

# How Training and Experience Affect the Benefits of Autonomy in a Dirty-Bomb Experiment

## Human Robot Interaction

David J. Bruemmer  
Curtis W. Nielsen  
David I. Gertman

March 2008

The INL is a  
U.S. Department of Energy  
National Laboratory  
operated by  
Battelle Energy Alliance



This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint should not be cited or reproduced without permission of the author. This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights. The views expressed in this paper are not necessarily those of the United States Government or the sponsoring agency.

# How Training and Experience Affect the Benefits of Autonomy in a Dirty-Bomb Experiment

David J. Bruemmer  
Idaho National Laboratory  
P.O.Box 1625  
Idaho Falls, ID, USA  
(208) 526-4078

David.Bruemmer@inl.gov

Curtis W. Nielsen  
Idaho National Laboratory  
P.O.Box 1625  
Idaho Falls, ID, USA  
(208) 526-8659

Curtis.Nielsen@inl.gov

David I. Gertman  
Idaho National Laboratory  
P.O.Box 1625  
Idaho Falls, ID, USA  
(208) 526-1756

David.Gertman@inl.gov

## ABSTRACT

A dirty-bomb experiment conducted at the INL is used to evaluate the effectiveness and suitability of three different modes of robot control. The experiment uses three distinct user groups to understand how participants' background and training affect the way in which they use and benefit from autonomy. The results show that the target mode, which involves automated mapping and plume tracing together with a point and click tasking tool, provides the best performance for each group. This is true for objective performance such as source detection and localization accuracy as well as subjective measures such as perceived workload, frustration and preference. The best overall performance is achieved by the Explosive Ordinance Disposal group which has experience in both robot teleoperation and dirty bomb response. The user group that benefits least from autonomy is the Nuclear Engineers that have no experience with either robot operation or dirty bomb response. The group that benefits most from autonomy is the Weapons of Mass Destruction Civil Support Team that has extensive experience related to the task, but no robot training.

## Categories and Subject Descriptors

I.2.9 [Robotics]: *Autonomous vehicles, operator interfaces, sensors.*

## General Terms

Algorithms, Measurement, Performance, Design, Reliability, Experimentation, Security, Human Factors.

## Keywords

Human-robot interaction, seamless autonomy, map-building, expert user.

## 1. INTRODUCTION

This paper describes an experiment which used three distinct groups of "expert users" to detect and localize radiological sources within a large industrial facility located in the Idaho National Laboratory (INL) critical infrastructure test range complex (CITRC). The experiment was focused on measuring the effectiveness of automated behaviors for dirty bomb detection and localization and radiation plume tracing. Additionally, the experiment provided an opportunity to evaluate how participants' background and training affected the way in which they used and benefited from robot autonomy. Some questions considered in this study are how different kinds of experts use a robot's autonomous behaviors and how the behaviors affect their performance. As a first approach, we considered two criteria by which to group users. The first is their level of robot operational expertise. In this experiment, this criterion considers how much experience the user actually had driving the robot in question – an iRobot Packbot. The second criterion is the level of training and experience with emergency response to dirty bombs and other radiological characterization tasks. With these two criteria in mind, the experiment plan divided participants into three groups: 1) Explosive ordinance disposal (EOD) soldiers trained in dirty bomb response as well as between 1 -2 years of regular, hands-on use of the Packbot in-theater; 2) Weapon of Mass Destruction Civilian Support Teams (WMD-CST) with extensive training in dirty bomb response, but no previous robotics use or training; 3) nuclear engineers (NE) with extensive understanding of radiological characterization, but no training in either robot operation or dirty bomb response. All of these participants could be considered "experts," and yet they represent very different kinds of experts. It remained to be seen how the behaviors and interface tools developed through years of iterative testing and development would perform in the hands of these three user groups.

For some time there has been discussion in the field of Human-Robot Interaction (HRI) regarding the role of novice and expert users in usability studies [2, 4, 7, 10, 12]. Of particular interest is the question of what exactly constitutes an expert and how various kinds of expertise (i.e. robotic design experience; application understanding; remote driving experience; video game playing; etc.) may affect the way various interaction schemes and behaviors are used and accepted. An early study by Marble et al. [7] observed that expert teleoperators would often resist the use of autonomous behaviors and suggested that the benefits of

autonomy might be seen most clearly in the hands of novices. On the other hand, the study pointed out that many of the complaints regarding the behaviors voiced by the expert operators were valid and that better training and system understanding might alleviate user frustration with the behaviors. Sellnar et al. [11] also pointed out that it is hard to extend results from a study with novice users performing an assembly task to actual experts such as astronauts. They hypothesize that some of the important performance tradeoffs exhibited by novices (i.e. between response time and accuracy) may not be seen once a greater level of expertise and knowledge is gained. If that is the case, the benefits of assistive robotic behaviors and support tools may prove to be less significant for experts than for novices.

