Measuring and Comparing Response Robot Capabilities and Remote Operator Proficiency

Version 2020B

Test Director:
Adam Jacoff
Intelligent Systems Division
National Institute of Standards and Technology
U.S. Department of Commerce

Sponsor:
Science and Technology Directorate
U.S. Department of Homeland Security

Internet
RobotTestMethods.nist.gov

Email
RobotTestMethods@nist.gov

Acknowledgements

Acknowledgements

This work was sponsored by **Philip Mattson** and **Kai-Dee Chu** from the Department of Homeland Security, Science and Technology Directorate, through an interagency agreement with the National Institute of Standards and Technology (NIST).

The NIST Team includes:
Adam Jacoff, Ann Virts, Raymond Sheh, Kamel Saidi, Kenny Kimble.

Dozens more people have contributed to the development and validation of these test methods. They include FEMA urban search and rescue task force teams, firefighters, law enforcement, collaborating test facilities, other civilian and military organizations, and commercial manufacturers. There are far too many to mention, but some of the ongoing (non-commercial) collaborators are listed below, roughly in order of their involvement:

Disclaimer

Commercial equipment shown in this document are for illustrative purposes only. This does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products identified are necessarily the best available for the purpose.

Measurement Units

The International System of Units (a.k.a. SI Units) and U.S. Customary Units (a.k.a. Imperial Units) are used throughout this document. Approximate equivalents in each system of units enable use of readily available materials in different countries. This avoids excessive purchasing and fabrication costs. The differences between the stated unit dimensions are insignificant for comparison of test method results, so each set of units are considered standard for the purposes of these test methods.

Collaborators

Tom Haus, Los Angeles Fire Dept. & CA-TF1, CA

Parry Boogard, Valley Regional Fire Authority & WA-TF1, WA

Clint Arnett, TEEX/Disaster City & TX-TF1, TX

George Hough, Fire Dept. of New York City & NY-TF1, NY

Jim Ingledue, Virginia Beach Fire Dept. & VA-TF2, VA

Mark Hundley, Virginia Beach Fire Dept. & VA-TF2, WA

Michael O'Shea, FAA UAS Integration Office (formerly U.S. DOJ)

Martin Hutchings, Sacramento Sheriff & IAB, CA

John Delaney, Arlington County Fire, Dept., & IAB, VA

Mike Marino, Prince George's County Fire Dept. & IAB, MD

Coitt Kessler, Austin Fire Dept., TX

Chris Sadler, York County Fire Dept., VA

Andy Moore, Southwest Research Institute, San Antonio, TX

Al Frazier, Grand Forks County Sheriff's Dept., ND

Ben Miller, CDPS COE for Aerial Technology Fire Fighting, CO

Mark Blanks, Virginia Tech University, VA

Daniele Nardi, Sapienza Universita di Roma, Italy

Max Delo, ESF-13, U.S. Marshals Service, DOJ

Bryan Gillespy, ESF-13, U.S. Marshals Service, DOJ

Gabriele Ferri, NATO CMRE, Italy

Howie Stockhowe, Virginia Beach Fire Dept, Virginia Beach, VA

Tony Galladora, Montgomery County Police, MD

Satoshi Tadokoro, Tohoku University, Sendai, Japan

Tetsuya Kimura, Nagoaka Univ. of Technology, Nagoaka, Japan

Bob Gann, CDPS COE for Aerial Technology Fire Fighting, CO

Andy Olesen, Canadian Explosives Technicians Assoc., Canada

Tom Prentice, Reveille Peak Ranch, Burnet, TX

Michael Leo, Fire Department of New York City, NY

Luke Bergan, New South Whales Police Dept., Sydney, Australia

Katie Thielmeyer, Woodlawn Fire Dept. OH

Oliver Huke, RACE Test Facility, UKAEA, Oxfordshire, United Kingdom

Project Overview

Safety | Capabilities | Proficiency

Objectives:

Develop the measurements and standards infrastructure necessary to quantitatively evaluate robotic capabilities and remote operator proficiency.

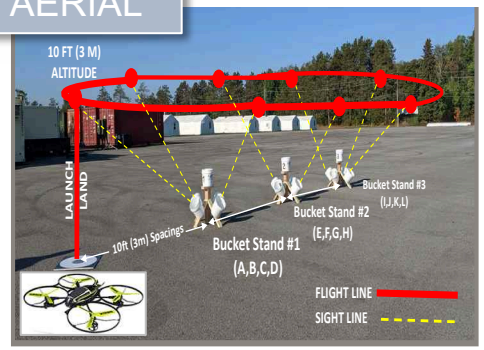
Outcomes:

Test methods, performance metrics, and data collection tools to facilitate integration of emerging technologies for hazardous and essential missions.

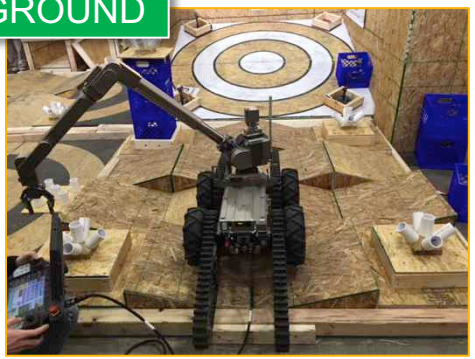
Impacts:

- Objective test methods help researchers and manufacturers push the state of the science by measuring progress and highlighting breakthroughs.
- Resulting quantitative performance data helps compare systems, specify purchases, and train with measures of remote operator/pilot proficiency.

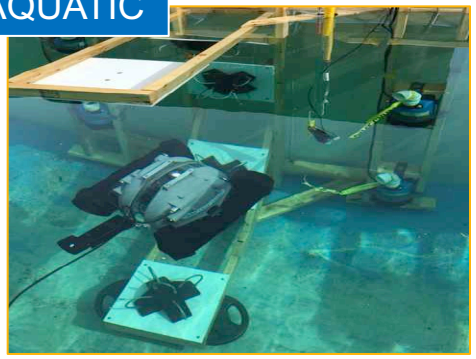
AERIAL



GROUND



AQUATIC



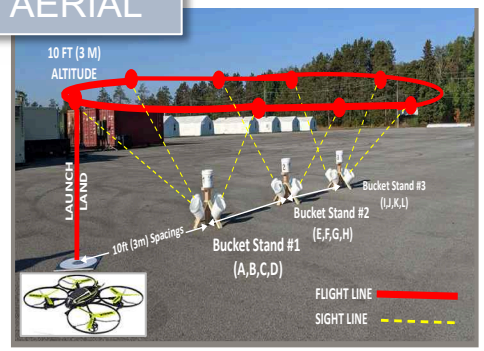
Project Approach

Safety | Capabilities | Proficiency

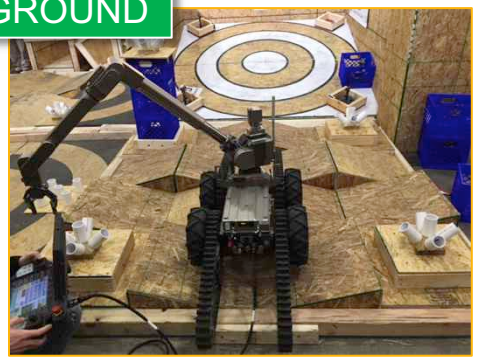
REPEAT

- ***Develop*** reproducible test methods that are cheap and easy to conduct.
- ***Measure*** combinations of existing capabilities and emerging technologies.
- ***Inspire*** innovation using tests to communicate operational needs and gaps.
- ***Guide*** purchasing and deployment decisions with objective data.
- ***Focus*** training with repeatable tasks to measure and compare proficiency.
- ***Identify*** readiness issues with equipment and/or training through comparisons with local, regional, or national averages.

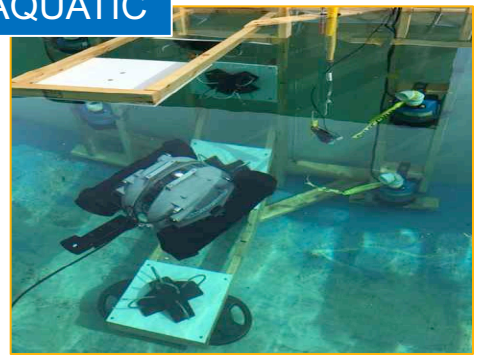
AERIAL



GROUND



AQUATIC



Same Tests Help Different Users

Safety | Capabilities | Proficiency

Robot Developers

- Understand missions through tangible test apparatuses
- Practice and refine robot designs, make trade-off decisions
- Highlight “Best-In-Class” capabilities

Responders, Soldiers, and Other Users

- Compare robots with objective data, not marketing
- Specify purchases based on existing combinations of capabilities
- Align expectations with deployment decisions

Program Managers

- Describe objectives with a collection of tangible tasks
- Challenge conventional approaches and inspire innovation
- Measure baseline capabilities and document progress

Comprehensive Suites of Standard Test Methods

Safety | Capabilities | Proficiency

Mobility
Dexterity
Endurance
Sensors
Radio Comms
Durability
Logistics
Safety
Autonomy
Proficiency


Standard Test Lanes

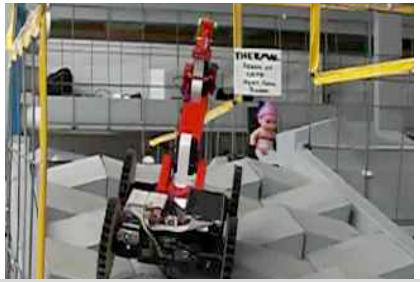

Scorable Scenarios


Repeatable Tasks


Competitions


120cm (48in) Scale

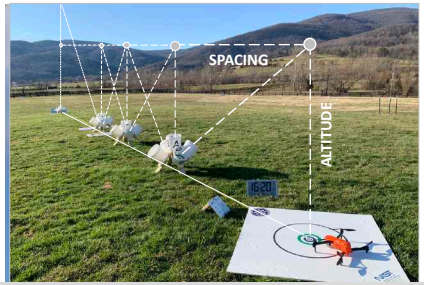

60cm (24in) Scale


30cm (12in) Scale


Confined Access


Aerial (netted)


Aerial (hangar)

