NISTIR 7443

Maze Hypothesis Development in Assessing Robot Performance During Teleoperation

Salvatore P. Schipani
Elena R. Messina

September 28, 2007

Table of Contents

List of Figures

List of Tables

ABSTRACT

National Institute of Standards and Technology (NIST) personnel had the opportunity to assess 14 prospective Urban Search and Rescue (USAR) robots, for the purposes of developing performance standards which currently do not exist. During this exercise, a maze configuration – hypothesized as potentially valid test methodology – was assessed. Among the findings, significant differences in completion and decision making times found between platforms enabled classifying these based on performance characteristics. Also revealed was that errors in navigation and encounters with walls correlated with times taken in making decisions… the longer it took to make a decision, the greater the chance this decision was incorrect. Results validated the hypothesis of a maze as beneficial in eliciting data necessary for human controlled robot performance assessment.

1. INTRODUCTION

Test performance standards for application-specific Urban Search and Rescue (USAR) robots providing valid replicable assessment measures do not exist, thus little or no guidance may be offered to local, state, or federal agencies regarding their utilization or procurement. In 2004, the Department of Homeland Security (DHS) Science and Technology (S&T) Directorate initiated an effort with the National Institute of Standards and Technology (NIST) to formulate comprehensive criteria related to the development, performance testing, and certification of available and anticipated robotic technologies, specifically directed toward application in USAR scenarios. To encourage collaboration between USAR responders and system developers, and in hopes of generating standards consensus among those interested, a third response robot evaluation exercise was conducted by NIST at the Montgomery County Fire Rescue Training Academy in Rockville, Maryland, particularly targeting the needs of DHS/Federal Emergency Management Agency (FEMA) USAR professionals. Operational standards deemed of concern included mobility, sensing, navigation, planning, integration into operational caches, and consideration of the human factor.

Individual characteristics of current production robots utilized for USAR vary. In light of recent national security concerns, this reality brings to the forefront a necessity for categorizing the operational capabilities of tools (for example, robots) and methods used by emergency response professionals in conducting duties. Any attempt at the organization of such information must address the requirements of emergency response professionals, and offer recommendations for system attribute improvement as discovered. In August of 2006, NIST personnel had the opportunity to assess 14 robots with potential for application during USAR situations based on visual sensors, mobility, logistic cache packaging, radio communications, and human factors in operations. This document reports on one proposed measure of performance, a subset of the decision making process referred to as operator time to acquire situation awareness, when attempting to teleoperate a robot within a maze, a scenario hypothesized as a valid test methodology given observed apparatus methods of control and assumed tasks.

1.1 BACKGROUND

1.1.1 Maze Rationale

Mazes derived directly from their predecessors, ancient labyrinth designs. This symbol and its family of derivatives may be traced back over 3500 years, however its origins remain a mystery. As opposed to a maze, labyrinths have no false pathways or dead ends, but rather consist of one single meandering way leading from entrance to center. Conversely, mazes may possess many paths, enticing or impairing anyone attempting to maneuver through. These have become accepted exercises in direction finding, providing paths to follow, some correct and others erroneous. As such, they are considered highly respectable tests of navigational skills, and attempted by many.

Correlations between maze performance and traditional psychometric measures of spatial ability have affirmed the relationship [1,2], especially as vestibular information from the inner ear as well as kinesthetic feedback from muscles has been shown to provide important cues regarding direction of heading and distance information [3,4]. The rationale becomes particularly acceptable subsequent to reviews of factor analytical studies for large spatial batteries yielding multiple spatial dimensions [5,6,7]. Optic flow also provides motion and movement cues necessary to navigate through environments, offering a visual analyses of motion which we have come to anticipate and rely on. Unfortunately, during teleoperation, such visual cues become the *only* aid presented [8,9], rendering tasks such as remote control especially difficult. Given that these cues are often disturbed during teleoperation due to issues in transmission, it should be expected that maze navigation become increasingly difficult.