We submit that both novice and expert studies can be used effectively at different stages of the system development process. Ideally, they can each be used to answer different research and development questions. Past experiments with novice users allowed us to improve the reliability and value of the basic behaviors, control scheme and interface [2]. They provided an opportunity to involve many participants, including a range of ages and abilities. This statistical power inherent with a relatively large sample size makes it possible to consider some of the more fundamental benefits and limitations to autonomy and interaction methods. Using novice studies also permits a low-cost, “short fuse” spiral of design and testing and a means to consider a variety of platforms, radios, operator control units and behaviors. On the other hand, there are limitations to what can be accomplished with novice studies. For robotic capabilities to impact real-world operations, it becomes necessary to measure performance in the hands of end-users. These studies can build on lessons learned from experiments with novices and can begin to consider a broader range of application-centric and mission-level performance criteria. The studies with expert users should ideally move towards real-world sensors on task-appropriate, and fieldable platforms.

To understand the differences between expert and novice users, it is helpful to consider the psychology literature on the subject. Bransford, Brown and Cocking [1] identified four key principles of experts' knowledge and their potential implications for learning and instruction:

1. Experts notice features and meaningful patterns of information that are not noticed by novice groups.
2. Experts organize knowledge in ways that reflect a deep understanding of their subject matter.
3. Experts' knowledge cannot be reduced to sets of isolated facts but, instead, is highly contextualized.
4. Experts are able to flexibly retrieve important aspects of their knowledge with little attentional effort.

Zimmerman and Schunk [13] explain that experts self regulate in terms of time and information and know when to apply knowledge. If this is true, it stands to reason that tools such as mapping, plume tracing, and intelligent robot behaviors might be used differently by experts than by novices and perhaps even by different sets of experts depending on how they have been trained.

## 2. SYSTEM DESIGN

In order to build and evaluate an intelligent robot system that was appropriate to the tasks and missions of emergency response personnel, we first held preliminary meetings with domain users where they were asked to express their “wish list” for future robotic technology. One of their requests was for a “click and go” behavior that would allow them to drive a robot down range without user input. Moreover, they specified that as the robot approached the critical or target area, they, as operators, wanted more direct control and less robot autonomy. Our response to this request was to provide a dynamic autonomy tasking tool for navigation that previously was shown effective in a search task with novice users [9]. The tool supports seamless autonomy which allows the operator to seamlessly change the autonomy of the robot between more direct human control and more autonomous robot control.

Another request from the domain users was that the users wanted better situation awareness of the overall area around the robot since it is possible for them to become disoriented in terms of position, orientation, and relative localization. Moreover, the users commented that they are currently using paper and pencil to draw a map of the robot's environment and the location of the hazards within the environment which is often an inaccurate and time consuming task. They reported that, ideally, a map of the environment would be produced automatically by the robot and would show hazardous areas within the map. Furthermore, users asserted that the representation should show both personnel safety thresholds as well as a dynamically scaled gradient illustrating “hotter” and “colder” areas specifically emphasizing local differences and changes. In response to this request, simultaneous localization and mapping (SLAM) techniques developed by Stanford Research Institute (SRI) were used by the robot to build maps of the environment [5]. The map of the physical environment was then augmented with an abstraction of the radiation readings in the area using a virtual 3D interface.

### 2.1 The Robot

In response to the recommendations of the domain users, the INL suite of robot behaviors, the Robot Intelligence Kernel (RIK) [2, 3] was integrated with an iRobot Packbot. To perform this integration, A RIK payload that slides into the payload bay of a standard Scout Packbot was developed that includes a 360 degree scanning laser, a 1 GHz processor, and a long range freewave radio. Additionally, a radiation detection payload was built using an AMP-50 Germanium tube-based low range radiation detector for use in low dose rate fields. The detector provides a linear response from 10  $\mu$ R/h to 4 R/h.

### 2.2 The Interface

The information from the radiation sensor is transmitted to the interface and presented via an on-screen meter, and an abstraction of the data is represented in the virtual 3D interface that we have used previously [8, 9]. Specifically, a radiation footprint the size of the robot is marked in the map where the radiation reading took place. The color of the footprint is determined based on the level of the reading. Background readings are shown in gray color, low readings are green, medium readings are yellow and high readings are red. Each level of reading also shows a gradient with lower levels being lighter and higher levels being darker to illustrate the direction towards the highest readings.



