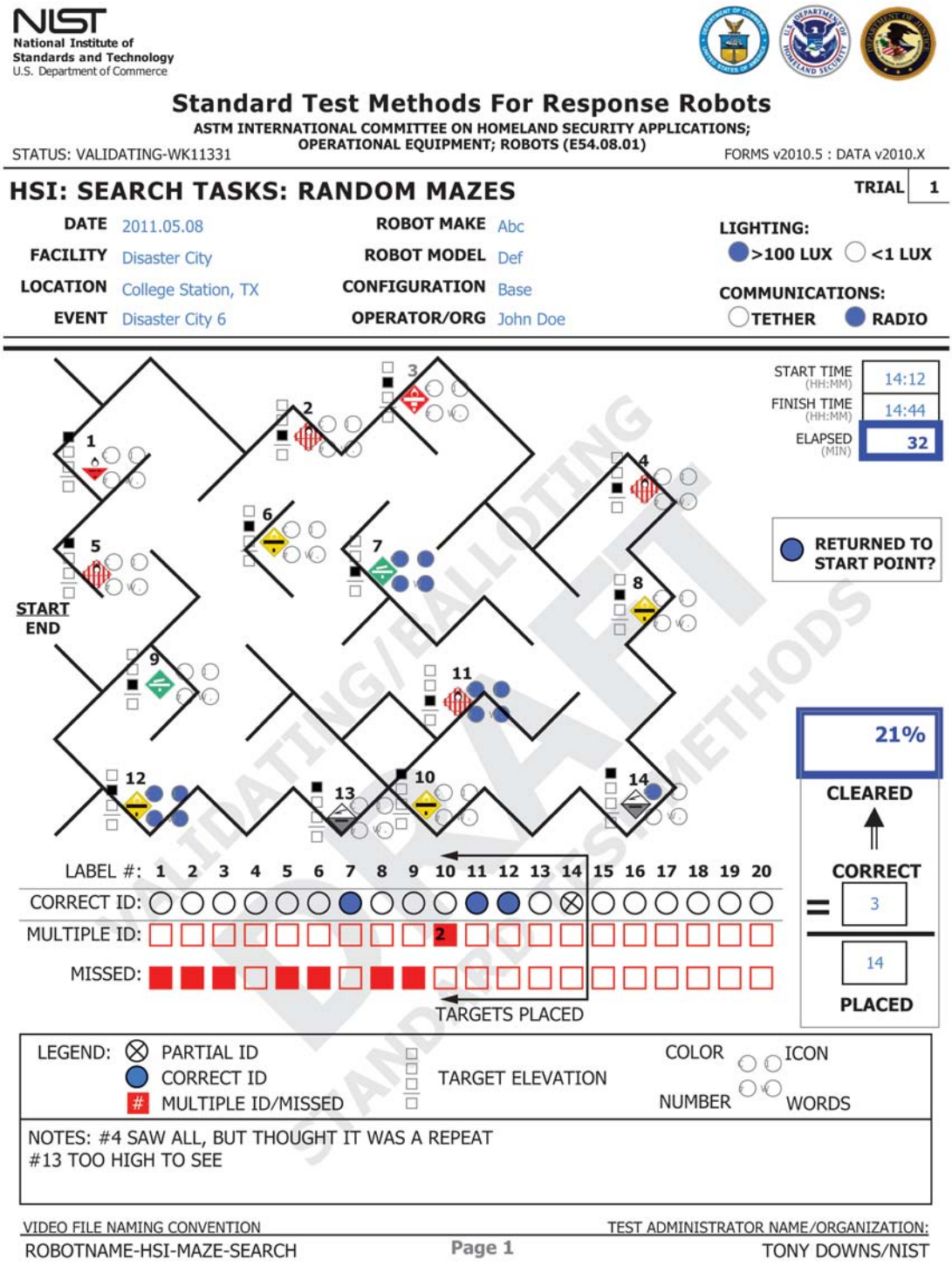

FIG. 5 Testing Result Illustration

This designation is also used with the occurrences of minor errors considered inconsequential to the overall outcome of the test so that the test can continue through completion.

10.4 Fig. 5 illustrates how such a test form can be filled out.

NOTE 9—The implementation of the test form is not standardized. As such, the resulting forms can be different while conforming to this specification.

11. Precision and Bias

11.1 *Precision:*

11.1.1 This test method, as a part of the HSI test suite, seeks to quantitatively measure the capabilities of robots intended to operate in human-scale structures and environments involving possibly multiple-day long operations, kilometer-range long distances, and a myriad of obstacles and terrain types.