1.1.2 Acquiring Situation Awareness

Though several definitions of Situation Awareness (SA) are posed in literature [10,11,12], SA is normally defined in terms of goals with particular decision tasks directed to a specified effort [13,14]. One definition offered, encompassing the essence of what most researchers care to relate, is *"The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future"* [15]. Endsley and Garland [14] further define levels of SA as: *Level 1*, the perception of cues: *Level 2*, an extension of cue perception, including the integration of multiple pieces of information plus the determination of their relevance to goals. Here, meaning must be considered as subjective interpretation (awareness) and objective significance (situation) [16, page 3], so that at this level one is able to derive operational relevance and significance from prior data, and; *Level 3*, the ability to forecast future events. SA is normally depicted as an operator's internal state model within an environment [17,14], causing designers to consistently question how well particular systems support one's ability to acquire necessary information. This design concern is exaggerated in dynamic situations and under operational constraints, thus observing the acquisition and eventual degree of SA has become a frequently used measure of performance.

Time has been shown a critical affecting factor in acquiring both *Levels 2* (comprehension) and *3* (future event projection) SA [18,19,20]. This is particularly the case in teleremote operations, as operator SA must be derived from a combination of the environment and integrated system's displays, and then interpreted by the operator at afforded instances and in short intervals [14]. Here, sufficient information must be provided through a remote interface so as to compensate for cues once perceived directly [21], an unfortunate scenario commonly found deficient. The collection of whatever information presented is assumed a subset of that derived from the environment and internal system parameters, however only a portion may be displayed via existing (visual) interfaces. With the majority of teleoperated systems currently deployed, operators are given minimal control of which information may be collected other than that presented via the visual channel, and are often restricted in transmitting commands to request further knowledge arrived at in such ways as by the autonomous selection of directions of traverse or specifying areas of sensor coverage [22]. Such deficiencies in data acquisition not only lengthen the time required for information collection, but also inhibit assimilation.

In goal driven processing such as that which takes place during teleoperation, an operator actively seeks information required for attainment of the goal, during which the mental model is claimed as existing underlying knowledge therefore the basis for SA [23]. Smith and Handcock (1995) support this view of SA as behavior directed toward goal achievement, describing it as the *"…up-to-the minute comprehension of task relevant information"*. Referred to as cognition-in-action, Lave [24] claims *"SA fashions behavior in anticipation of the task-specific consequences of alternative actions"*. Over time, a pattern-recognition thus action-selection sequence becomes routine, developing to a level of response automaticity [25]. Such automaticity can positively affect SA by reducing demands on limited attentional resources, but only if proper information is retrieved, comprehended, and adequately assimilated. When one's goal is to eventually emplace a system (robot) at a specified location, an internal model of previously traversed terrain with appropriate continued or corrected model for subsequent route direction becomes essential. This has been shown difficult when using existing teleremote visual displays due to inadequate cueing for guidance, and lack of available space for displaying previous information, thus SA is compromised.

2. METHOD

2.1 PARTICIPANTS

Personnel operating robots during this exercise were engineering professionals representing their respective product. Each had extensive experience not only in robot operation, but also in development. Additionally, each vendor-operator was made aware that the performance of their product would be compared to competitors during the exercise, thus it behooved them to offer their best operator for the assessment. Personal observations substantiated the fact that each participant could be considered proficient in robot manipulation, thus the level of expertise was deemed a fixed factor. In all, 14 participants were involved, one each from all robot vendors appearing for the test.

2.2 MATERIALS

2.2.1 Test Course

In this particular maze (see Figure 1), there exists one possible solution with only a single main branch leading to correct termination, having an approximate solution length of 2,117.29 centimeters (833.58 inches) which consists of 21 wall segments equating to 21.17 meters (69.47 feet). Traveling forward, the maze possess three left turns, three right turns, three straight-aways, two left curves, no right curves, two irregular curves, two ramps, four junctions, no crossroads, loops, or roundabout passages, and two dead-ended isolation points (designated points 1 and 2 in the Figure 1 diagram). Additionally, two route enticements were constructed at which light was visible hinting at clear passage however actually blocked, with only short possible deviation lengths within the two provided false passages of 115.57 and 346.71 centimeters (45.5 and 136.5 inches).