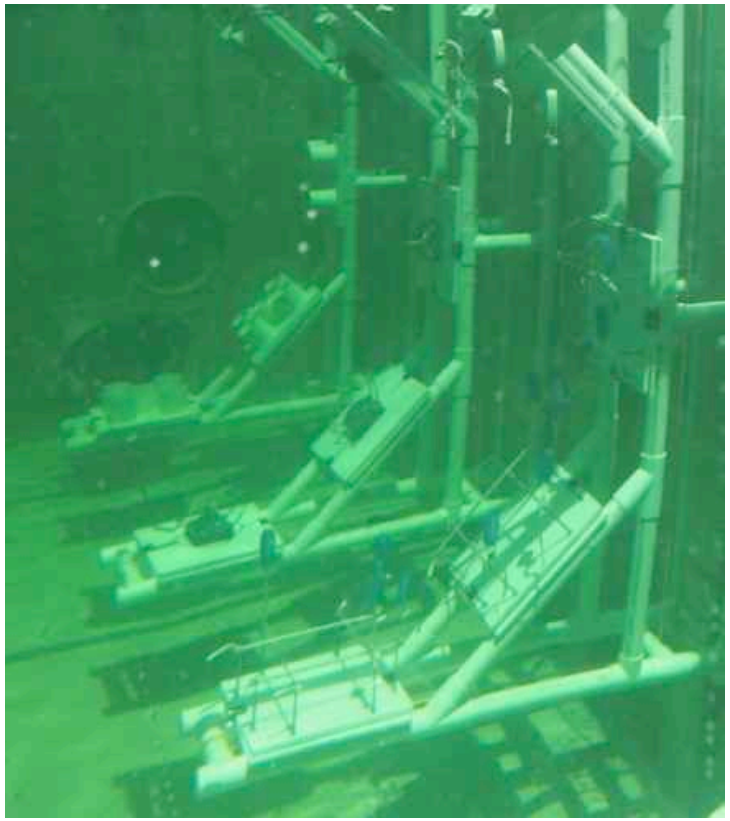


Aerial (outdoor)


Underwater (tank)

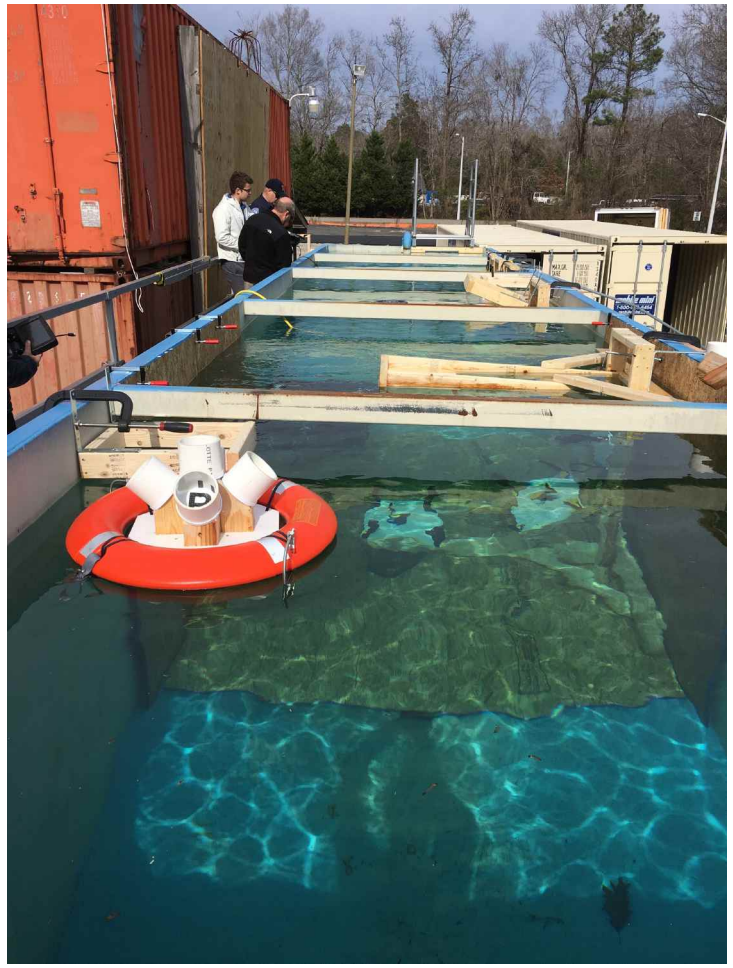
Underwater ROV Test Methods

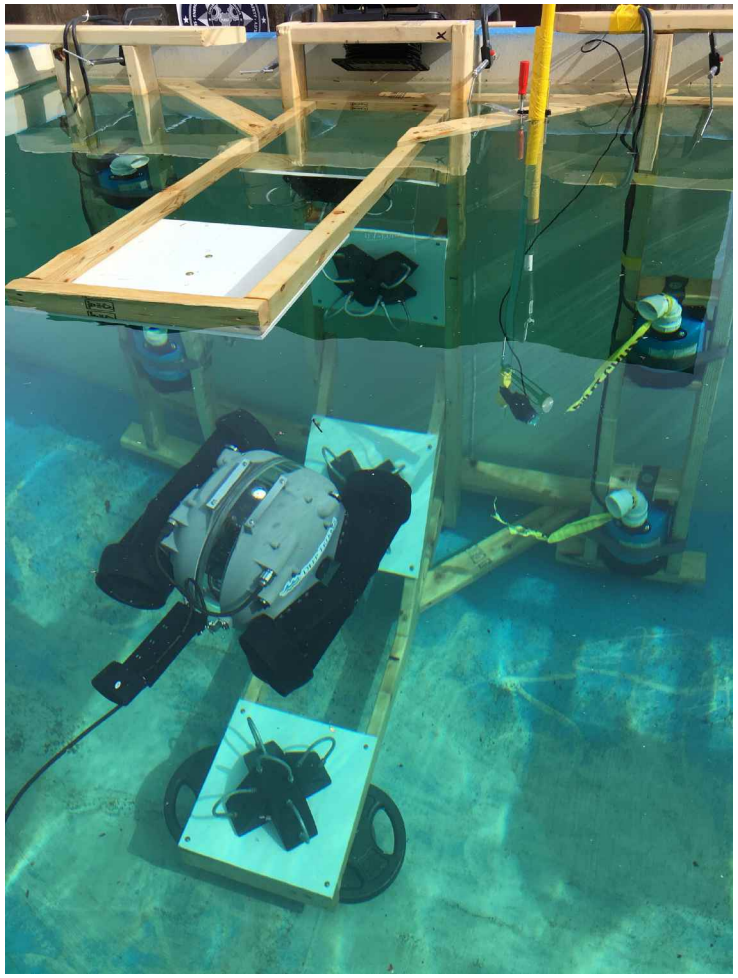
Safety | Capabilities | Proficiency



Underwater ROV Test Methods

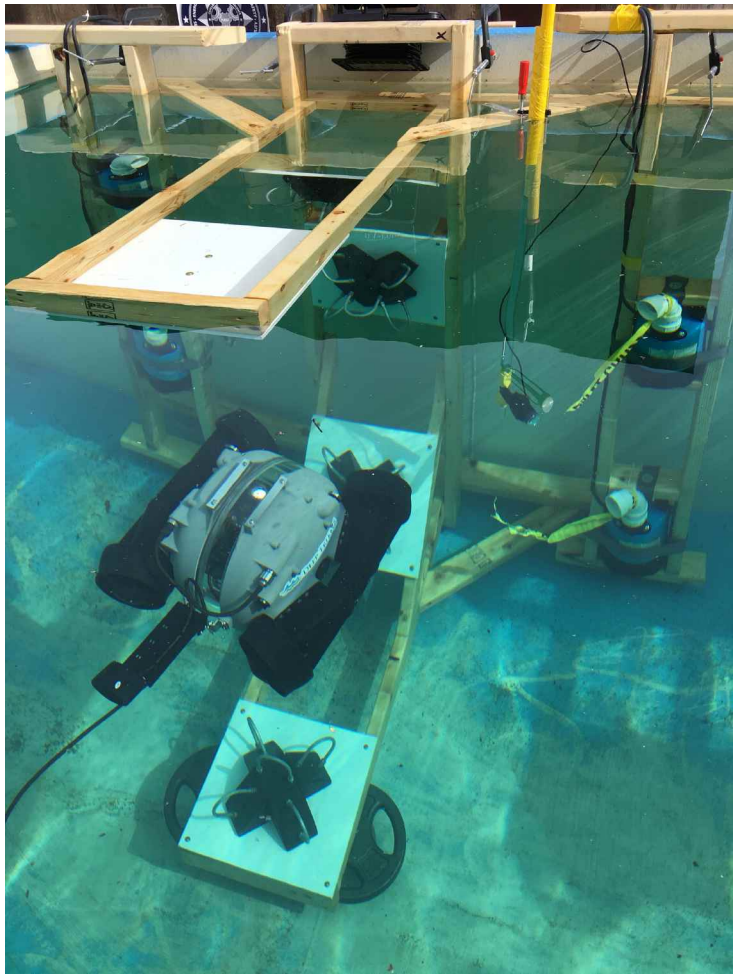
Safety | Capabilities | Proficiency





Underwater ROV Test Methods

Safety | Capabilities | Proficiency



Test Methods for sUAS

Safety | Capabilities | Proficiency

Small Unmanned Aircraft Systems (sUAS)











Initial focus is VTOL, but some tests apply to forward flying aircraft when scaled up to the appropriate orbit radius.

