Improving Ad-Hoc Team Performance Using Video Games

by

Jeffrey David Craighead

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy Department of Computer Science and Engineering College of Engineering University of South Florida

Co-Major Professor: Robin Murphy, Ph.D.
Co-Major Professor: Lawrence Hall, Ph.D.
Jennifer Burke, Ph.D.
William Kearns, Ph.D.
Dewey Rundus, Ph.D.

Date of Approval: July 13, 2009

Keywords: Robotics, Simulation, Training, Search and Rescue, Communication

© Copyright 2009, Jeffrey David Craighead
Dedication

This work is dedicated to my wife Jamie and our son Alex.
Acknowledgements

I would like to thank the Clint Arnett, Brian Smith, Daniel May, and Jeff Bolich from Disaster City and Kimberly Mallett from Texas A&M University for their assistance in arranging and conducting my experiments. Without their help I would not have finished on time.
Table of Contents

List of Tables iii
List of Figures iv
Abstract vi

Chapter 1 Introduction 1
  1.1 Motivating Examples 5
  1.2 Dissertation Organization 7

Chapter 2 Related Work 8
  2.1 Robot Games and Simulators 10
  2.2 The Unity Engine 20
  2.3 Games for Education 24
  2.4 Video Game Score Keeping 27
  2.5 Training 29
  2.6 Fractal Path Analysis 32
  2.7 Chapter Summary 35

Chapter 3 Approach 36

Chapter 4 SARGE Implementation 38
  4.1 Inuktun Extreme VGTV 39
  4.2 Real-Time Fractal Path Analysis 41
  4.3 Environments 43
  4.4 Data Logging 46
  4.5 Chapter Summary 47

Chapter 5 Experiments 48
  5.1 Part One 50
  5.2 Part Two 53
  5.3 Analysis 57
  5.4 Demographics 58
  5.5 Dependent Measures 73
  5.6 Chapter Summary 78
<table>
<thead>
<tr>
<th>Chapter 6</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Correlations</td>
</tr>
<tr>
<td>6.2</td>
<td>Team Training Has No Effect on Field Score</td>
</tr>
<tr>
<td>6.3</td>
<td>Fractal D Has a Non-Linear Relationship with Search Score</td>
</tr>
<tr>
<td>6.4</td>
<td>A Driver’s Prior Gaming Experience Increases Fractal Dimension</td>
</tr>
<tr>
<td>6.5</td>
<td>A Driver’s Prior Gaming Experience Improves Field Score</td>
</tr>
<tr>
<td>6.6</td>
<td>A Driver’s Prior Gaming Experience Improves Search Score</td>
</tr>
<tr>
<td>6.7</td>
<td>A Driver’s Increasing Age Decreases Fractal Dimension</td>
</tr>
<tr>
<td>6.8</td>
<td>A Driver’s Age Has a Negative Relationship with Search Score</td>
</tr>
<tr>
<td>6.9</td>
<td>A Driver’s Age Has No Effect on Field Score</td>
</tr>
<tr>
<td>6.10</td>
<td>Observation: Team Training Increases Intra-Team Communication</td>
</tr>
<tr>
<td>6.11</td>
<td>Discussion of Findings</td>
</tr>
<tr>
<td>6.12</td>
<td>Changes for the Future</td>
</tr>
<tr>
<td>6.13</td>
<td>Chapter Summary</td>
</tr>
</tbody>
</table>