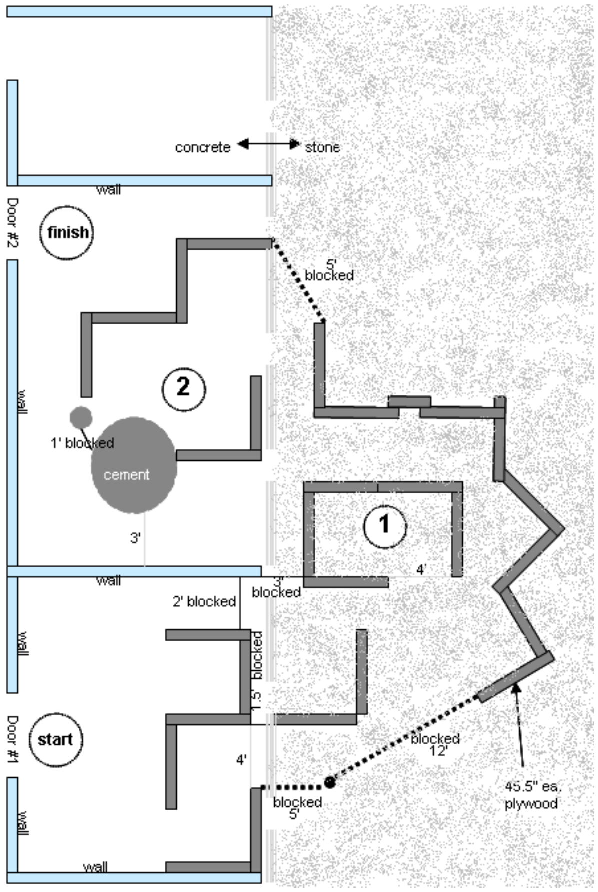

Figure 1: Original Maze Configuration (not to scale)

2.2.2 Robots

Following are descriptions of the 14 participating robots, accompanied by individual dimensions (taken from the NIST draft publication "Response Robots – DHS/NIST Sponsored Evaluation Exercises" Pocket Guide, Version 2006.1). To provide for anonymity, each has been designated a number in place of name.

- Robot #1: Width 57.15 centimeters (22.5 inches), length 86.36 centimeters (34 inches), height 38.1 centimeters (15 inches), weight 52.16-63.5 kilograms (115-140 pounds), turning diameter 0 centimeters (0 inches) (skid-steer, tracks), maximum speed 8.369 kilometers per hour (5.2 miles per hour), non tethered (tether option), remote teleoperation control, sensor include black and white camera (optional biological, chemical, and temperature sensors), five degrees-of-freedom 132.08 centimeters (52 inches) horizontal reach end effector (*i.e.*, manipulator);

- Robot #2: Width 30.988 centimeters (12.2 inches), length 42.164 centimeters (16.6 inches), height 15.24 centimeters (6 inches), weight 6.35 kilograms (14 pounds), turning diameter 0 centimeters (0 inches) (skid-steer), maximum speed 2.286 meters per second (7.5 feet per second), non tethered, remote teleoperation control, sensor include black and white camera (with options for thermal, acoustic, infra-red, and visual wide-angle sensing), no end effector;

- Robot #3: Width 25.4 centimeters (10 inches), length 35.56 centimeters (14 inches), height 16.51 centimeters (6.5 inches), weight 6.35 kilograms (14 pounds), turning diameter 50.8 centimeters (20 inches), maximum speed, 1.829 meters per second (6 feet per second), non tethered, remote teleoperation control, sensors include color and infrared cameras, no end effector;

- Robot #4: (no data available);

- Robot #5: Width 55.88 centimeters (22 inches), length 68.58 centimeters (27 inches), height 63.5 centimeters (25 inches), weight 56.7 kilograms (125 pounds), tracked skid-steer turns on center, maximum speed 10.46 kilometers per hour (6.5 miles per hour), non tethered, eyes on and remote teleoperation control with way-point following and drive intent, sensors color video camera and laser range scanner, no end effector;

- Robot #6: Width 57.15 centimeters (22.5 inches), length 86.36 centimeters (34 inches), height 63.5 centimeters (25 inches), weight 52.16-63.5 kilograms (115-140 pounds), turning diameter 0 centimeters (0 inches) (skid-steer, tracks), maximum speed 8.369 kilometers per hour (5.2 miles per hour), non tethered (tether option), remote teleoperation control, sensor include black and white camera (optional biological, chemical, and temperature sensors), five degrees-of-freedom 132.08 centimeter (52 inch) horizontal reach end effector;

- Robot #7: (no data available);

- Robot #8: Width 34.29 centimeters (13.5 inches), length 52.07 centimeters (20.5 inches), height 30.48 centimeters (12 inches), weight 11.34 kilograms (25 pounds), turning diameter 0 centimeters (0 inches) (skid-steer, tracks), maximum speed 6.437 kilometers per hour (4 miles per hour), non tethered, no tether, eyes on and remote teleoperation control, sensor black and white camera, no end effector (i.e., manipulator);