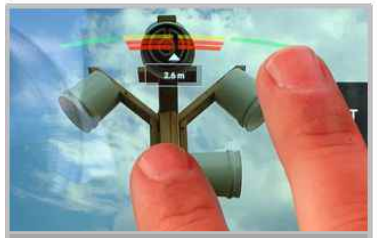
Repeatable Maneuvering and Payload Functionality Tests

Safety | Capabilities | Proficiency

MEASURE & COMPARE


SMALL SYSTEMS


LARGE SYSTEMS

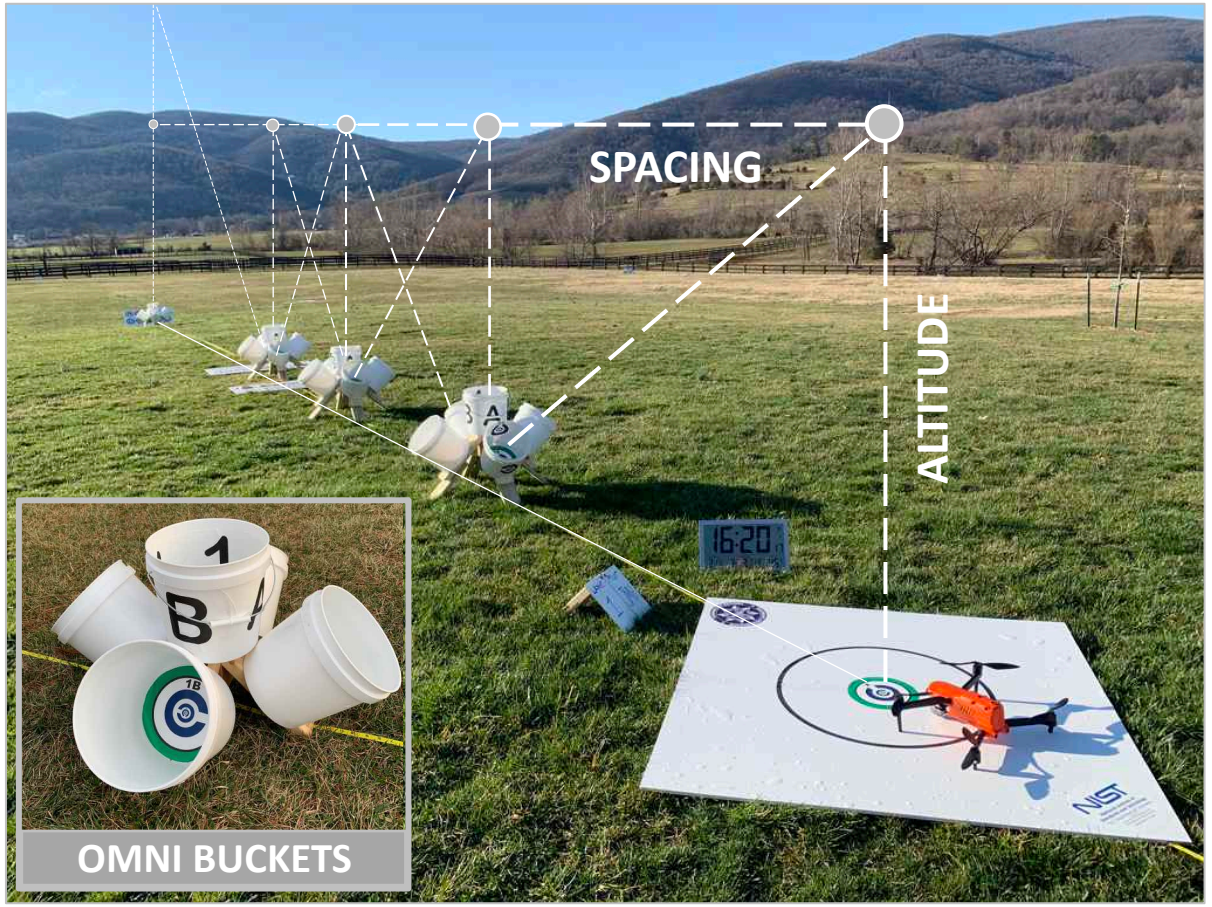

INTERFACES


PROCEDURES

SENSORS


MANEUVERING

SCALABLE TEST LANES (ALTITUDE = SPACING)