Figure 2. The modified Packbot used for the experiment.

In addition to the visual representation, the interface also made use of an audible radiation sound of a more traditional radiation meter as shown in Figure 2. To set the operating threshold of the meter, the user clicked the mouse in the bottom portion of the window on the desired level. As the meter rose from 0 to 10 for the given level, the rate of audible clicking would increase as well. Anything over 10 would saturate the meter for the given level and would sound the same until the user changed the thresholds. Figure 3 shows a map of the environment with the radiation abstractions.

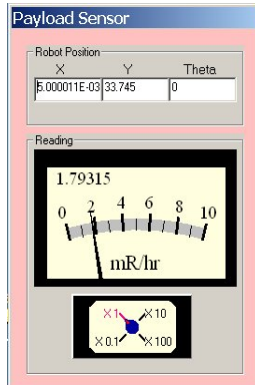


Figure 1 Software radiation meter

### 3. EXPERIMENT

To evaluate the usefulness of the robot system and the interface, an experiment was designed to address three questions with respect to the use of an intelligent robot for a dirty-bomb task.

- How does user background and experience affect performance?
- How does an automatically generated map affect performance?
- How does a robot with intelligent behaviors affect performance?

#### 3.1 Task

For the experiment participants were asked to use the INL modified Packbot and interface to explore a building in search of two radiation sources. Upon localizing the sources the users were asked to record the information on a paper floor plan that was 80% correct and return the robot to the start location. Each participant performed the experiment with three different interaction schemes that modified either the interface or the autonomy on the robot.



Figure 3 Virtual 3D interface illustrating map and abstracted radiation readings

The first condition showed the radiation meter as well as a live video feed from the robot. The operator controlled the robot via a joystick and the robot was allowed to prevent collisions with obstacles by inhibiting movement towards detected obstacles. Figure 4a shows the interface for the first condition.

The second condition was similar to the first except that the interface was augmented to include the robot's map of the environment along with the radiation footprints. The operator viewed the robot and environment information from an egocentric perspective above and behind the robot. The operator controlled the robot via the joystick and the robot was allowed to prevent collisions by inhibiting movement towards detected obstacles. Figure 4b shows the interface for the second condition.

The third condition provided a third person perspective or bird's eye view of the environment. The operator controlled the robot by moving a target icon with a mouse around the environment. Target Mode uses automated path planning to generate and follow a path towards the target icon (see Figure 4c and [9]). The farther the target is moved from the robot, the more initiative is granted to the robot in terms of which behaviors are used. For instance, when the target is close, the robot does not perform path planning, whereas further targets involve the robot planning a route. The study by Nielsen and Bruemmer [9] shows the value of the target mode when used by a novice population to perform visual search and detection tasks

#### 3.2 Participants

We invited 18 participants who could all be classified as domain experts to participate in this study. The participants were divided into three groups based on prior training and experience. The first group of participants included seven Explosive Ordnance Disposal (EOD) personnel from the Army, Navy, and Air force who all had prior experience with in-field dirty bomb response and in-theatre robot experience using iRobot's Packbot. The second group included five Weapons of Mass Destruction Civil Support Team (CST) personnel and one combat engineer who all had prior experience with dirty bomb response but no prior robot experience. The third group included five Nuclear Engineers (NE)