| Chapter 7 | Conclusions and Future Work |

List of References

Appendices

- Appendix A: Demographic Survey
- Appendix B: IRB Approval Letter

About the Author
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Participant Gender</td>
<td>59</td>
</tr>
<tr>
<td>Table 2</td>
<td>Participant Age</td>
<td>60</td>
</tr>
<tr>
<td>Table 3</td>
<td>Participant Occupation</td>
<td>61</td>
</tr>
<tr>
<td>Table 4</td>
<td>Participant Education</td>
<td>63</td>
</tr>
<tr>
<td>Table 5</td>
<td>Participant Technology Experience</td>
<td>65</td>
</tr>
<tr>
<td>Table 6</td>
<td>Differences in Demographic Means by Team Training Group</td>
<td>72</td>
</tr>
<tr>
<td>Table 7</td>
<td>Descriptive Statistics of Dependent Measures</td>
<td>74</td>
</tr>
<tr>
<td>Table 8</td>
<td>Correlations with the Dependent Measures</td>
<td>83</td>
</tr>
<tr>
<td>Table 9</td>
<td>Inter-item Correlations from the Survey</td>
<td>86</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1  The Unity IDE  23
Figure 2  Fractal Dimension Example  34
Figure 3  Simulated Inuktun VGTV  40
Figure 4  Slalom Course  44
Figure 5  Step Course  44
Figure 6  Search Course  45
Figure 7  Disaster City Collapse Layout  56
Figure 8  Participant Gender  59
Figure 9  Participant Age  60
Figure 10  Participant Occupation  62
Figure 11  Participant Education  64
Figure 12  Experience Level with R/C vehicles  66
Figure 13  Experience Level with Robots  67
Figure 14  Experience Level with Video Games  68
Figure 15  Experience Level with Video Cameras  69
Figure 16  Experience Level with Search Equipment  70
Figure 17  Experience Level with Computers  71
Figure 18  Search Score Histogram  74
Figure 19  Mean Fractal Dimension Histogram  75
Figure 20  Field Score Histogram  76
Figure 21  Communication Frequency Histogram  77
Figure 22  Mean Field Score vs. Team Training  88
Figure 23  Mean Fractal Dimension vs. Training Search Score  91
Figure 24  Mean Fractal Dimension vs. Prior Gaming Experience  93
Figure 25  Prior Gaming Experience vs. Field Score  95
Figure 26  Prior Gaming Experience vs. In-Game Search Score  97
Figure 27  Median Age vs. Gaming Experience Level  98
Figure 28  Mean Fractal Dimension vs. Age  100
Figure 29  Effect of Driver Age on Search Score  103
Figure 30  Effect of Driver Age on Field Score  105
Figure 31  Communication Level vs. Team Training  107
Improving Ad-Hoc Team Performance Using Video Games

Jeffrey David Craighead

ABSTRACT

This dissertation examined the effects of distributed, multiplayer training video games on the performance of distributed teams of robot operators. Two hypotheses were tested, the first hypothesis stated that online, game-based team training will improve the performance of an ad-hoc team versus an ad-hoc team formed of individually trained teammates. The second hypothesis stated that the fractal dimension of a robot’s path can be used as an indicator of its operator’s skill. Forty-one volunteers participated in an experiment in which they played a distributed, online training game which showed them the basics of operating an Inuktun Extreme VGTV for a search task. The participants were divided into two groups, one group trained in pairs as a team while the other group trained individually. The results showed that team training has no effect on the number of items found in a search by an ad-hoc team; however, team training does significantly impact the amount of information sharing between team members. The results also showed that the fractal dimension of a robot’s path is quadratically related to the operator’s effectiveness in a search task. Additionally, a participant’s age and prior video game experience are related to their score obtained in a search task using a robot.
Chapter 1

Introduction

This dissertation investigates the effect of training teams of robot operators in a
distributed gaming environment by comparing single operator training versus distributed,
multiple-operator training for two-person, ad-hoc teams. The first research question is:

Does the performance of distributed, ad-hoc teams of robot operators improve if the team
members have previously participated in online, multiplayer robot operator training?