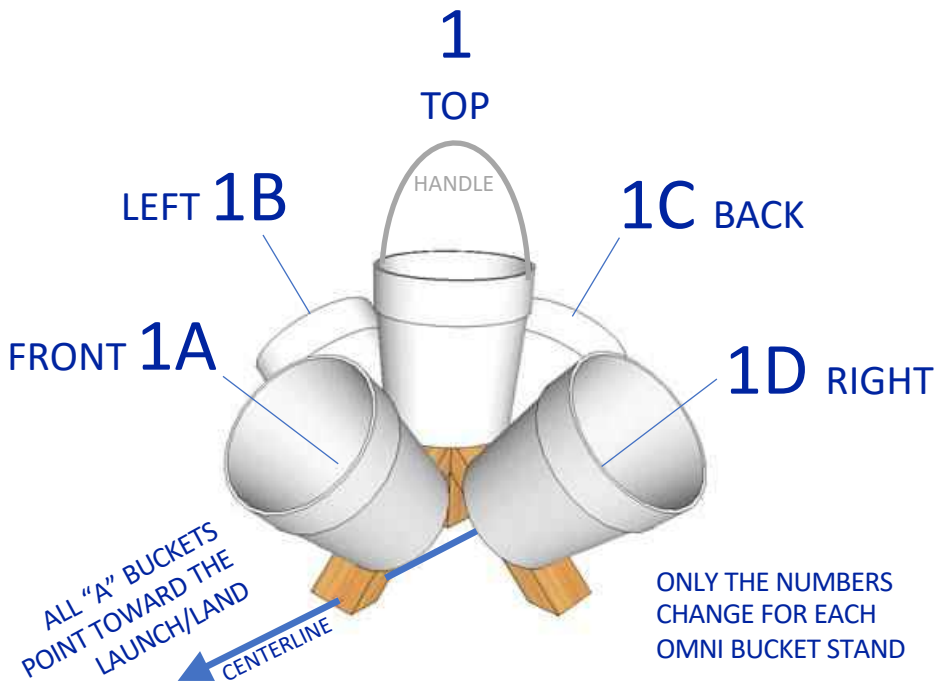
OMNI BUCKETS

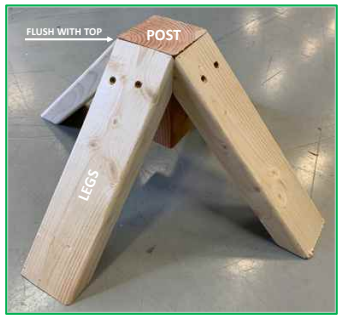
Omni Bucket Stands

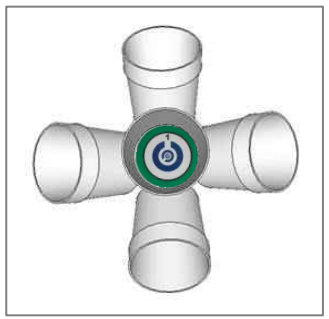
Open Test Lane

WHITE BUCKETS & GREEN RINGS IN STANDARD TEST LANES

SHOWN NOT ALIGNED





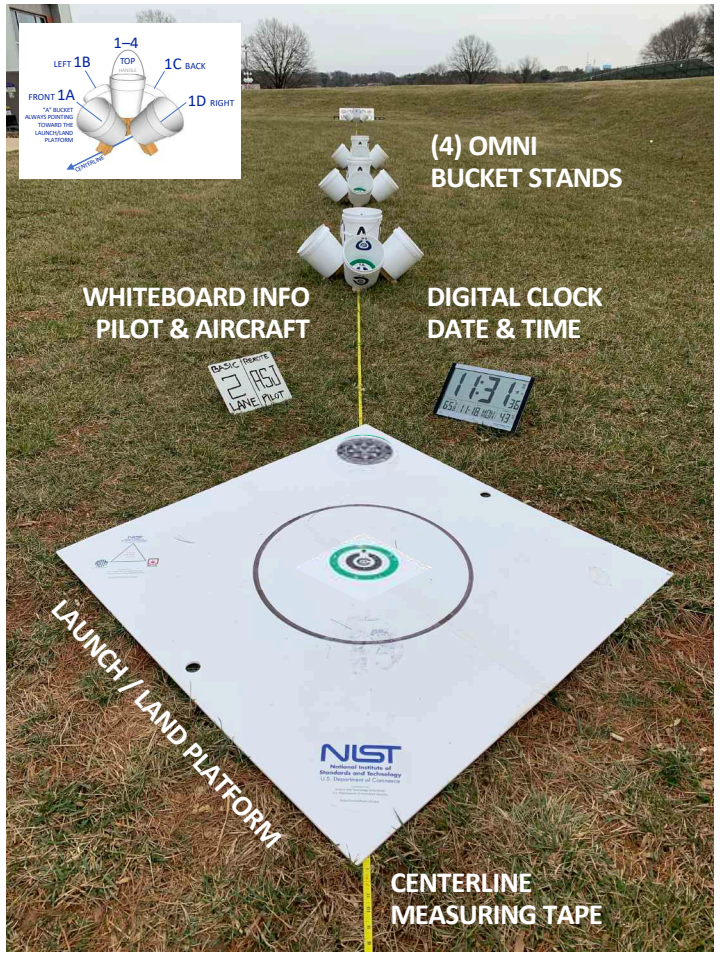


BLACK BUCKETS & COLOR RINGS EMBEDDED INTO SCENARIOS

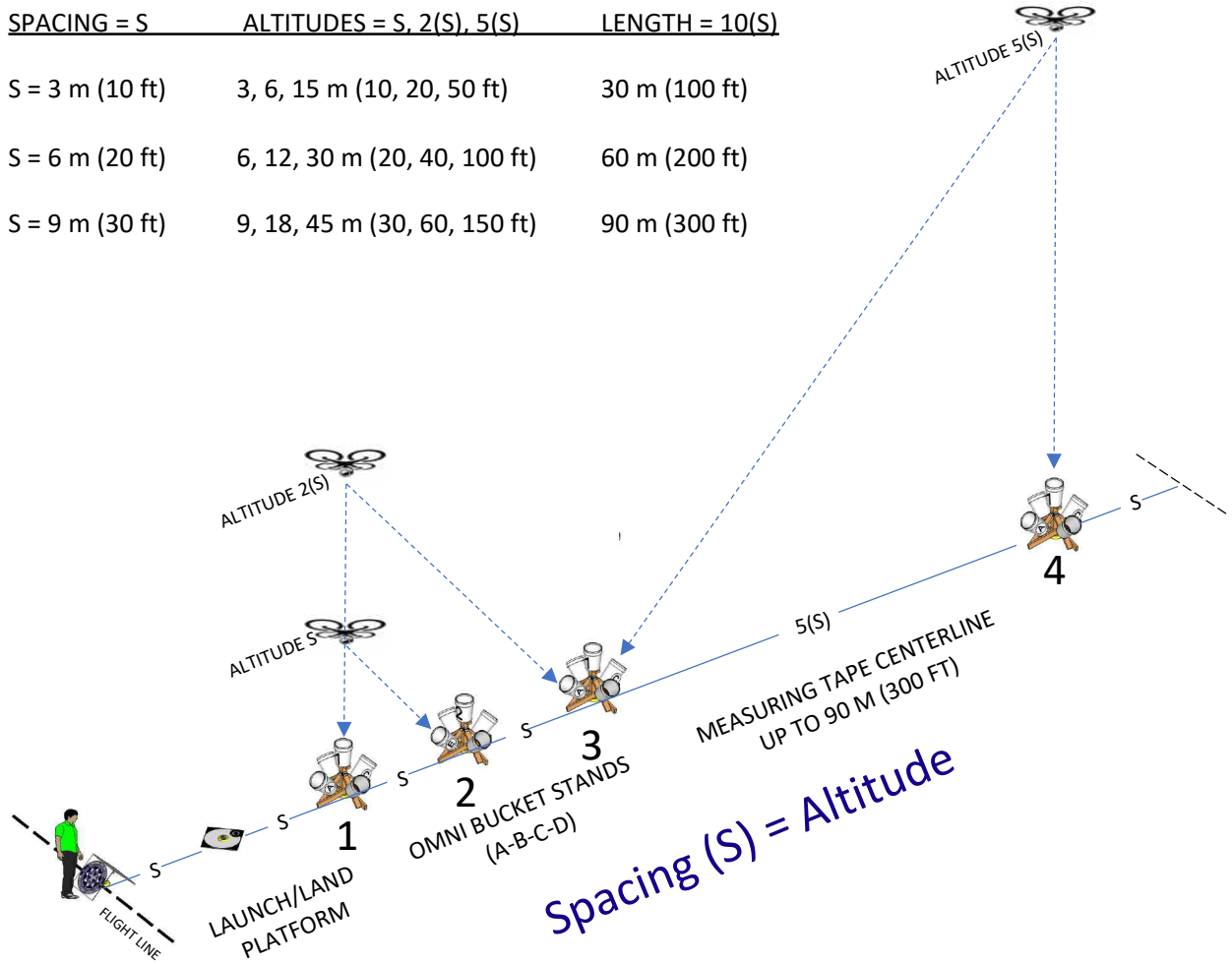
SHOWN ALIGNED

Scalable for Indoor/Outdoor Venues

Open Test Lane



SPACING = S	ALTITUDES = S, 2(S), 5(S)	LENGTH = 10(S)
S = 3 m (10 ft)	3, 6, 15 m (10, 20, 50 ft)	30 m (100 ft)
S = 6 m (20 ft)	6, 12, 30 m (20, 40, 100 ft)	60 m (200 ft)
S = 9 m (30 ft)	9, 18, 45 m (30, 60, 150 ft)	90 m (300 ft)



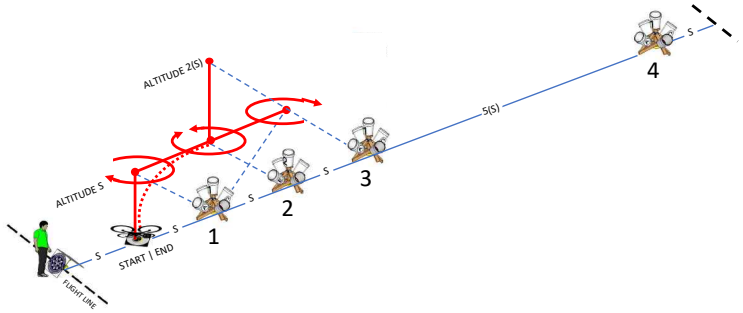
Circuit Training and Credentialing

Open Test Lane

Position

MAN/PAY 1

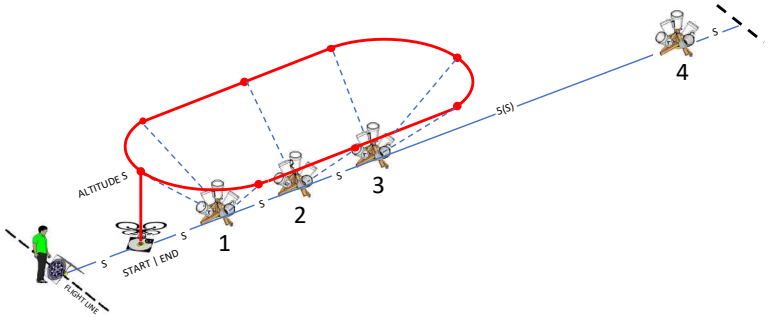
- Hover stably
- Basic maneuvers
- Land accurately
- 20 Buckets in 1 lap



Traverse

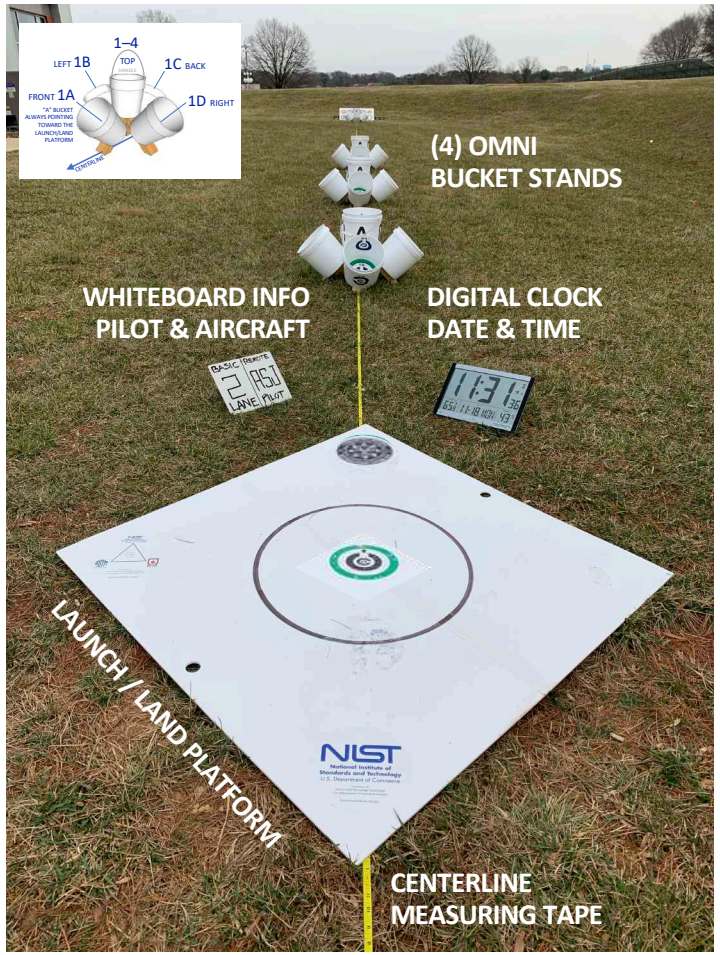
MAN /PAY 2

- Fly sideways along a line
- Left and right directions
- Land accurately
- 20 Buckets in 2 laps