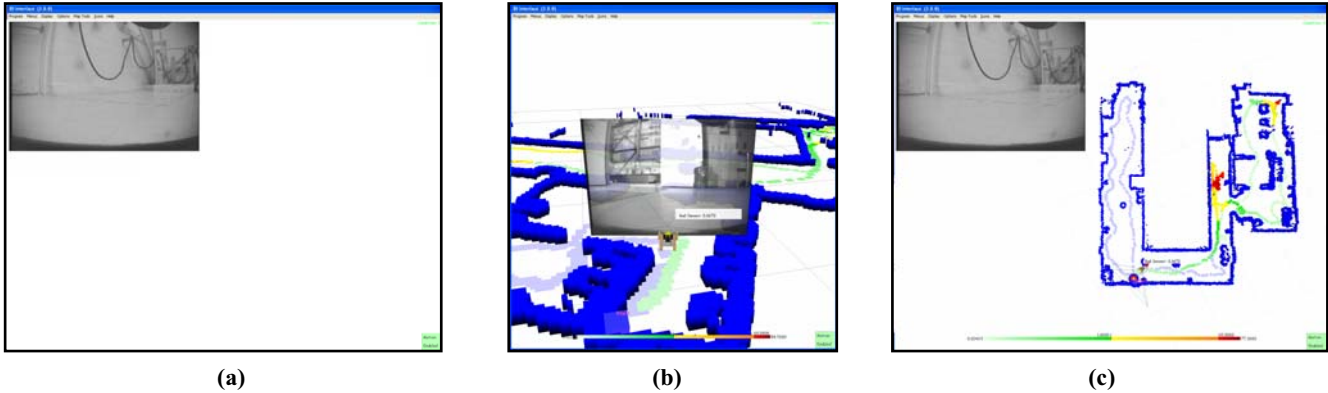


Figure 4. Interfaces conditions used for the experiment: a) Joystick, b) Joystick + Map, c) Target + Map

who had experience with radiological hazards but no experience with emergency response, dirty bombs, robotics, or operation of radio-controlled vehicles.

### 3.3 Environment

The facility for the experiment was an abandoned building at the INL that had a main hallway with a number of side rooms. The environment was mostly planar with large obstacles placed throughout the environment. The rooms were occupied by common office furniture and some mechanical equipment. Figure 5 shows the facility and some of the rooms within the facility.

### 3.4 Experiment Design

Participants wirelessly operated the robot from within a trailer located outside and near the building. Subjects remained inside the trailer during their experiment and were not allowed to enter the building until the end of the experiments for the day. For each test condition, two radiation sources were placed inside the building, one in a relatively open area and the other in a more hidden location. The radiological materials used for the experiment consisted of two sealed Cesium-137 sources each with source strength of approximately 24 millicuries (mCi). Each source produced a Geiger-Mueller (GM) ionization detector reading of approximately 2.1 R/hr at one foot and 10 mR/hr at 7.5 feet. One source was shaped like a pellet, about 0.75 inches long and 0.5 inches in diameter. The second source was attached to the end of a rod that was 0.5 inches in diameter and 16 inches long.

The exact locations of the radiation sources were changed between conditions. Participants were informed that the paper map on which they were to illustrate their understanding of the location of the sources was only about 80% correct but it was the most recent intelligence available. Figure 6 shows the map of the facility that we provided the users. The grayed out areas were not used in the experiment and the top left room was used for training the operators. The top right set of rooms had incorrect walls as illustrated by the drawing where the crossed out walls were not really there and the black wall should have been there but was not on the map. The illustration also shows the four doors that were used as starting places for the robot. For the training session, the robot was placed inside door C. For the conditions of the experiment, the robot was placed alternately between doors A, B,

and D. When the robot started at the A location, the door on the left of the main hallway was closed and the sources were hidden at locations A1 and A2. When the robot started at the B or D doors the door on the right side of the main hallway was closed and the sources were hidden at locations B1, B2 or D1, D2 respectively. The order of test conditions and assignment of subjects were randomized and counterbalanced to minimize learning effect.

### 3.5 Training

All subjects were provided both classroom training and an eleven page manual. Additionally, each participant was given supervised hands on experience with all three conditions prior to the experiment. These hands on training sessions were performed from the same trailer as was used in the experiment and utilized the same robot and same interface. However, the training took place in the boiler room section of the environment (Door C) which was not accessible during the experiment. During the training, a single source was used to provide the user with the experience of actually finding and localizing a source. These hands on training sessions were generally completed in less than 15 minutes. Participants verified that they felt they received sufficient training prior to completing the training and starting the experiment. Participants were advised to place equal importance with logging the radiation levels as with pinpointing the source locations.

### 3.6 Metrics and measurements

For the experiment, a variety of objective performance criteria were developed including how many sources the participants found and the accuracy with which they could pinpoint the position of this source within a paper floor plan of the building. Another metric was the accuracy with which the participants could identify the start location within the floor plan. The participants were not told where the robot was started and as a result it often took several minutes of exploration before the user was able to identify where they were in the map and, by extension, where they had started. This proved to be a useful indication of the awareness the user had regarding the robot's position and also of their general understanding of the environment. For both of these metrics, significant effort was taken to develop an appropriate scoring method. Each source and start location was scored by the same two researchers using a