This type of training is similar to logging on to a multiplayer, team-based game
periodically to refine one's skills as a robot operator. The novelty of this dissertation is the
examination of the research question using a distributed, game-based robot simulation
game, leveraging the successful work of Cromby [Cromby 1996] and Rose [Rose 2000]
in transfer of training in virtual worlds. Additionally this dissertation will test a second
research question that involves the use of a new real-time fractal path analysis tool. The
second question is: Does the fractal dimension of a tele-operated robot’s path have a
relationship with the skill of it’s operator when conducting a search task? Fractal path
analysis has previously been used to examine the migratory behavior of animals [Nams
The lack of familiarity and training with robots among US&R teams has been well documented by Burke, Murphy, and others [Burke 2003, 2004, Murphy 2004]. More recent work such as that by Nourbakhsh, Lewis, and others [Nourbakhsh 2005] has focused on improving human-robot team performance by attempting to make the robots navigate autonomously and provide more information to the user. However, according to Burke, a robot operator is already overwhelmed by the large amount of information on which they must act and with the unfamiliar viewpoint that they must take when operating a remote vehicle. Her work suggests that it is necessary to add more humans to a team to simply monitor the sensor data for important items and events; or else limit the information bombarding the operator, through computer mediation, to that which is immediately relevant in order to reduce the cognitive load on the robot operator and improve situational awareness. Burke cites Endsley's [Endsley 1988] three level model for situational awareness.

“Perception (Level 1) is detection of sensory information: the perception of elements in the environment within a volume of time and space. Comprehension (Level 2) is divided into two subcategories: identification and interpretation. Identification is defined as comprehension of perceived cues in terms of subjective meaning (eg., identifying objects, locations, and victims). Interpretation is defined as comprehension of perceived cues in terms of objective significance or importance to the situation. Projection (Level 3) is defined as the perceived cues in terms of objective significance or importance to this situation.” [Burke 2004b]
A tool that allows operators to develop skills that decrease the time to attain situational awareness while maintaining active control of the robot could be useful. As discussed in several articles [Green 2003, Bekebrede 2005, Fabricatore 2000, Aitkin 2004, Mayo 2007, Chatham 2007] games have the potential to enhance learning and improve reflex actions when they include skill tests appropriate for the given task. Additionally when these games are assigned as homework to complement traditional classroom learning the potential for improvement is even greater. As part of this dissertation a robot operator training game & simulation environment called the Search and Rescue Game Environment (SARGE) [Craighead 2008a, Craighead 2008b] was developed to answer the research questions.

The experiment consists of two groups of two-person teams that varied on type of training received. Group 1 consisted of teams of randomly paired participants that received individual operator training in the video game. Group 2 consisted of teams of randomly paired participants that received the same individual operator training as Group 1 and also received a multiple operator training scenario with a random partner. These two groups participated in a field trial in which they used a real robot to conduct a search task. During the field trial, the participants in group Group 2 were repaired and stratified on the role of the participant (driver or observer) during team training to eliminate a bias towards learning a role instead of learning to work with a teammate. This allows us to determine if team training with a random partner is beneficial to distributed, ad-hoc teams in the real world. The first hypothesis for this experiment is that ad-hoc teams consisting
of operators who receive team training will perform better in the real world than ad-hoc
teams consisting of operators who have only received individual operator training due to
their experience working in a distributed, ad-hoc team. Performance in this dissertation
refers to the number of victims or victim related objects a team identifies. Additionally
this dissertation hypothesizes that team performance will have either a positive [Voshell
2007] or negative [Clarke 2007] linear relationship with the fractal dimension of path of
the robot, this will be tested using the real-time fractal path analysis tool.
1.1 Motivating Examples