- Robot #9: Width 53.34 centimeters (21 inches), length 76.2-86.36 centimeters (30-34 inches), height 30.48 centimeters (12 inches), weight 27.67 kilograms (61 pounds), turning diameter 0 centimeters (0 inches) (skid-steer, tracks), maximum speed 3.219 kilometers per hour (2 miles per hour), non tethered, fiber optic cable tether (for data, video, and audio), remote teleoperation control, sensors include black and white camera (optional biological, chemical, and radiological sensors), five degrees-of-freedom 111.76 centimeter (44 inch) end effector;

- Robot #10: Width 40.64 centimeters (16 inches), length 63.5 centimeters (25 inches), height 19.304 centimeters (7.6 inches), weight 11.34 kilograms (25 pounds), turns in place, maximum speed 1.341 meters per second (4.4 feet per second), non tethered, remote teleoperation and telemetry control, sensor black and white camera, end effector (i.e., manipulator) six degrees-of-freedom with 106.68 centimeter (42 inch) reach;

- Robot #11: Width 40.64 centimeters (16 inches), length 68.58 centimeters (27 inches), height 19.05 centimeters (7.5 inches), weight 21.77 kilograms (48 pounds), turning diameter 0 centimeters (0 inches), maximum speed 8.047 kilometers per hour (5 miles per hour), non tethered, remote teleoperation control, sensor black and white camera on short non-extending boom, no end effector;

- Robot #12: Width 40.64 centimeters (16 inches), length 68.58 centimeters (27 inches), height 19.05 centimeters (7.5 inches), weight 21.77 kilograms (48 pounds), turning diameter 0 centimeters (0 inches), maximum speed 8.047 kilometers per hour (5 miles per hour), non tethered, remote teleoperation control, sensor black and white camera on three-rod extending boom, no end effector;

- Robot #13: Width 50.8 centimeters (20 inches), length 55.88 centimeters (22 inches), height 45.72 centimeters (18 inches), weight 6.804 kilograms (15 pounds), turning diameter 0 centimeters (0 inches), maximum speed 5.633 kilometers per hour (3.5 miles per hour), non tethered, remote teleoperation control, sensor black and white camera, no end effector;

- Robot #14: Width 27.432 centimeters (10.8 inches), length 42.672 centimeters (16.8 inches), height 13.97 centimeters (5.5 inches), weight 6.35-9.072 kilograms (14-20 pounds), turning diameter 0 centimeters (0 inches) (skid-steer, tracks), maximum speed 0.4572 meters per minute (1.5 foot per minute), 30.48 meter (100 foot) polyurethane multi-cord tether, remote teleoperation and eyes-on control, sensor black and white tilt camera, no end effector.

3. PROCEDURE

Participants were directed – upon the experimenter command "begin" – to teleoperate assigned robotic platforms traversing pathways through the unfamiliar maze, and do so within the shortest time possible. They were further instructed to operate carefully enough to limit or avoid encounters with path walls. Their informed consent to participate and to allow a video record made of their system was agreed upon prior to test initiation, at which time operator sightedness was screened. Participants were permitted to ask questions concerning test methods and purpose prior to testing, or at any time during the test. They were instructed that they were to complete four iterations, two in forward and two in reverse, until reaching their goals which were open doorways located at the beginning and end of the maze.

3.1 DATA COLLECTION

Time data collection was recorded in seconds and performed manually utilizing hand-held stop watches (one recording total maze traverse time, the second monitoring time spent in decision points), and on digital video in order that post-test evaluations of performance could be made. Video records were taken via hand held roving camera, with camera person consistently positioned behind the robot thus completely out of robot camera view to ensure that no visual cues were offered to operators.

4. EXPERIMENTAL DESIGN

The experiment was treated as a 2 *x* 2 *x* 1 factorial, where two levels of traverse exist (forward and reverse), with two instances of dead ended isolation points, and this applied between the performance of 14 robotic platforms given one level of operator proficiency.

4.1 DEPENDENT MEASURES

Dependent measures were averaged "maze completion times" traveling forward and reverse, averaged "decision making times" recorded at points specified within the maze, "errors" in direction of traverse when exiting aforementioned decision points, and observed "encounters" made with maze walls.