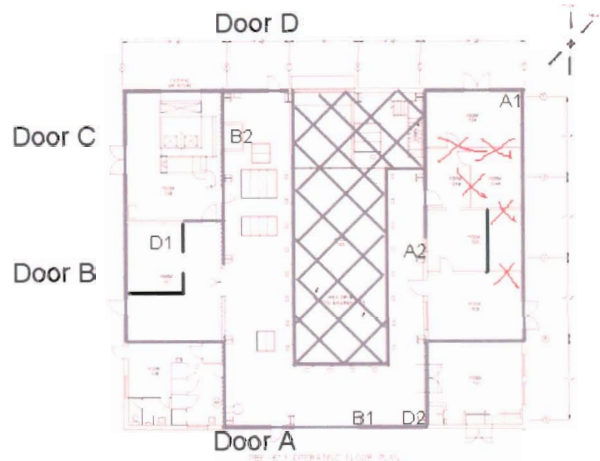
**Figure 5. Facility used for the experiment**

rubric that produced a numerical score up to 20 points. The rubric used distance between the real and marked location and also related the specified location to the nearest feature such as a corner or obstacle. For example, the score would decrement by a point if the source was marked in the wrong corner. Scores below 12.0 were placed in the “Fail” category which meant that the operator had minimal understanding of the environment.

A variety of metrics focused on human-robot interaction were developed including the percent of time that the user was actively interacting with the interface (% Active Time); the number of times the operator gave erroneous or impossible commands to the robot (Operator Confusion); the number of times the robot was required to prevent a collision (Collisions). Note that all of these metrics were normalized to account for differences in the sizes of the areas searched. We were also interested in understanding participants’ subjective feelings toward autonomy and, specifically, how they rated their understanding of the system, mental demand, physical demand, effort, frustration and feeling of having been in control. These metrics were assessed using self-assessment questionnaires developed by a cognitive psychologist on the team. All workload measures were determined by administering the NASA Task Load Index (TLX).

#### 4. RESULTS

A 3 x 3 factorial analysis of variance followed by a Tukey multiple comparison procedure was performed to analyze the results. When comparing subjective assessment scores across the EOD, CST and NE user groups several overarching differences were noted. For example, those in the NE group recorded a significantly higher workload than those in the EOD/CST groups ( $p=0.0004$ ). The average workload for NE was 8.7 on a 20-point scale, when compared to the EOD/CST groups (4.6/5.9). Findings for group differences on mental demand ratings were also significant, ( $p=0.0001$ ). Those in the NE group reported a mental



**Figure 6. The floor plan of the environment, showing the location of the radiation sources**

demand significantly greater than the other two groups; mean mental demand for NE =12.4, EOD = 5.4, and CST = 8.2 on a 20-point scale. Effects for the frustration scale associated with workload were also significant ( $p=0.004$ ). Those in the NE group experienced an average frustration of 9.3, (on a 20-point scale) while those in the EOD/CST group only experienced a frustration of 4.3/6.0.

Trends regarding operator performance were noted as well. Performance self report ratings were based upon a scale ranging from 1 - 20 where 1 represents the best possible performance. When these scores were summed across display conditions, those in the NE group reported an average of 8.7, while those in the EOD/CST groups reported that they performed better (5.3/4.9). In terms of other self-assessments including situation awareness measures, analysis of the data did not show significant background effects for how participants reported awareness of environmental surroundings or their comprehension of the robot’s actions.

Group effects for navigational error instances per 1,000 square feet were also significant ( $p=0.01$ ). The EOD group (0.99) has a significantly smaller number of navigational error instances than do the CST/NE groups (2.07/2.05).

There was not a significant difference between operators having and not having a map. Reviews of the data do suggest that there are several fundamental performance differences attributable to intelligent behaviors (Target Mode). Across all user groups, the use of Target Mode reduced the percent of time that users spent interacting with the interface ( $p=0.0001$ ). When participants were in Target Mode, they averaged a significantly smaller percentage of time enacting commands with either the joystick or mouse (36%) than in Joystick/Joystick + Map modes (81%/83%). Also, across all user groups, operators averaged significantly fewer instances of confusion ( $p=0.003$ ) when using Target-Mode (1.3) than those in Joystick/Joystick + Map (5.1/.5.8.). Effects for navigational error instances as a function of control mode were also significant, ( $p= 0.0009$ ). Collisions were less likely in the

Target Mode. The robot took initiative to protect itself from collisions approximately three times less when using Target Mode (0.7) compared to Joystick/Joystick + Map (2.0/2.1).

Radiation source localization data was also examined for performance effects based on control mode and group. The mean average source localization scores presented in Table 1 provide strong evidence that the use of autonomy and mapping improves source localization accuracy. Adding the map also increased performance from an overall average of 14.2 for the Joystick condition to 15.5 for the Joystick + Map condition. Using the Target Mode increased this aggregate score to 17.7.