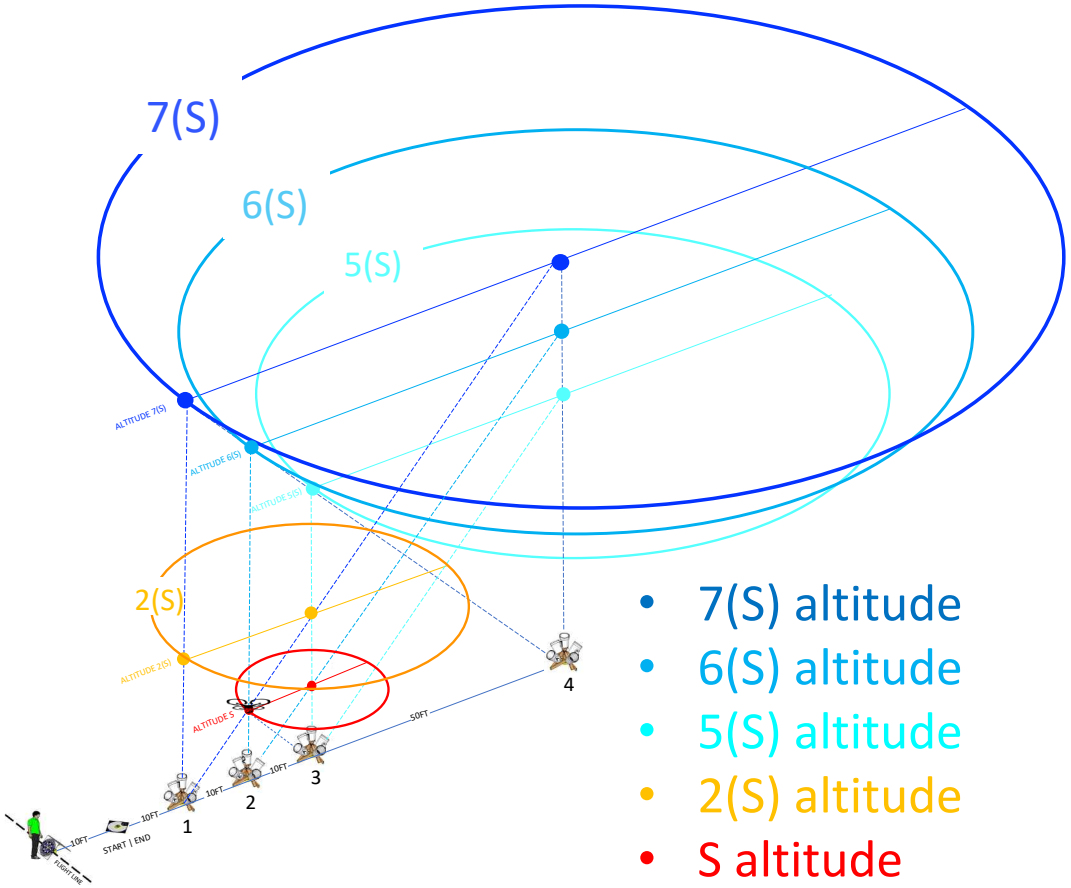


Evaluate Various Flight Paths and Sensors

Open Test Lane

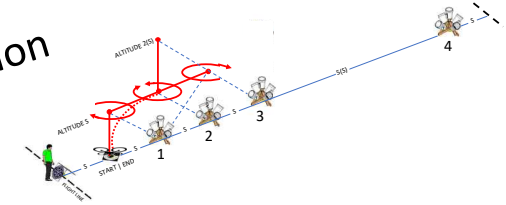


5 Different Orbits in Every Lane Spacing
(S) = 10ft, 20ft, 30ft, or other

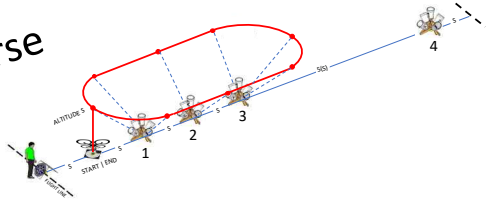


- 7(S) altitude
- 6(S) altitude
- 5(S) altitude
- 2(S) altitude
- S altitude

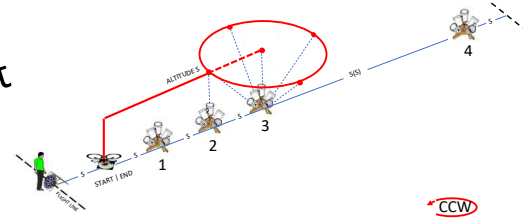
Position



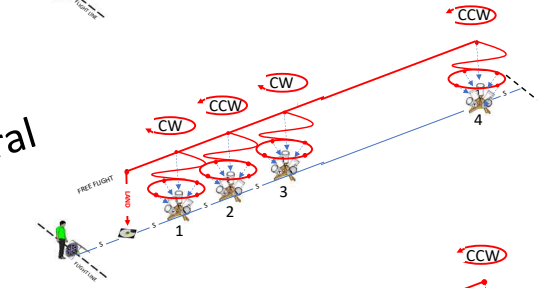
Traverse



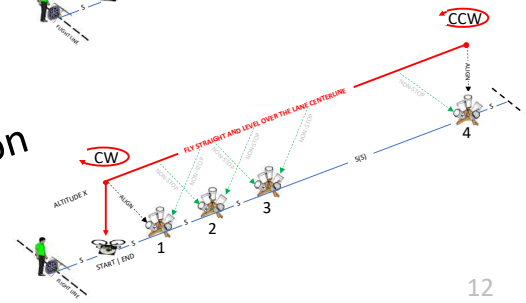
Orbit



Spiral



Recon

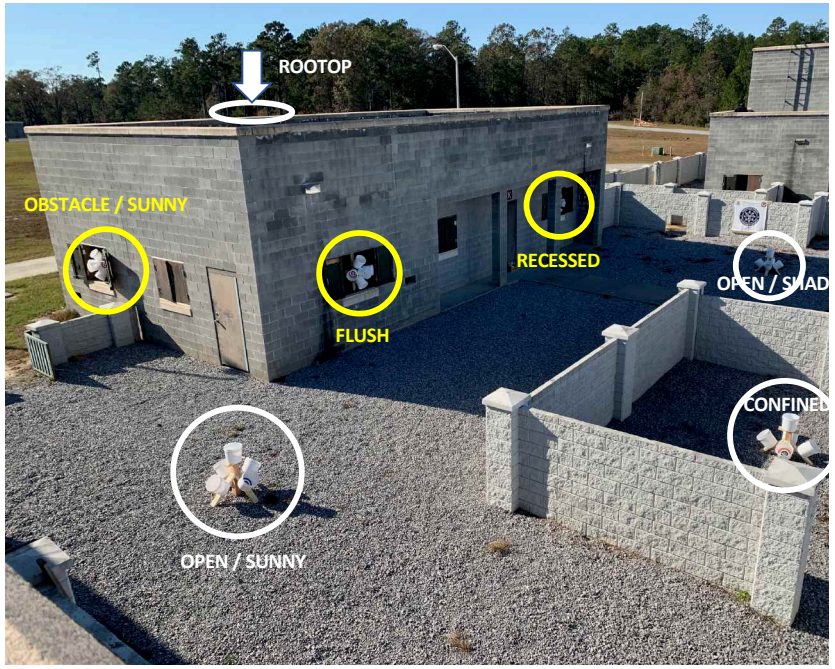


Embedded Into Scenarios as Repeatable Scoring

Training and Evaluation

WIDE AREA SEARCH
(DOWNWARD OBJECTS)

VEHICLE INSPECTION
(EXTERIOR AND INTERIOR)

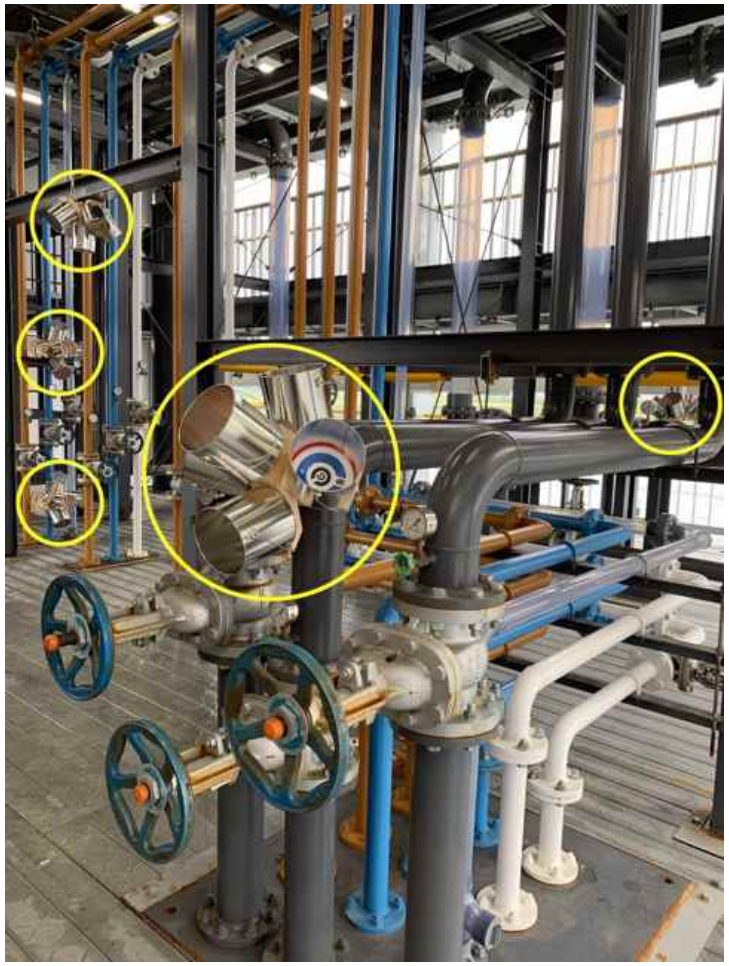


BUILDING EXTERIOR SEARCH
(DOWNWARD OBJECTS)

Embedded Into Scenarios as Repeatable Scoring

Training and Evaluation

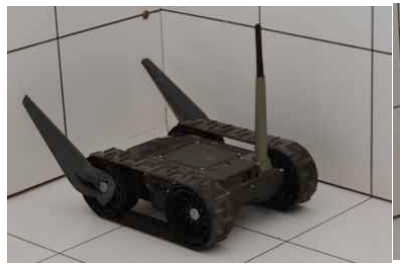
METAL BUCKETS
BLEND INTO THE ENVIRONMENT



Standard Disaster Response Robot Challenge and Plant Disaster Prevention Challenge, World Robot Summit, Japan

Test Methods Scale to Measure All Robot Sizes

Safety | Capabilities | Proficiency



iRobot 110 FirstLook
2.4kg (5.2lbs)

Qinetiq Dragon Runner 10
4.5kg (10lbs)

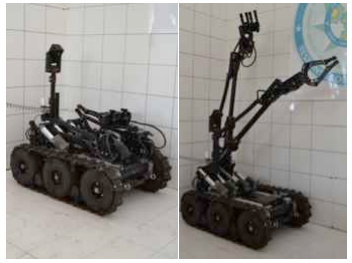
iRobot 310 SUGV
13.2kg (29lbs)

ICOR Caliber Mini
27kg (65lbs)

Remotec Titus
61kg (135lbs)

ICOR Caliber T5
64kg (140lbs)

Cobham Telemax
80kg (175lbs)



ICOR Caliber MK3
84kg (185lbs)

Remotec HD-SEL
111kg (245lb)

iRobot 710 Kobra
166.5kg (367lbs)

Remotec F6B
220kg (485lb)

WM Robotics Knight
249kg (550lbs)

Remotec Mark 5-A1
358kg (790lbs)

Remotec Wolverine
367kg (810lbs)

Howe & Howe Thermite RS1 & RS3
550kg (1200lbs) 1200 Gallons per Minute

"Start Remote, Stay Remote?"

Safety | Capabilities | Proficiency

Aerial Example

Aerial Example

Remotely controlled R/C helicopters (10 years ago). An expensive hobby!

Remote Control of all Degrees of Freedom
-- Hard to Remotely Operate --

Intelligent Control Systems

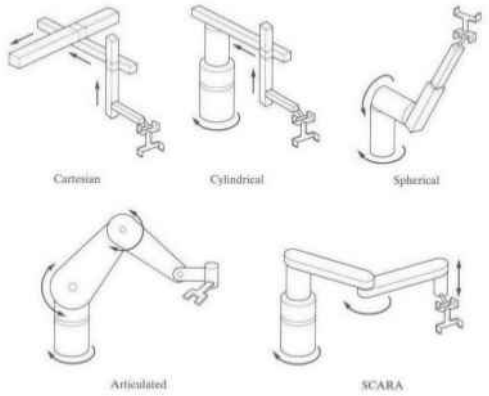
Multi-rotor systems that self level, hold station, and move in intuitive ways are much easier to fly.

Onboard Sensors and Coordinated Control
-- Easy to Remotely Operate --

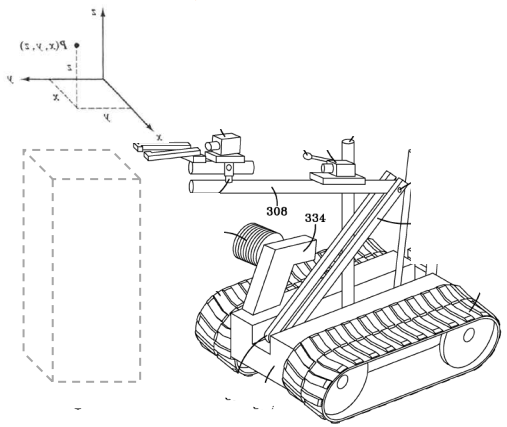
Dexterity Example

Measuring Assistive Capabilities

Industrial pedestal robotics (circa 1980's) meets mobility in rugged environments.



Rugged robot manipulators with 6 DOF coordinated control move in ways that are easy for remote operators to gain proficiency.