Times were recorded for total maze completion in each direction (separate forward and reverse recordings), and during instances at which robots entered into and lingered in designated dead-ended isolation points. For total completion time, recording began as test director instructed participants to "begin" each trial, and ended once the robot reached the step-sill of exit doors located at either end of the maze. Each participant completed two forward and two reverse iterations.

For instances in which participants entered a dead-ended isolation area (*i.e.*, decision eliciting 'traps'), total time spent within was recorded. Time data collection for this began when the most forward portion of a robot crossed a horizontal imaginary line at the entrance of the isolation area, and ended as the most forward portion again crossed this line exiting. This data was treated as the time necessary for participants to gain situation adequate awareness, sufficient for participants to realize that they had entered a dead end in the maze and to reach a decision on how to properly exit.

As participants exited dead-ended decision points, their direction of traverse was recorded for correctness. The accurate direction could be determined by experimenter observation as being the most obvious direction of course traverse within which one might successfully complete the maze. Finally, robot encounters with walls (*e.g.*, "hits") were recorded as each participant teleoperated through pathways.

5. RESULTS

Following (see Table 1) find descriptive statistics for averaged *Maze Completion Time*, *Decision Making Time*, wall *Hits* (encounters), and *Errors* in direction traversed.

	Mean	Std. Dev.	Std. Error	Count	Minimum	Maximum	# Missing
av. Comp Time	1.944	.896	.240	14	.380	3.320	0
av. Decision Time	18.429	8.145	2.177	14	6.000	32.500	0
Hits	2.786	3.017	.806	14	0.000	11.000	0
Errors	2.286	1.383	.370	14	0.000	5.000	0

Table 1: Descriptive Statistics
(Times presented in seconds, Hits & Errors as unit segments)

Figure 2 presents maze completion times, showing robots 2, 8, 9, 10 and 12 displaying lowest times to complete the maze (averaging 1.14 minutes, or 68.4 seconds), and robots 5, 6, 7 the highest (averaging 3.23 minutes, or 193.8 seconds).

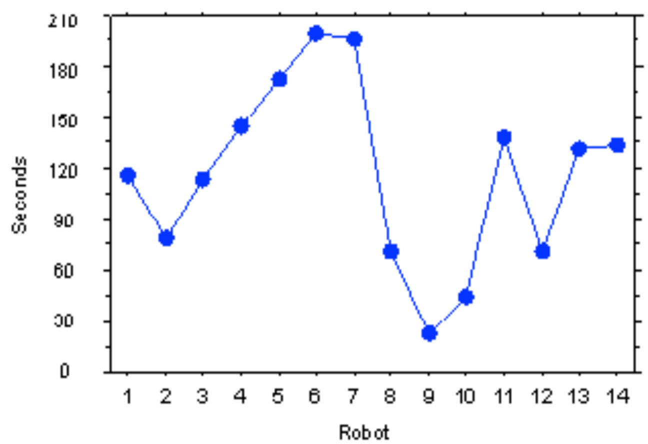

Figure 2: Maze Completion Times

Robots 1, 8 and 9 displayed lowest decision making times (*i.e.,* Situation Awareness gaining time) averaging 6.93 seconds, and robots 6, 11, and 13 the highest averaging 31 seconds (see Figure 3).

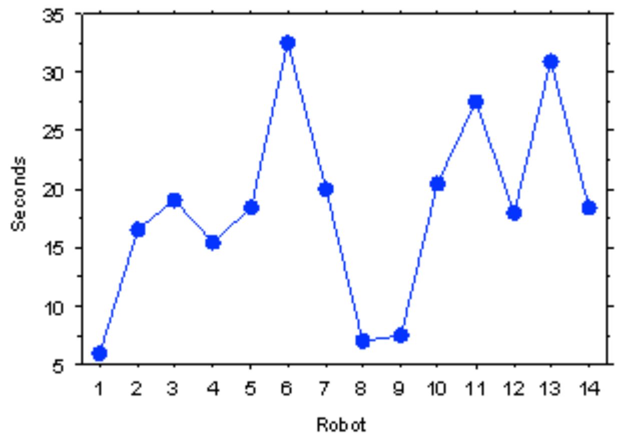

Figure 3: Decision Making Times

No statistically significant difference found among robots for average maze completion times (*p* = 0.68), the most frequently attained ranging from 2.14 minutes to 2.48 minutes (128.4 seconds -148.8 seconds). It may be assumed that – being a first attempt – the current maze configuration did not provide sufficient distance to evoke performance anticipated. Future maze investigations employing increased areas of traverse should resolve this issue. However, three categories may be delineated from the data when observing performance groupings which ranged from slightly greater or less than 1.0 minute, on average 2.2 minutes, and slightly greater or less than 3.0 minutes (60 seconds, 132 seconds, and 180 seconds respectively) (see Figure 4). There was a significant difference found between *forward* and *reverse* times to complete the maze (*p* = 0.003). Times in reverse were shorter, obviously an indication that operators were becoming familiar with the test course.