Further inspection of Table 1 reveals that this improvement was most pronounced for the CST user group which attained a 10.2 (Fail) average for Joystick Mode; and 17.8 (Good) for Target Mode (an “Excellent score” was determined to range from 18 to 20). The EOD group also benefited from using Target Mode, but due to their initial good performance the differences between conditions are less pronounced. On the other hand, NE users do not show any benefit to using the map and show very little improvement when using the target.

Although subjects differed in terms of performance scores; robot initiative; operator confusions; and subjective workload and preference measures, the results showed no statistically significant difference between modes or user groups for the time to completion, i.e., time to localize the two radiation sources and exit the building. On average, participants took about 15 minutes to complete the task. Time was not considered an important metric by the experimenters since participants had been told to emphasize source detection and localization reporting rather than speed.

Overall, two participants preferred Joystick mode, three preferred Joystick+Map mode and thirteen preferred Target Mode. Interestingly, 11 of the 13 EOD/CST end users (85%) preferred Target Mode whereas in the NE group, only 40% (2 out of 5) preferred Target Mode.

## 5. DISCUSSION

As expected, the EOD group’s considerable field experience with robot teleoperation seems to have made them better joystick drivers than the other two groups. They committed less than half of the navigational errors than the other two groups that had no experience with robot teleoperation. When using Joystick mode, EOD personnel also performed significantly better than the other two groups on the task of localizing the sources.

When localizing sources, the use of Target Mode was helpful across all user groups, but seemed to be most helpful for the CST group that had no prior experience with robots, but years of training regarding how to search for and localize radiological hazards. In Joystick mode the CST group received an average localization score of 10.2. However, when these same participants are given the Target Mode, they achieved an average localization score of 17.8. Interestingly, the autonomy raises the average performance of the CST group to just below what the EOD user group achieved when using the Target Mode. This implies that by

**Table 1. Average Map Localization Scores**

	Joystick Only	Joystick + Map	Target Mode
Explosive Ordnance Disposal (EOD)	16.7 Good	17.0 Good	18.4 Excellent
Weapons of Mass Destruction Civil Response Team (CRT)	10.2 Fail	15.3 Fair	17.8 Good
Department of Energy Nuclear Engineers (NE)	15.4 Fair	13.5 Poor	16.2 Good
Overall	14.2 Fair	15.5 Fair	17.7 Good

providing the right interaction tools, it is possible, in terms of source localization, to eliminate the robotic training effect that separates the “robot unexperienced” CST user group from the “robot experienced” EOD group.

Readers should remember however, that performance differences in robot initiative and in active time as well as preferences and estimates of workload dimensions still exist between the individual groups. Debriefing questions indicated that the Target Mode encouraged the CST and EOD participants to enact the highly effective source localization strategies their military training had taught them to use. Put simply, they believed that Target Mode allowed them to think about the task in the same terms that they had been trained with over their years of experience. Joystick Mode did not support this carry-over of the task to their training. On the other hand, the NE users had no such training to fall back on and either created their own exploration / localization methodology or simply wandered with no observable method. This may be why some of the NE users did not benefit from the use of autonomous behaviors. As an example, one of the lowest performers on the localization task was an NE participant who, when in target mode, tried to exit the building, hoping to assess radiation levels from the exterior of the building. He experienced significant frustration when this strategy proved ineffective.

The original hypothesis concerned the benefits of autonomy to those who have extensive training and expertise. The question was whether autonomy would provide as much benefit to experienced and highly trained users as it has done for previous studies with novices. The results indicate that autonomy can provide benefits to experts both in terms of objective performance and subjective measures. To assess the benefit experienced by the most skilled users, a separate analysis considered the percent of EOD localization ratings which were considered excellent (scores of 18 or above) in the different control modes. While using Joystick Mode, 57% of the EOD group received a score in the “Excellent” category (i.e. within roughly one meter of the real source), in the Joystick + Map Mode, 67% received an “Excellent score,” when using Target Mode, 100% of the EOD personnel achieved a score within this category. It appears that autonomy benefits even those users who are the most experienced and well trained.