There are several motivating cases for distributed robot operations. One motivation for this dissertation is the operation of a reach-back device called Survivor Buddy. The Survivor Buddy is a robot mounted, wireless device with two way audio and video communication capabilities and may include other sensors to monitor a victim's health. The Survivor Buddy camera and display are mounted on a pan/tilt unit which is under the control of the operator. The device is intended to provide a means of communication with responders, off site medical teams, and potentially the victims' family. An off site medical team might be interested in scanning the victim for injuries using the onboard camera and other sensors integrated into the device as well as talking with the victim; on the other hand a victim's family might only be interested in talking with the victim for comfort. Given these proposed uses, the operator of the Survivor Buddy device will not be the operator of the robot on which it is mounted, thus the robot operator and Survivor Buddy operator must coordinate such that the robot position facilitates optimal placement of the Survivor Buddy for the given task (communication or medical assessment using sensors). Providing remote users experience operating the robot and Survivor Buddy as a team in a simulated environment may improve the coordination and overall performance of a real world ad-hoc team of which they are a member.
A second motivation is Remote Shared Visual Presence (RSVP) [Burke 2006] via a data network. A remote shared visual presence system allows one or more remote parties to view the live stream of data from a robot and provide feedback on a task to the robot operator. One can see that this is the precursor to the use of a Survivor Buddy. A robot, outfitted with a Survivor Buddy could be sent into a disaster site under the control of one on-site operator. The operator is in communication with one or more search specialists or structural engineers that are off site, forming an ad-hoc team. The off site team members receive the same video feed that is available to the robot operator. These team members provide feedback to the operator and help locate critical information about the area and possible victim locations. If a victim is located the team configuration changes as the search specialists and engineers are replaced with a doctor or a family member.

These two examples are limited to the search and rescue domain using a small ground robot; however, the use of RSVP and shared control for ad-hoc teams of operators of robots or robot mounted equipment applies to other domains as well such as UAV and USV operations for surveillance and scientific exploration.
1.2 Dissertation Organization

This dissertation is organized into nine chapters. Chapter 1 provided an overview of the research question, hypotheses and motivation for proposed work. Chapter 2 provides a detailed discussion of the related work in the areas of robot simulation, training, and fractal-path analysis. Chapter 3 discusses the approach this dissertation will use to answer the research question. Chapter 4 explains the additions made to the SARGE simulation environment for this dissertation. Chapter 5 details the experiments conducted as well as the demographics of the participants. Chapter 6 details the significant results of the study and discusses the significance of these findings, then provides suggestions for making the experiment more effective. Finally, Chapter 7 provides a summary of the findings and the contributions of this work and discusses several future projects.
Chapter 2

Related Work

In the last few years there has been a renewed interest in robot simulators. There are many commercial and open source simulators available such as Webots, USARSim, SimRobot, Player/Stage/Gazebo, Microsoft's Robotics Studio and many others. Each of these simulators is designed to serve a similar purpose: to provide a virtual version of real robots and environments to aid in development of control algorithms for both research and educational purposes. There have been attempts to validate USARSim's physics simulation fidelity as well as verify its capabilities for HRI research [Wang 2005, Carpin 2006, Carpin 2007]. Wang even stresses the advantages of USARSim over the others for HRI because of its high fidelity physics and graphics. However, USARSim was originally built as a visualization and simulation environment, not as a user friendly application, which is apparent from the non-trivial installation and execution procedures.

There has been an interest in using computer games for training and HRI research. Richer and Drury [Richer 2006] examined the usefulness of various video game features for HRI research. Fabricatore [Fabricatore 2000], Pivec [Pivec 2003], Chamberlin [Chamberlin 2003], Atkin [Aitkin 2004] and others examine the potential of video games
for learning. The general consensus is that video games have great potential for use in
education and training beyond the simple reading, writing, and mathematics of the past.
Not only can games be used as an entertaining teaching tool for young students, but for
adults as well in a variety of fields.

Each section below highlights the important, or most similar works in each of the
areas discussed above. Section 2.1 discusses the general capabilities of games and
simulators that have been used for robotics research. Section 2.2 discusses the advantages
of the Unity engine, the engine of choice for developing SARGE, over the other
simulators and engines reviewed. Section 2.3 discusses several recent works with 3D
environments used for education. Section 2.4 discusses recent techniques used in video
game scoring and match making. Section 2.5 discusses works related to training in virtual
environments. Section 2.6 discusses how fractal path analysis works and how it has been
used in the past. Section 2.7 provides a summary of this chapter.
2.1 Robot Games and Simulators