TABLE 1 Testing Results for Random Maze Search for Robot A

robot #A					
height range	trgt #	# of fea. cleared	C/R for covering a target (via clearing at least 1 feature)	R/C for clearing a target	C/R for clearing a target once covered it
A:	1	0	5 % / 60 %	2 % / 60 %	6 % / 60 %
1.6 m – 2.4 m	5	0	0 % / 80 %	0 % / 80 %	2 % / 80 %
(63 in – 96 in)	10	0		(0 out of 5)	(0 out of 1)
	13	1			
	14	0			
B:	3	0	22 % / 60 %	0 % / 60 %	5 % / 60 %
0.8 m – 1.6 m	6	2	5 % / 80 %	0 % / 80 %	1 % / 80 %
(31½ in – 63 in)	8	0	(2 out of 4 trgts)		(0 out of 2)
	12	2			
C:	2	4	59 % / 60 %	59 % / 60 %	80 % / 60 %
0 m – 0.8 m	4	4	26 % / 80 %	26 % / 80 %	55 % / 80 %
(0 in – 31½ in)	7	4	(4 out of 5 trgts)	(4 out of 5 trgts)	(4 out of 4 trgts)
	9	4			
	11	0			

Key: trgt – Target; fea. – Feature; C – confidence level; R – reliability level

Therefore, coarsely testing a greater variety of robot capabilities more often is preferable to establish the overall capabilities of a given robot configuration. For this reason, the apparatus settings related to this test method sparsely placed target in the maze—14 HAZMAT labels in all the possible units, each with four possible sides and three possible height ranges for the labels to be posted, in this particular test. While test apparatuses could be developed to test the HSI capabilities to smaller increments, those are considered too fine for the operational conditions associated with human-scale structures and environments and would increase the overall testing time per robot. As such, finer incremental testing is considered outside the scope of this testing approach.

11.1.2 Repeatability testing is required to validate the test method by showing that, for a variety of robot lengths, weights and locomotion types, the test method produced repeatable results for a range of apparatus settings—multiple targets types and the height ranges. The apparatus and metric scale increments produced clear delineations between successful and unsuccessful attempts and demonstrated that the test method is suitable for evaluating the obstacle-negotiation capability.

11.1.3 Tables 1-3 provide a set of repeatability testing results for a representative three out of all the participating robots. The apparatus is set with 14 targets at locations unknown to the operators. Further, the targets are placed on the wall and are divided into the following three height categories to increase the complexity of the search: A. between 1.6 m and 2.4 m (63–96 in), B. between 0.8 m and 1.6 m (31½ –63 in), and C. 0–0.8 m (0–31½ in). The lighting condition is ambient light. Each height range contains four or five targets. The following precision and bias statistical features are computed based on the testing results:

Confidence level for operator covering a target with 60 % reliability
Confidence level for operator covering a target with 80 % reliability
Confidence level for operator clearing a target with 60 % reliability
Confidence level for operator clearing a target with 80 % reliability
Confidence level for operator clearing a target once the target is covered, with 60 % reliability
Confidence level for operator clearing a target once the target is covered, with 80 % reliability

11.1.3.1 For example, as illustrated in Table 1, the operator for Robot A covered target #13 but was not able to clear it by using the robotic system. The operator covered two targets in height range B but was not able to clear either. The operator has at least 59 % level confidence and 60 % level reliability in clearing a target in the height range C.

11.1.3.2 For example, as illustrated in Table 2, the operator for Robot B covered and cleared three targets in height range A and one target in height range B. The operator has at least 26 % level confidence and 80 % level reliability in clearing a target in the height range C.

11.1.3.3 For example, as illustrated in Table 3, the operator for Robot C covered and cleared one target each in height ranges A and B. The operator has at least 32 % level confidence and 60 % level reliability in clearing a target in the height range C.

11.1.4 As specified in Section 1, it is recommended that users of this test method consider the scope of the test as it applies to their own projects. Performance in this test method alone shall not be considered as the collective indication of the performance of the robot's HSI subsystem nor of the entire robotic system. Testing across the entire suite of applicable test methods is essential to determine the capabilities of the robot in general.

11.2 *Bias:*

11.2.1 One variable that was found typically to introduce a bias was the operator's familiarity with the test method. The operator's performance was normally lowest when she/he did not have prior practices. The performance typically improved to a stable level once the operator practiced sufficiently.

11.2.1.1 There are additional human factors that can introduce biases, including the skill level, fatigue level, and level of concentration of the operator. An operator who obtained proper training and possessed abundant field experiences could perform at a higher level, particularly when all the robotic capability was fully exercised.

11.2.2 Onboard sensing capability can affect the task performance. The range(s) and the field of view of the camera(s) can affect how the operator is able to see the test apparatus and control the robot accordingly.