One of the most important questions for the stakeholder of the experiment was the question of whether autonomy could be used to address the high variance between current emergency response operators. Put simply, the question was whether autonomy would enable tighter bounds on the expected performance of the end users (EOD/CST). When we consider the scores for localization accuracy, the standard deviation for EOD teams was 5.36 for Joystick, 4.26 for Joystick + Map and 1.40 for Target Mode. Likewise, for CST the standard deviation was 5.65 for Joystick, 6.98 for Joystick + Map and 1.33 for Target Mode. This indicates that the introduction of autonomy did indeed provide a tighter distribution that enables us to place tight bounds on the expected performance of the end-users. Figure 7 shows the scores for localization accuracy for the end-user group consisting of EOD and CST personnel. For Target Mode, there is not a single target user (EOD & CST) who scores below a 15 for source localization accuracy.

The study also points to major differences in the subjective experience of the user groups. In debriefs, the CST and EOD groups explained that they are trained to handle a great deal of stress and workload when dealing with a dirty bomb scenario. This may explain why the NE group reported significantly higher overall workload, frustration and mental demand than CST or EOD groups. In fact, the reported frustration experienced by the NE group was twice that of the expert EOD group. Another notable result is that the end-users overwhelmingly preferred the more autonomous conditions whereas the NE group did not. This contradicts the stakeholder hypothesis that users with teleoperation expertise would be less likely to prefer the autonomous control to those without teleoperation experience.

## 6. CONCLUSION

Previous research at INL indicated that it is important to understand how the benefits of autonomy may differ among users with different backgrounds [7]. To understand differences in how autonomy affects the various user groups, it is helpful to consider psychology literature comparing novice and expert users. Marianne La France writes that “One characteristic yet elusive difference is that experts seem to focus more on goals than do novices. For example, when presented with a case description, experts are more likely than novices to see it in terms of what plan is being put into effect rather than on what has just taken place” [6].

LaFrance goes on to explain that this ability to recognize situations based on experience and orchestrate task elements accordingly is one crucial factor that separates experts from novices: “Experts have greater episodic memory than novice groups... In contrast to semantic memory, which can be described as the knowledge of facts, hierarchically-arranged, episodic memory is the knowledge of situations culled from experience. So, even if a novice were to possess as much semantic memory of a domain as an expert, the expert would still be more knowledgeable about how to use it” [6]. This is consistent with the more experienced EOD group navigating and localizing sources better than the other two groups.

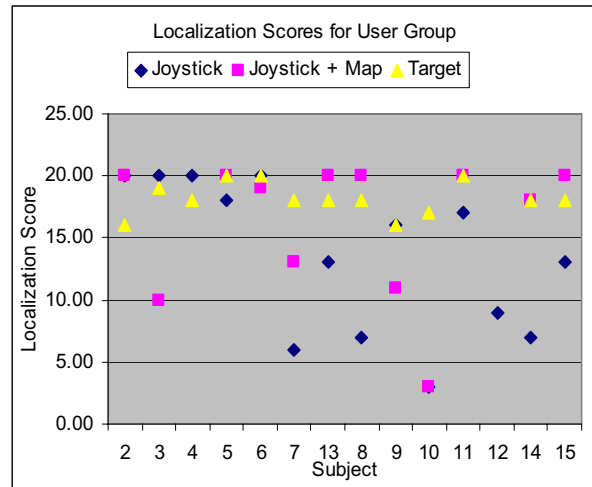


Figure 7. Comparison of the localization scores for each of the end-users.

In general, the Target Mode makes better use of the CST and EOD participants’ episodic memory because they tend to think about the process and the strategy instead of discrete tasks and direct robot control. Although it is hard to be sure, we suspect that the map-based, “intentional control” provided by the Target Mode supports strategic thinking, thereby enabling EOD and CST users to realize the full benefit of their prior experience and training. The NE group had less experience to bring to the task, experienced higher workload, greater frustration, and more navigational errors than did the EOD and CST participants.

The results indicate that significant performance benefits can be realized by providing autonomous robotic capabilities to the user community. The use of a map alone without intelligent behaviors is not sufficient to gain significant performance differences. While these results are encouraging, it is not the case that autonomy will always be preferred by end-users. Other studies performed at the INL indicate that end-users may resist the use of autonomy if they cannot develop an accurate mental model that could predict robot behavior [7]. In the current study, the use of autonomy seemed to support the experts’ mental model.

We believe the Target Mode facilitates goal-based thinking and tasking on the part of the user. In formal debriefs, participants felt that using the Target Mode allowed them to focus more on their strategy for detecting and localizing the dirty bomb than on the navigation of the robot, whereas the Joystick condition required greater attention to navigation. A representative comment was: “Target [mode was my preferred mode of operation] because it freed me to focus on radiation. I could trust the target to get between areas.” Likewise, analysis of the %-active time indicates that during Target Mode operation, users spent less time interacting with the interface. They also reported in debriefings that they spent more time thinking about how to sequence the overall search plan.