Previous work has shown that simulation in virtual environments can train aircraft pilots, which is the most commonly known form of professional grade simulation. The US Navy has used Microsoft's Flight Simulator application and the X-Plane flight simulator has received certification by the US Federal Aviation Administration (FAA) for training general aviation pilots. Both applications are commercial products available to any consumer for less than $100. Other work such as that by Stottler, et al. [Stottler 2002, Stottler 2000] have applied simulation based training to other vehicles and systems such as armored ground vehicles and a Naval tactical action station. However these tasks were all first-person in nature, that is the pilot was actually in the vehicle. The operation of remote vehicles is more difficult as the pilot must mentally project themselves into the position of the vehicle to maintain situational awareness. There has been little published work in the use of simulation for training unmanned vehicle operators with the exception of unmanned underwater vehicles (UVV or ROV). This exception is most likely due to the fact that UUVs have been in widespread use by Naval and research groups for several decades [Pioch 1997, Seet 2001]. More recent work has focused on increasing the fidelity of simulations of autonomous unmanned vehicles [Carpin 2007]. These works have largely ignored the need for operator training. An additional challenge with newer vehicles is the fact that many of these vehicles are or will often be operated by two or more physically separated individuals connected by high speed networks. This poses challenges to team communication and coordination that were not issues for the UUV.
work. This section presents a review of the literature related to robot simulation as well as available robot simulators and simulation engines.

Alexander, et al. in "From Gaming to Training: A Review of Studies on Fidelity, Immersion, Presence, and Buy-in and Their Effects on Transfer in PC-Based Simulations and Games" [Alexander 2005] argue that commercial games and game engine based simulations have the potential to provide an environment that is high-fidelity representation of reality. Nielsen and Goodrich in "Comparing the Usefulness of Video and Map Information in Navigation Tasks" [Nielsen 2006] used the Unreal2 game engine with the USARSim modification to examine how video and maps affect human interaction with a robot while navigating the robot through an environment. Stephen Hughes and Michael Lewis in "Robotic Camera Control for Remote Exploration" [Hughes 2004] use the Unreal2 game engine with USARSim to study the effects of camera placement on the human control of robot mounted cameras.

The remainder of this section highlights the important features and uses of the currently available robot simulators and their shortcomings. Section 2.2 provides a discussion of why the Unity environment was chosen over the others for the development of SARGE.

eyeWyre Studio is a development environment and simulator for BASIC Stamp 2 micro-controller based robots. The eyeWyre provides physics simulation in small
environments. The environments and robots are limited to those provided with the package, severely limiting its usefulness in a research environment requiring custom robots. As of this writing eyeWyre Studio is no longer available.

FlightGear is a open source simulator that uses by default a blade element analysis, similar to X-Plane. Global scenery is available for FlightGear. Aircraft models must be created in an external 3D modeling application and an XML file describing the various aircraft features must be created by hand. FlightGear has been used for various academic projects. For example, Summers, et al. in [Summers 2002] used FlightGear to simulate a UAV carrying environmental sensors and Cervin, et al. in [Cervin 2004] used FlightGear to create an interface for a real UAV. Kurnaz, et. al used FlightGear to develop a fuzzy logic based control algorithm for a UAV [Kurnaz 2007]. FlightGear is available as a free download under a GPL license. The entire source code is available for modification and is under constant development. The application runs on Windows, Mac, and Linux operating systems.

MATLAB is a numerical simulation environment that supports visualization via a Virtual Reality toolkit. Toolboxes are available that provide quick access for building various robot controllers based on fuzzy logic, neural networks, and genetic algorithms. MATLAB can communicate via network with other simulation environments that may provide better physics simulation and visualization for rapid controller prototyping. The MATLAB, Simulink, and the VR Toolbox set runs on Windows, Mac, and Linux and is

Microsoft Flight Simulator provides detailed visuals for the aircraft and environment. It uses a lookup table driven flight model for aircraft simulation which is less accurate than the blade element analysis approach used by X-Plane; however, Flight Simulator has been used by the US Navy as a training aide for pilots. Microsoft provides an SDK as a download for Flight Simulator which provides access to simulator data via a network, weather, terrain, scenery, instrumentation, and aircraft creation. Aircraft must be created in a 3rd party 3D modeling application such as GMax or Lightwave. Flight Simulator is often used for visualization purposes in combination with MATLAB for UAV controller development as in [Bayraktar 2004] and as an all-in-one simulation environment as in [Feldstein 2004]. As of February 2009, development of Flight Simulator has been discontinued by Microsoft.