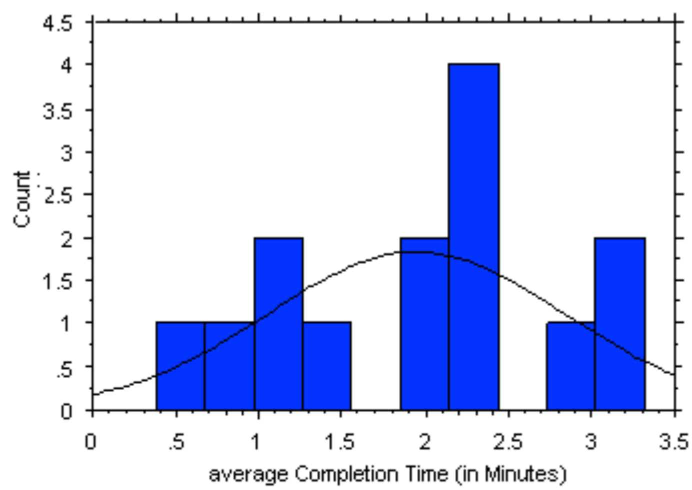

Figure 4. Histogram of Distributed Completion Times

There was a significant difference found among robots concerning averaged decision making times ($p = 0.001$), as individual attributes of particular platforms apparently aided or hindered performance during the challenge. There was not a significant difference found between times to decide at isolation area 1 versus 2 ($p = 0.891$), revealing the two similar in nature. The most frequently attained decision making times ranged from 16.6 seconds to 19.25 seconds. Here again, robots could be grouped per three categories of performance of from slightly greater or less than 7.5 seconds, averaged 18.5 seconds, or slightly greater or less than 30 seconds (see Figure 5).

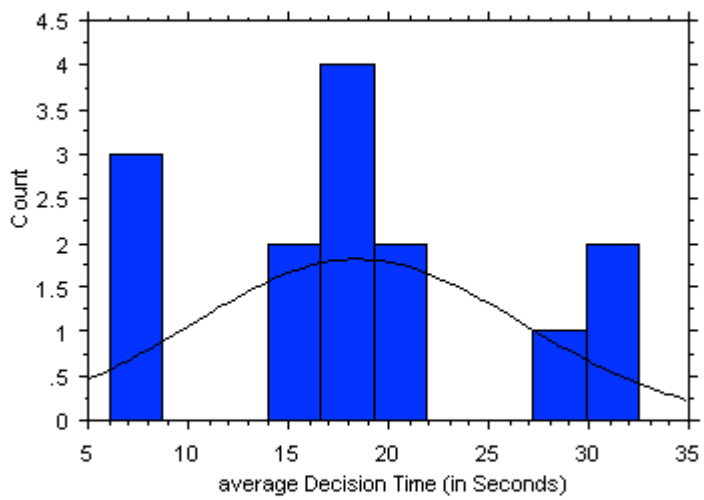

Figure 5. Histogram of Distributed Decision Times

A significant difference was found among robots concerning hits (wall encounters) ($p = 0.001$). In reviewing video recordings, it would appear as if particular robots acted out-of-control due to inferior or transmission lagged control response, no or poor methods of halting forward movement, or poor camera views provided the operator.

There was a significant difference found among robots concerning errors ($p = 0.048$). Errors were also found correlated with increased times spent in making decisions (r = 0.67). This would appear to support the notion that the longer it took to make a decision as to which direction to move next, the more this decision (the direction of traverse selected) was found incorrect. No significant correlations were observed between averaged completion times and decision making times, revealing these entities distinct (r = 0.543). However, averaged wall hits data correlated highly with errors made in correct direction of traverse (r = 0.864), suggesting confusion in the selection of subsequent travel direction due to post-collision trauma.

For comparative purposes, individual performance is displayed in Table 2.