11.2.3 Yet another variable that was found to introduce a possible bias was the lighting conditions. In some situations,

TABLE 2 Testing Results for Random Maze Search for Robot B

robot #B					
height range	trgt #	# of fea. cleared	C/R for covering a target (via clearing at least 1 feature)	R/C for clearing a target	C/R for clearing a target once covered it
A:	1	4	32 % / 60 %	32 % / 60 %	70 % / 60 %
1.6 m – 2.4 m	5	4	8 % / 80 %	8 % / 80 %	45 % / 80 %
(63 in – 96 in)	10	0	(3 out of 5 trgts)	(3 out of 5 trgts)	(3 out of 3 trgts)
	13	0			
	14	4			
B:	3	0	8 % / 60 %	8 % / 60 %	33 % / 60 %
0.8 m – 1.6 m	6	4	1 % / 80 %	1 % / 80 %	18 % / 80 %
(31½ in – 63 in)	8	0	(1 out of 4 trgts)	(1 out of 4 trgts)	(1 out of 1trgts)
	12	0			
C:	2	4	59 % / 60 %	59 % / 60 %	80 % / 60 %
0 m – 0.8 m	4	4	26 % / 80 %	26 % / 80 %	55 % / 80 %
(0 in – 31½ in)	7	4	(4 out of 5 trgts)	(4 out of 5 trgts)	(4 out of 4 trgts)
	9	0			
	11	4			

Key: trgt – Target; fea. – Feature; C – confidence level; R – reliability level

TABLE 3 Testing Results for Random Maze Search for Robot C

robot #C					
height range	trgt #	# of fea. cleared	C/R for covering a target (via clearing at least 1 feature)	R/C for clearing a target	C/R for clearing a target once covered it
A:	1	0	5 % / 60 %	5 % / 60 %	33 % / 60 %
1.6 m – 2.4 m	5	0	0 % / 80 %	0 % / 80 %	18 % / 80 %
(63 in – 96 in)	10	4		(1 out of 5 trgts)	(1 out of 1 trgts)
	13	0			
	14	0			
B:	3	0	8 % / 60 %	8 % / 60 %	33 % / 60 %
0.8 m – 1.6 m	6	4	1 % / 80 %	1 % / 80 %	18 % / 80 %
(31½ in – 63 in)	8	0	(1 out of 4 trgts)	(1 out of 4 trgts)	(1 out of 1 trgts)
	12	0			
C:	2	4	32 % / 60 %	32 % / 60 %	70 % / 60 %
0 m – 0.8 m	4	4	8 % / 80 %	8 % / 80 %	45 % / 80 %
(0 in – 31½ in)	7	4	(3 out of 5 trgts)	(3 out of 5 trgts)	(3 out of 3 trgts)
	9	0			
	11	0			

Key: trgt – Target; fea. – Feature; C – confidence level; R – reliability level

certain degrees of variations in the lighting conditions occur, which can result in different search results. However, a test sponsor can elect to test only under a specific and consistent lighting condition in order to prevent from this potential bias.

NOTE 10—A reason causing the possible variations is this standard's approach of specifying apparatuses that can be implemented with nominal costs and efforts. As such, sophisticated control for the environmental conditions might not be affordable. Users of the test results should reference the lighting conditions.

12. Measurement Uncertainty

12.1 Proper use of this test method to measure the random maze search capability will result in an uncertainty of one half of the elapsed time unit, or 30 s as well as an uncertainty of the recognition of one target. Section 11.1.1 specifies that finer resolutions are insignificant for this test method.

13. Keywords

13.1 abstain; emergency responder; emergency response; flat-floor terrain element; HSI; human-scale; OCU; operator control unit; operator station; oriented strand board; OSB; repetition; robot; test suite; urban search and rescue; US&R

ASTM International takes no position respecting the validity of any patent rights asserted in connection with any item mentioned in this standard. Users of this standard are expressly advised that determination of the validity of any such patent rights, and the risk of infringement of such rights, are entirely their own responsibility.

This standard is subject to revision at any time by the responsible technical committee and must be reviewed every five years and if not revised, either reapproved or withdrawn. Your comments are invited either for revision of this standard or for additional standards and should be addressed to ASTM International Headquarters. Your comments will receive careful consideration at a meeting of the responsible technical committee, which you may attend. If you feel that your comments have not received a fair hearing you should make your views known to the ASTM Committee on Standards, at the address shown below.

This standard is copyrighted by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, United States. Individual reprints (single or multiple copies) of this standard may be obtained by contacting ASTM at the above address or at 610-832-9585 (phone), 610-832-9555 (fax), or service@astm.org (e-mail); or through the ASTM website (www.astm.org). Permission rights to photocopy the standard may also be secured from the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, Tel: (978) 646-2600; http://www.copyright.com/