Microsoft's Robotics Studio attempts to be a complete end-to-end solution for robot control software development. It provides a distributed service based architecture that can run on Windows PC and several microcontroller based platforms and a
simulation environment that uses the PhysX physics engine. Robotics Studio currently does not support user created content. Robotics Studio is under active development and most recently Microsoft has sponsored a 2008 Robotics FIRST competition in Seattle, Washington to promote Robotics Studio. Additionally the preview/beta version of Robotics Studio 1.5 was released in May 2007. Moreno, et al. [Moreno 2007] integrated a cognitive agent into the Robotics Studio services for deliberative control of a simulated robot.

MissionLab is described in ``Behavior-Based Formation Control for Multirobot Teams'' [Balch 1998], Balch and Arkin, presents the use of reactive behaviors for formation control in simulation and on real robots using two architectures, AuRA and the UGV Demo II architecture. The current version of MissionLab uses a distributed architecture, allowing various pieces of the simulation to run on different machines. This also allows the same interface to be used to control real robots. Visualization is provided by both 2D and low fidelity 3D OpenGL displays. The current version of MissionLab supports all examined vehicle types: UGV, UAV, UUV, and USV. The 3D display can render terrain generated from a height map along with a low polygon count model of robotic vehicles. The 2D displays can render maps and image overlays as well as custom graphs. MissionLab does not appear to support any physics simulation which would be necessary for vehicle simulation. Extensive documentation is provided, but C/C++ is the only language supported.
The Player/Stage/Gazebo project is an open source project that provides a 2D and 3D environment for robot testing. Stage and Gazebo are networkable simulation environments, Player defines an interface for robots and sensors to communicate with Stage and Gazebo. Stage is a simple 2D environment that provides basic collision detection and range sensor modeling. Gazebo is a 3D environment that brings the basic simulations of Stage into the 3rd dimension. Gazebo provides a camera sensor as well as the ability to use complex objects in the environment. Gazebo presents a simple low fidelity OpenGL based visualization of the environment. While Stage does not support physics simulation, Gazebo can make use of the Open Dynamics Engine (ODE) physics engine. Recent research with Player/Stage/Gazebo has used it almost exclusively for simulation of navigation and control algorithms. Karimian, et al. extended the simulator to examine the use of audio for inter-robot communication [Karimian 2006]. Rusu, et al. extended the Player/Stage/Gazebo architecture and simulation environment to support wireless sensor networks, manipulators, and 3D laser sensors [Rusu 2007]. Kim, et al. used Player/Stage to compare a non-random cleaning algorithm for a vacuuming robot to the random algorithm used by the Roomba [Kim 2007]. Skubic, et al. used the Player/Stage environment to examine the use of hand drawn sketches for controlling a team of robots using a Tablet PC [Skubic 2007].

Simbad [Hugues 2006] is an open source Java based 3D robot visualization environment. It does not support any physics calculations, only simple collision detection with objects placed in a flat world. The goal of this simulator is to provide a simple
environment to test robot controllers and AI algorithms, the support for high fidelity visualization is not present. The standard sensors are sonar, camera, light, and bump sensors. Robots are represented as simple geometric primitives. As an open source project the ability to add new sensors is present in this simulator. Simbad will run on any operating system with a Java client with the Java3D library. Simbad has been used by Hartland, et al. to explore the effectiveness of robotic controllers which are evolved in simulation and then transferred to a real robot [Hartland 2006].