Robot	Completion Time	Decision Time	Errors	Hits
1	average	best	best	best
2	average	average	average	best
3	average	average	average	best
4	average	average	average	best
5	poor	average	average	best
6	poor	poor	poor	poor
7	poor	average	poor	poor
8	best	best	best	best
9	best	best	best	best
10	best	average	average	best
11	average	poor	poor	poor
12	best	average	average	best
13	average	average	poor	best
14	average	average	average	best

Table 2: Performance as a function of Dependent Measures

6. DISCUSSION

At present, performance standards for Urban Search and Rescue (USAR) designated robots are nonexistent, thus little guidance may be offered to local, state, or federal agencies regarding their purchase or use. A precursor to successful search and rescue operations if employing a robot is the ability to teleoperate the system satisfactorily, attaining directional cures from the remote visual display as possible. When one is driving, vestibular information and kinesthetic feedback provide additional cues regarding direction. However, during teleoperation, the *only* cues available are those presented visually, yet sufficient information must be attained via a remote interface in order to compensate thus discern most advantageous pathways. Intensifying this effort, situation awareness in such circumstances must be attempted while on-the-move, which becomes defined in terms of goal achievement with time the critical factor affecting acquisition. This document reports on one scenario hypothesized as valid methodology for assessing performance of such platforms, a maze test configuration employed as a navigation exercise.

Data collected included time to complete the maze, and also that necessary for gaining situation awareness when entrapped in either of two predestinated dead-ended isolation points. Data also included recordings of maze wall encounters, and errors made in direction of traverse. Digital video recordings were taken to enable *post hoc* analyses. Participants were directed to teleoperate their assigned platforms through the maze in the shortest time possible, while avoiding encounters with walls. Fourteen robots, potential candidates for deployment in USAR scenarios, were involved. Participants operating the robots were engineering professionals representing their respective product, each possessing extensive experience both in operation and platform development. Results revealed significant differences in time to gain situation awareness ($p = 0.001$), encounters with walls ($p = 0.001$), and errors made in direction of traverse ($p = 0.048$). Also uncovered was that increased times spent in making decisions correlated with erroneous subsequently selected directions of traverse ($r = 0.67$), supporting the notion that the longer it took to make a navigational decision the more this decision could be found incorrect. Finally, encounters with walls correlated highly with errors made in direction of traverse ($r = 0.864$), revealing confusion as a result of post-collision trauma.

Given results of the current exercise, utilization of a maze test approach for evaluating robot teleoperation appears rational, as the scenario elicited data sufficient to examine performance as intended. Forthcoming endeavors are expected to include increased maze distances and complexity, to ensure that appropriate pragmatic assessments may be made.

Anticipations are to submit the maze hypothesis to tests of validity and reliability in the near future. Generally accepted validity determinations involve criterion-oriented procedures such as predictive and concurrent, or are else-wise considered either content or construct [26]. For the test method in question, a predictive approach to validation appears most logical, as criterion-oriented validity "*involves the acceptance of a set of operations as an adequate definition of whatever is to be measured*." [27]. This will be attempted per performance criterion found necessary via repeated investigation, as well as by exploiting guidance offered from emergency response professionals. Reliability assessments should establish whether this type examination measures consistently. Concurrently, appropriate levels of maze complexity will be evaluated, and mathematical formulas aiding in maze construction developed for use by those not capable of testing at a NIST designated arena. Subsequently, results will be submitted through appropriate committee of the American Society for Testing and Materials (ASTM) to attain consensus as a national standard, as NIST personnel explore supplementary measurement methods deemed essential.

7. REFERENCES

[22] Adams, M.J., Tenney, Y.J., & Pew, R.W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors*, *37*(1).

[26] Anastasi, A. (1950). The concept of validity in the interpretation of test scores. *Educational Psychological Measurement*, *10*, 67-78.

[27] Bechtoldt, H. P. (1951). Selection. In S.S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, 1237-1267.

[5] Borich, G.D. & Bauman, P.M. (1972). Convergent and discriminant validation for the French and Guilford-Zimmerman spatial orientation and spatial visualization factors. *Educational and Psychological Measurement*, *32*, 1029-1033.

[3] Chapuis, N. & Scardigli, P. (1993). Shortcut ability in hamsters (Mesocricetus auratus): The role of environmental and kinesthetic information. *Animal Learning and Behavior*, *21*, 255-265.