Intelligent Control Systems

Mobility Example

Measuring Assistive Capabilities

Intelligent Control Systems

Centering Between Obstacles, Avoiding Obstacles, Wall Following, Exploration,

Automatic Appendages to Improve Efficiency and Reliability in Terrains and Obstacles

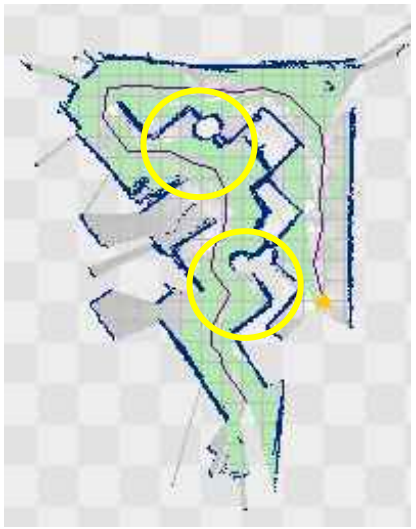
Localization and Mapping

Measuring Assistive Features

We use embedded mapping fiducials, half round shapes, on BOTH sides of walls to measure Coverage, Consistency, Local Accuracy, and Global Accuracy. This enables evaluation of ladar/lidar scanned maps of interiors with complex flooring (could be a killer app for aerials).

Low/High Fiducials

Useless Map



Broken Map

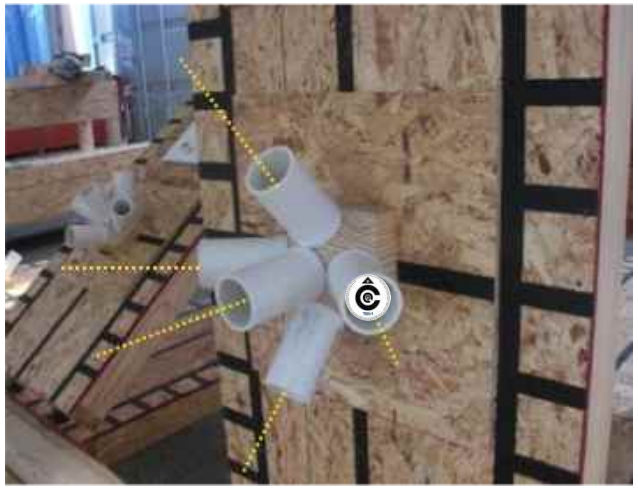
Consistent/Accurate Map

Standard Test Methods NOT Standard Robots

Safety | Capabilities | Proficiency

- **Apparatus**: A repeatable, reproducible, and inexpensive representation of a task you expect the robot to perform.
- **Procedure**: A script for the robot operator to follow.
- **Metric**: A quantitative way to measure the performance.

Visual Acuity Targets



Manipulator Dexterity Tasks

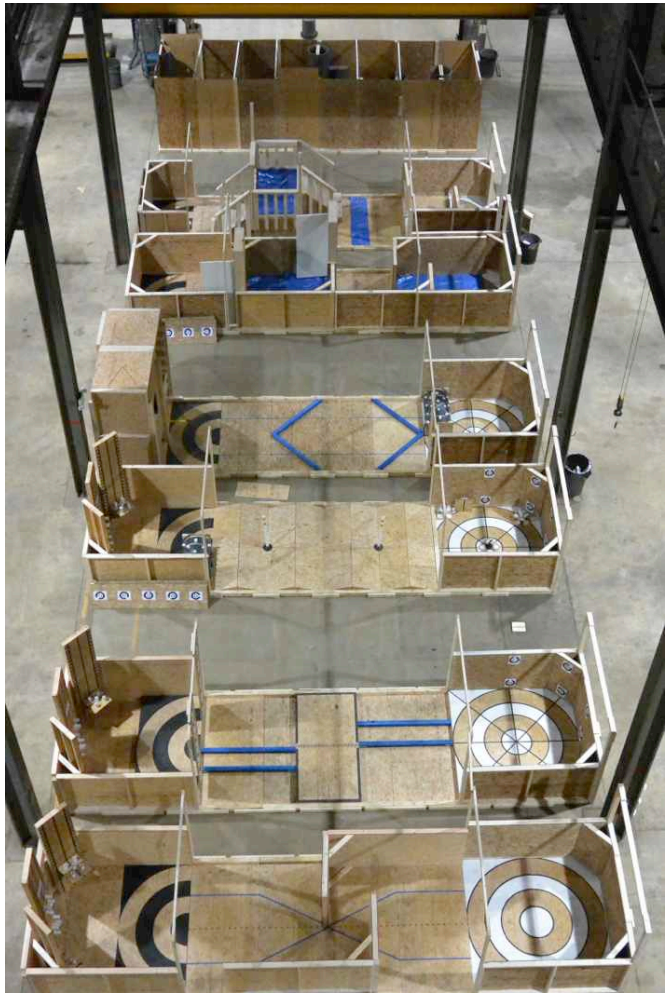
Mobility Terrains and Obstacles

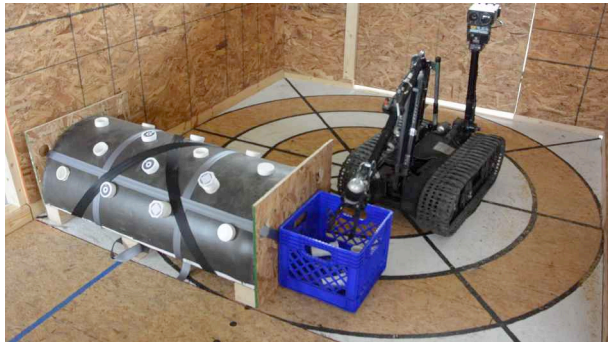
Apparatuses Scale to Intended Environments

Safety | Capabilities | Proficiency

120cm (4ft) Lateral Clearance Apparatuses

Safety | Capabilities | Proficiency

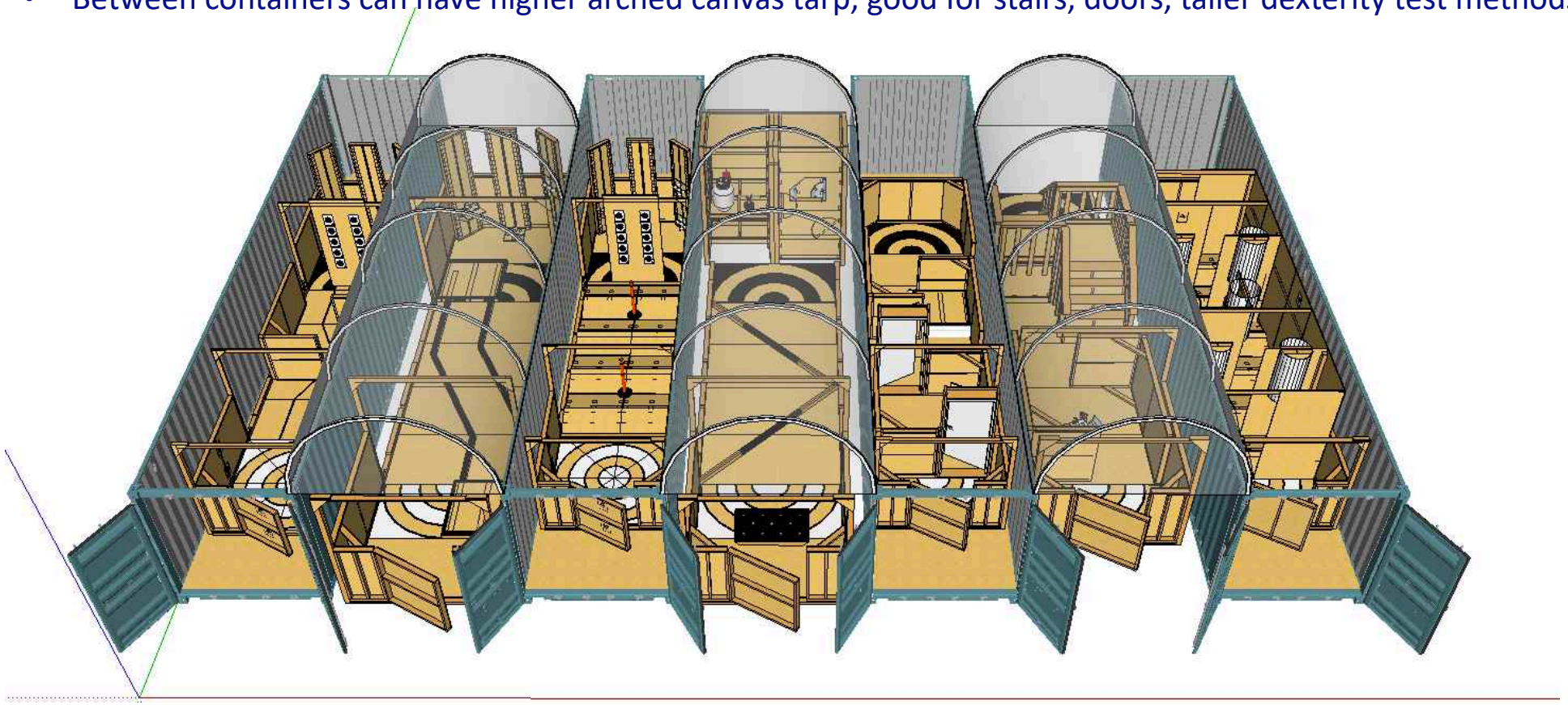




Parking Lot ISO Container Facilities (Rent or Buy)