In the summary, all participants were able to use the tools developed for this effort to complete the tasks and felt the field study to be realistic. When using Target Mode every one of the EOD and CST users was highly successful at localizing both sources (i.e. scored 16 or higher). Perhaps more significantly, every one of these users reported that they would use the system in a real mission. In terms of future work, the resulting system is being adapted for chemical sensing by the Edgewood Chemical and Biological Center in Maryland and will be rigorously evaluated by the Chemical School at Ft. Leonard-Wood for fielding with a Chemical, Biological, Radiological and Nuclear (CBRN) Unmanned Ground Vehicle.

## 7. ACKNOWLEDGMENTS

Our special thanks to Scott Hartley for his work on the incredible logistics required to gather the participants and perform this experiment. We would also like to thank the participants who provided so much valuable insight into how they view the role of robotics in their daily operations. Additionally, we want to thank Tophier McFarland, Alex Gertman, and Tim Klein for their help with the data analysis.

## 8. REFERENCES

- [1] J. D. Bransford, A. L. Brown, and R. Cocking (eds.). *How People Learning: Brain, Mind, Experience, and School*. Committee on Developments in the Science of Learning, National Research Council, 1999.  
<http://www.nap.edu/html/howpeople1/>
- [2] D. J. Bruemmer, D. A. Few, R. L. Boring, J. L. Marble, M. C. Walton, and C. W. Nielsen. "Shared Understanding for Collaborative Control." *IEEE Transactions on Systems, Man, and Cybernetics, Part A. Systems and Humans*, vol. 35, no.4, pp. 505-512, Jul. 2005.
- [3] D. J. Bruemmer, D. I. Gertman, C. W. Nielsen, D. A. Few, W. D. Smart. "Supporting Complex Robot Behaviors with Simple Interaction Tools." Accepted for publication in *Human Robot Interaction*, ed. Nilanjan Sarkar, by Advanced Robotics Systems International, Vienna, Austria. 2007.
- [4] J. L. Burke, R. R. Murphy, E. Rogers, V. J. Lumelsky, and J. Scholtz, "Final report for the DARPA/NSF interdisciplinary study on human-robot interaction," *IEEE Trans. Syst., Man, Cybern. C, Appl.Rev.*, vol. 34, no. 2, pp. 103-112, May 2004
- [5] K. Konolige. Large-scale map-making. In *Proceedings of the National Conference on AI (AAAI)*, San Jose, CA, 2004.
- [6] M. LaFrance. "The quality of expertise: implications of expert-novice differences for knowledge acquisition," *ACM SIGART Bulletin*, vol. 108 pp. 6-14, 1989.
- [7] J. L. Marble, D. J. Bruemmer, D. A. Few. "Mixed-Initiative Control for Remote Characterization of Hazardous Environments," In *Proc. of the 2004 Hawaii International Conference on System Sciences (HICSS)*, Waikoloa Village, Hawaii, January 2004.
- [8] C. W. Nielsen and M. A. Goodrich. Comparing the usefulness of video and map information in navigation tasks. In *Proceedings of the 2006 Human-Robot Interaction Conference*, Salt Lake City, UT, March 2006.
- [9] C. W. Nielsen and D. J. Bruemmer. "Hiding the system from the user: Moving from complex mental models to elegant metaphors," In *Proc. of the IEEE International Symposium on Robot and Human Interactive Communication (Ro-man)*, Jeju Island, South Korea, August 2007.
- [10] J. Scholtz, "Human-robot interactions: Creating synergistic cyber forces," in *Proc. Workshop Multi-Robot Syst.*, 2002, pp. 177-184.
- [11] B. Sellner, F. W. Heger, L. M. Hiatt, R. Simmons, and S. Singh, "Coordinated Multiagent Teams and Sliding Autonomy for Large-Scale Assembly," *Proceedings of the IEEE - Special Issue on Multi-Robot Systems*, Vol. 94, No. 7, July, 2006, pp. 1425 - 1444.
- [12] H. A. Yanco, J. L. Drury, and J. Scholtz, "Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition," *J. Human-Comput. Interaction*, vol. 19, pp. 117-149, 2004.
- [13] B.J. Zimmerman and D. H. Schunk. (Eds.). *Self-regulated Learning and Academic Achievement: Theoretical Perspectives* (2nd ed.). Mahwah, NJ: Erlbaum. 2001.