[19] Cohen, M.S., Freeman, J.T., & Wolf, S. (1996). Metacognition in time-stressed decision making: Recognizing, critiquing, and correcting. *Human Factors*, *38*(2), 206-219.

[12] Dominguez, C. (1994). Can SA be defined? In, M.Vidulich, C.Domingues, E.Vogel, & G.McMillan (Eds.), *Situation awareness: Papers and annotated bibliography* (AL/CF-TR-1994-0085), 5-15. Wright-Patterson AFB, OH.: Armstrong Laboratory.

[6] Ekstrom, R.B., French, J., Harman, H.H., & Dermen, D. (1976). *Kit of Factor Referenced Cognitive Tests*. Princeton, MJ: Educational Testing Services.

[15] Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. In, *Proceedings of the Human Factors Society 32nd Annual Meeting* (97-101). Santa Monica, CA.: Human Factors Society.

[14] Endsley, M.R. & Garland, D.J. (Eds.), (2000). *Situation Awareness Analysis and Measurement*. Mahwah, N.J.: Lawrence Erlbaum Associates.

[18] Endsley, M.R., Farley, T.C., Jones, W.M., Midkiff, A.H., & Hansman, R.J. (1998). *Situation awareness information requirements for commercial airline pilots* (ICAT-98-1). Cambridge, MA.: Massachusetts Institute of Technology International Center for Air Transportation.

[20] Endsley, M.R. & Robertson, M. (1996). Team situation awareness in aviation maintenance. In, *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society* (1077-1081). Santa Monica. CA.: Human Factors and Ergonomics Society.

[4] Etinne, A.S., Teroni, E., Maurer, R., Portenier, V, & Saucy, F. (1985). Short distance homing in a small mammal: The role of exteroceptive cues and path integration. *Experientia*, *41*, 122-125.

[16] Flach, J.M. (1995). Situation awareness: Proceed with caution. *Human Factors*, *37*(1), 149-157.

[10] Fracker, M.L. (1988). A theory of situation assessment: Implications for measuring situation awareness. In, *Proceedings of the Human Factors Society 32nd Annual Meeting* (102-106). Santa Monica, CA.: Human Factors Society.

[1] Gaulin, S.J.C & Fitzgerald, R.W. 1986). Sex differences in spatial ability: An evolutionary hypothesis and test. *American Naturalist*, *127*, 74-88.

[17] Gibson, J., Orasanu, J., Villeda, E., & Nygren, T.E. (1977). Loss of situation awareness: Causes and consequences. In, R.S. Jensen & R.L.A. (Eds.), *Proceedings of the 8th International Symposium on Aviation Psychology* (pp. 1417-1421). Columbus, OH.: The Ohio State University.

[25] Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492-527.

[24] Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life*. Cambridge: Cambridge University Press.

[21] Marshak, W.P., Kuperman, G., Ramsey, E.G., & Wilson, D. (1987). Situational awareness in map displays. In, *Proceedings of the Human Factors Society 31st Annual Meeting* (533-535). Santa Monica, CA.: Human Factors Society.

[7] Michael, W.B., Zimmerman, W.S., & Guilford, J.P. (1951). An investigation of the nature of the spatial relations and visualization factors in two high school samples. *Educational and Psychological Measurement*, *11*, 561-577.

[8] Mittlestaedt, H. (1983). The role of multimodal convergence in homing by path integration. In *Multimodal Convergences in Sensory Systems*, E.Horn (Ed.). New York: Gustav Fischer Verlag, 197-212.

[23] Mogford, R.H. (1977). Mental models and situation awareness in air traffic control. *International Journal of Aviation Psychology*, *7*(4), 331-342.

[11] Sarter, N.B. & Woods, D.D. (1991). Situation awareness: A critical but ill-defined phenomenon. The *International Journal of Aviation Psychology*, *1*(1), 45-57.

[2] Silverman, I. & Eals, M. (1992). Sex differences in spatial

abilities: Evolutionary theory and data. In the *Adapted Mind*, J..H. Barkow, L.Cosmides, and J.Tooby (Eds.). New York: Oxford, 533-549.

[13] Smith, K. & Hancock, P.A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors*, *37*(1), 137-148.

[9] Teroni, E, Portenier, V, & Etienne, A.S. (1987). Spatial orientation of the golden hamster in conditions of conflicting location-based and route-based information. *Behavioral Ecology and Sociobiology*, *20*, 389-397.