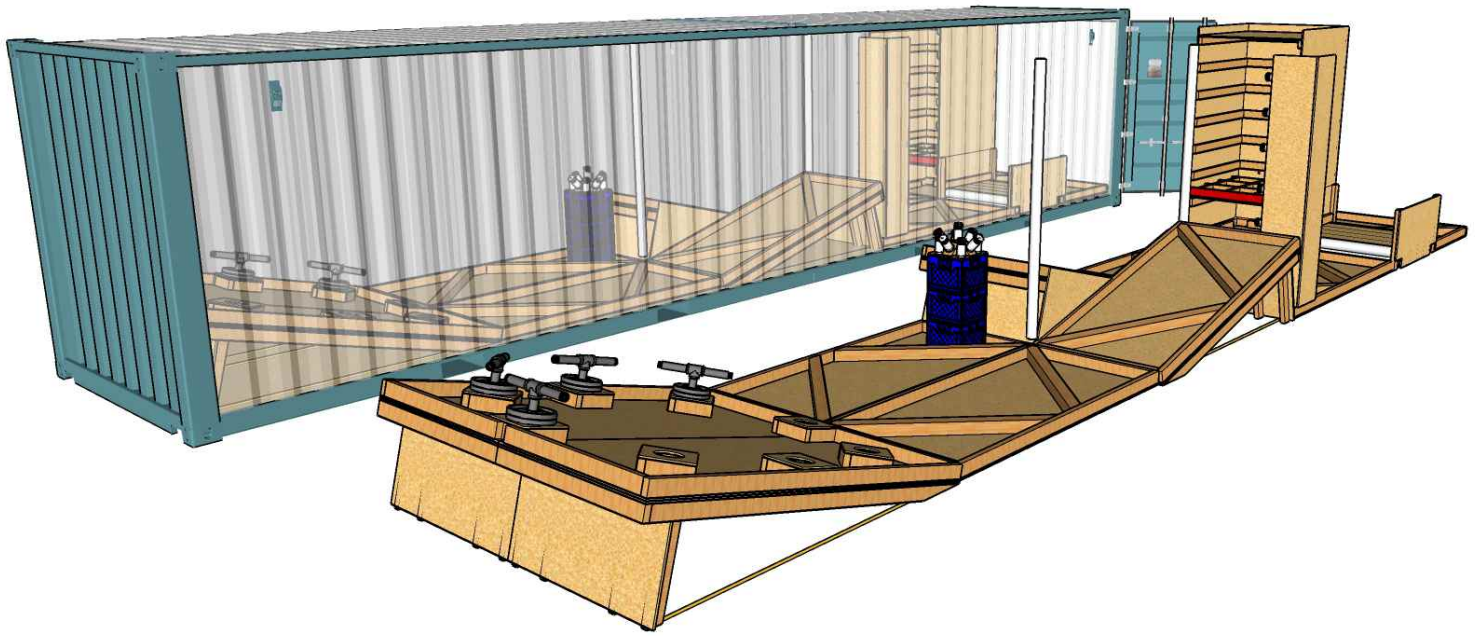
Safety | Capabilities | Proficiency

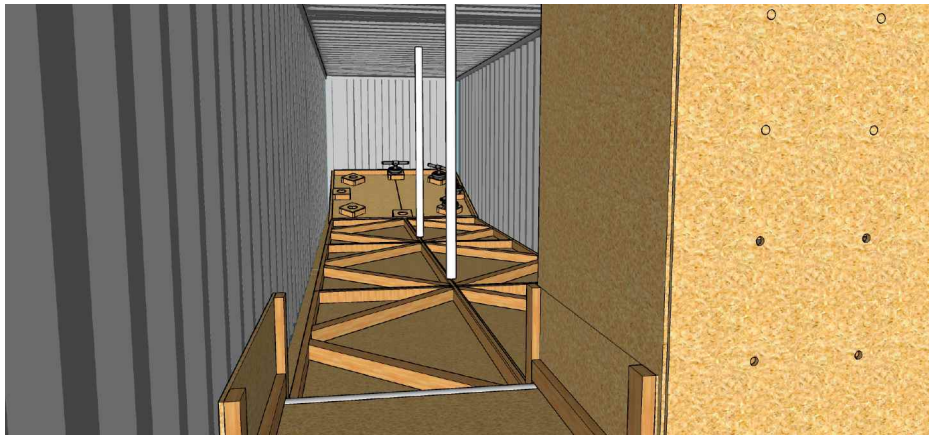
- ISOs have easy lights out testing, radio attenuated, all weather.
- Between containers can have higher arched canvas tarp, good for stairs, doors, taller dexterity test methods.

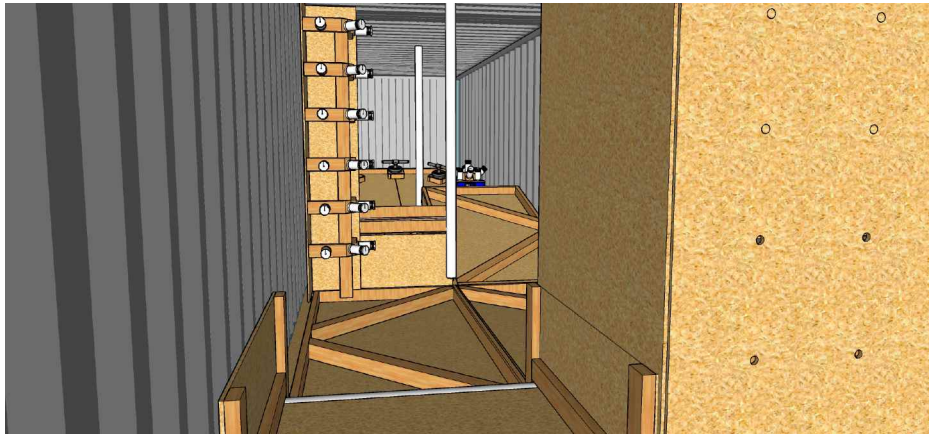


Parking Lot ISO Container Facilities (Rent or Buy)

Safety | Capabilities | Proficiency







60cm (24in) Lateral Clearance for Confined Environments

Safety | Capabilities | Proficiency

Scale: 60 cm (24 in) Lateral Clearance
Trains, Busses, Planes, Dwellings, Parked Cars, etc.

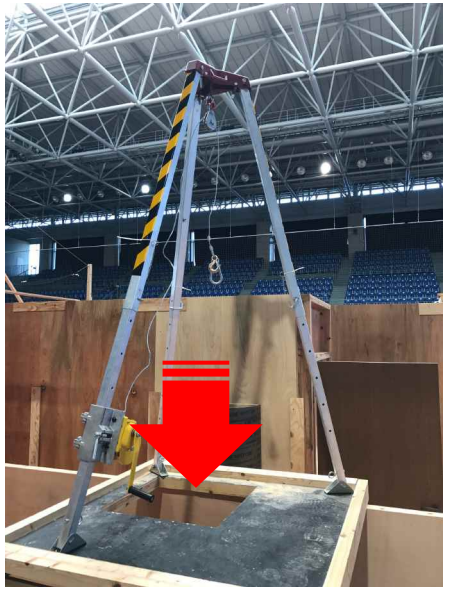
30cm (12in) Lateral Clearance for Confined Environments

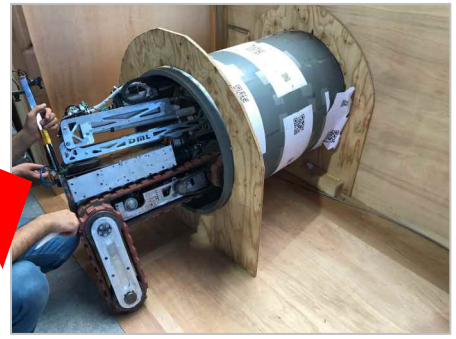
Safety | Capabilities | Proficiency

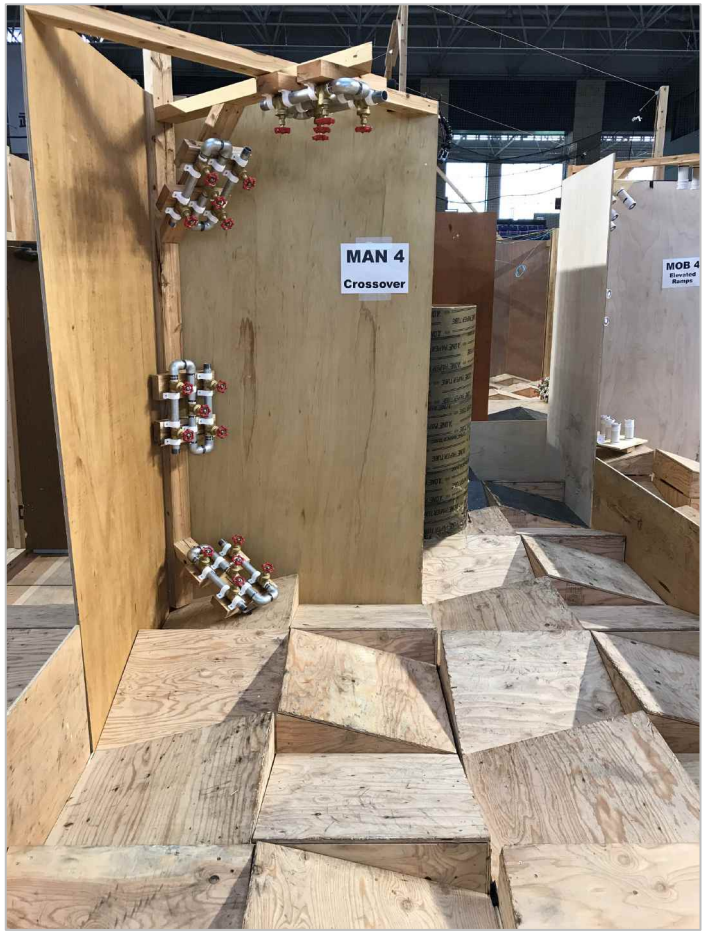
Confined Access, Throwable Robots, 3D Printed Robots (Disposable), STEM Kits

Robots Scale By Size, Weight, or Access Hole

Safety | Capabilities | Proficiency

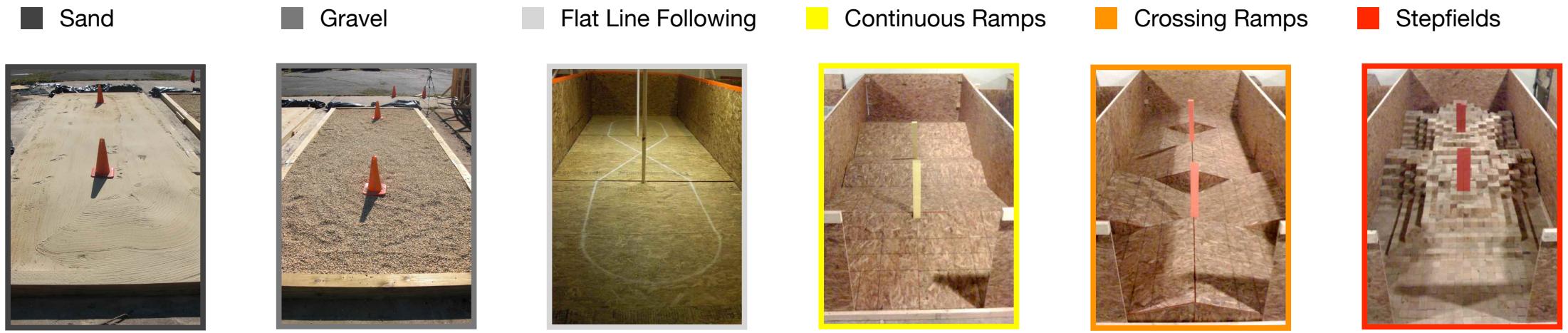






Compare Robot Capabilities or Proficiency

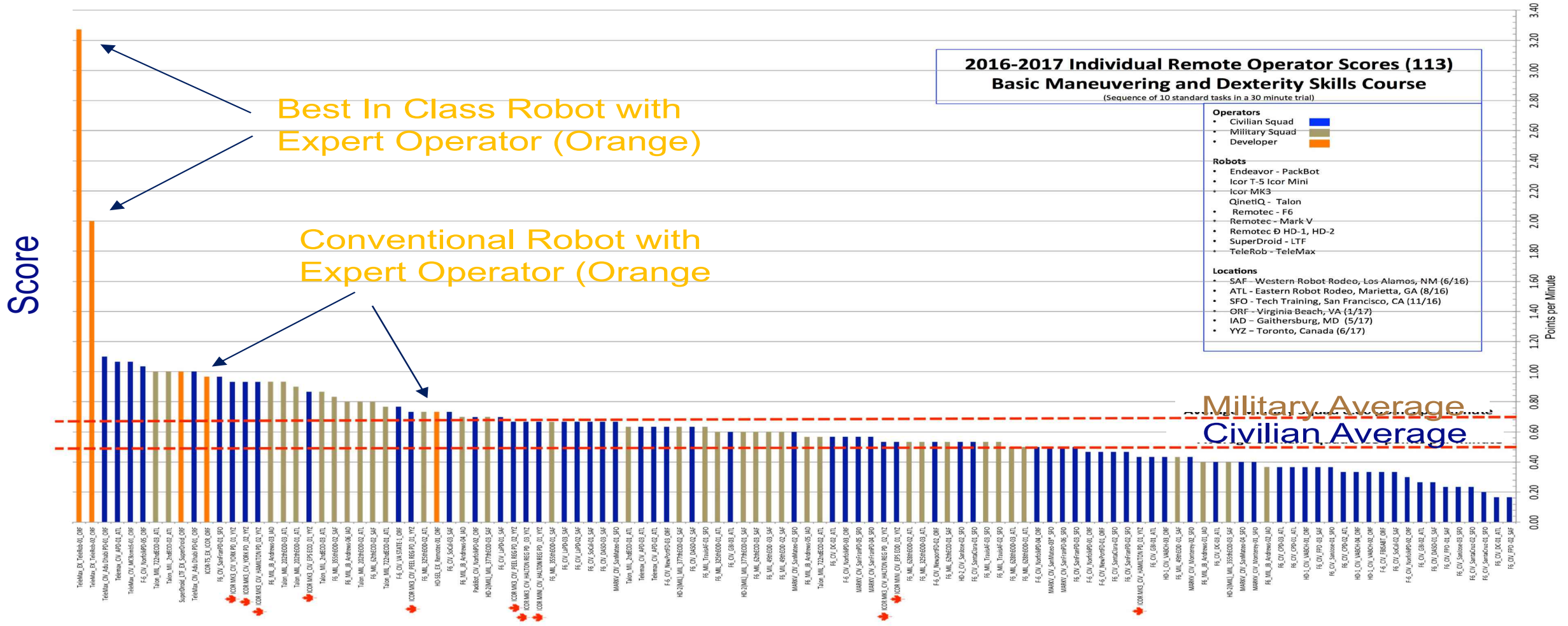
Safety | Capabilities | Proficiency



Average Rate of Advance on Terrain for At Least 100m (meters/minute)

Compare Robot Capabilities or Proficiency

Safety | Capabilities | Proficiency



100+ Bomb Techs Using Their Operational Robots (USA & Canada)

Simple Rules for Comparing Scores

Focus Training and Evaluate Proficiency for Credentialing

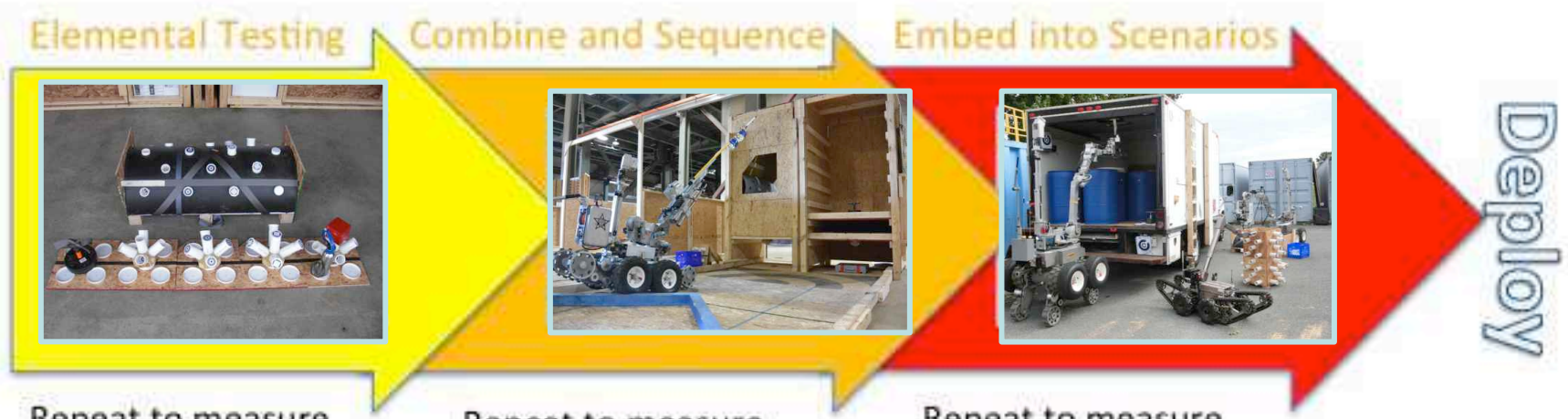
- Compare scores in the same size lanes or embedded scenarios
  Use ALL tests and scenarios that apply to the robot or to the intended missions.

- Compare scores with the same trial times
  Trial time limits help normalize fatigue across several tests, so novices don't get worn out unnecessarily. They provide enough time for an "expert" to perform a complete trial with 20 tasks. Longer times can be used for scenarios with embedded test apparatuses.

- Compare scores to "average" or "best-in-class" operators across organizations or regions
  - "Expert" operators designated by the manufacturer are used to capture the system's best possible performance in each test. Those scores are considered the 100th percentile of remote operator proficiency for that system. Your score is some percentage of that "expert" score (1-100%).
  - For training, use the entire time limit and track your scores. When they become repeatable, your learning phase is over. The best indicator of your proficiency in each test is the average of your last 5 trials.
  - We will post "average" and "expert" scores and rates for each aircraft in each test as we collect them.

Conducting Standard Test Methods

Safety | Capabilities | Proficiency



Repeat to measure ***capabilities.***	Repeat to measure ***trade-offs.***	Repeat to measure ***readiness.***
Identify and fix weaknesses.	Increase complexity in stepwise ways.	Involve uncontrolled variables.
Practice and evaluate task proficiency with quantitative scores in timed trials.	Practice and evaluate mission proficiency with quantitative scores in timed trials.	Practice and evaluate readiness with quantitative scores in timed trials.

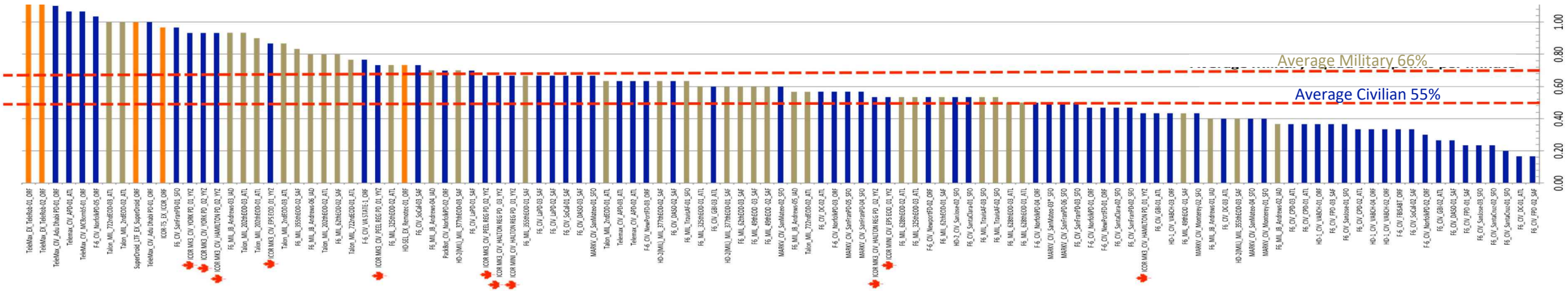
Set Your Minimum Thresholds for Pass/Fail

Focus Training and Evaluate Proficiency for Credentialing

- Organizations can set their own threshold for pass/fail in these tests based on their tolerance for reliability and/or efficiency. Complete trials are assumed.

- Measure everybody repeatedly over time and graph the results to help people understand their strengths and weaknesses. Then set minimum thresholds relative to the average or “expert” scores. Or adopt other organization’s thresholds as a central credentialing reference.

- At deployment time, each organization needs to consider their environmental variables, and mission complexity (night ops, BVLOS, etc.) to select a robot and operator that is likely to succeed.



Example proficiency data shown from bomb squads in ground robot tests

Related Awards

Project Recognition

Our same process originally applied to ground robots will work similarly for sUAS and remote pilots.

2020 Presidential Gears of Government Award

Recognizing people across the Federal workforce whose dedication supports exceptional delivery of key outcomes for the American people, specifically around mission results, customer service, and accountable stewardship.

For developing the first ever comprehensive suite of emergency response robot test methods and data collection tools to evaluate and improve bomb-disposal robots and operators. These efforts led to enhanced testing and use of advanced robot capabilities that enable emergency responders to perform extremely hazardous missions from safer standoff distances.

2019 Secretary Ron Brown Excellence in Innovation Award, U.S. Department of Commerce

The most prestigious singular honor given by the Department, chosen from among the Gold Medal Awards each year.

2019 Gold Medal Award, U.S. Department of Commerce

The highest award given by the Department for extraordinary contributions that impact key mission objectives.

Test Director:
Adam Jacoff
Intelligent Systems Division
National Institute of Standards and Technology
U.S. Department of Commerce

Sponsor:
Phil Mattson
Science and Technology Directorate
U.S. Department of Homeland Security

Internet
RobotTestMethods.nist.gov

Email
RobotTestMethods@nist.gov