SimRobot, described in "SimRobot - A General Physical Robot Simulator and its Application in RoboCup" [Laue 2006], is a physics based robot simulator with a 3D OpenGL based display. SimRobot uses the Open Dynamics Engine for physics calculations which gives it an edge over many custom simulators. The use of a custom OpenGL visualization environment however could be improved on by using a preexisting rendering engine like OpenSceneGraph. Robots and environments are specified using XML by specifying part types and positions. Several sensor types are supported, including cameras, range sensors, touch sensors, and actuator state. SimRobot was used by the German team for the 2005 RoboCup competition; however, it is not limited to RoboCup robots or environments. The Laue paper shows an office environment simulated in the SimRobot simulator.

USARSim is an open source urban search and rescue robot simulator USARSim, based on the Unreal2 engine, is primarily aimed at ground vehicles. The engine best
supports bipedal and wheeled robots; however, it is possible to add support for other robot types. It should be noted that the Karma physics engine used in the Unreal2 engine provides only basic simulation of forces on specific objects within the environment.

Robot parts and environment objects are created in 3rd party modeling applications. To create a complete robot the user must create a text file by hand which specifies the size and position of each part. Worlds can be created using an included utility. Robots can be programmed using UnrealScript or controlled over a network connection using USARSim's UDP control protocol. USARSim is used for the RoboCup Rescue competition's simulation league. The Unreal2 game engine is one of the dominant commercial simulator platform for robotics simulation for unmanned ground vehicles. It has been used to simulate robots, train army recruits and firefighters, as well as conduct studies on search and rescue tasks [Schneider 2005, Nielsen 2006, Wang 2003, Phongsak 2004].

Webots PRO [Michel 2004, Hohl 2006] is a ground robot simulator that uses the open source Open Dynamics Engine for its physics simulations and an extended VRML97 based environment. Webots provides several small built-in robots such as the Khepera, Pioneer2, and Aibo as well as the means to import custom robots from 3rd party modeling applications using the VRML97 format. World size is defined by the user and can be as large as needed. Webots PRO supports various sensor types such as camera, range finder, GPS, light sensors, etc; as well as effectors like grippers, limbs, and wheels. WebotsPRO can compile controllers created within the simulator to work on real robots.
given that the hardware is supported by the compiler. Hohl in demonstrates the remote control of and controller transfer to an Aibo robot through Webots.

X-Plane is a commercially available flight simulator developed by Laminar Research. X-Plane has received FAA certification as a training simulator when used with certain hardware configurations because of its high fidelity simulation of flight model and visualization. X-Plane uses blade element analysis to drive its flight model. Included with the package are the simulator, global scenery generated using data from NASA's terrain mapping radar missions, an airfoil designer, and a plane maker application. X-Plane has been used for testing and pilot training for the Carter Copter and SpaceShipOne experimental vehicles. Previous work has successfully used X-Plane to test a micro UAV controller developed in MATLAB [Ernst 2007].

robotSim is the visualization portion of Cogmation's robotSuite. The robotSuite enables visual coding of selected robots and sensors. The robotSim is created using the Unity game engine and provides several environments in which the robots can be run. The robotSim is limited to being used with the robotSuite.

This section presented an overview of recent work using robot simulators. All of the work is focused on vehicle simulation, not training operators. Additionally none of the works incorporate any type of game or training features. These works are strictly testing algorithms for navigation, planning, and object identification. Additionally the
majority of the environments are not suitable for game-based tutoring of robot control and those which may be suitable are inappropriate for non-engineers. The dearth of options shows there is a need for a new easy to use game-based application for tutoring robot operators.
2.2 The Unity Engine

The Unity game engine is developed by Unity Technologies. Unity integrates a custom rendering engine with the nVidia PhysX physics engine and Mono, the open source implementation of Microsoft's .NET libraries. The benefits of using Unity are numerous when compared to the engines and simulators discussed in Section 2.1. This section provides a breakdown of what we consider to be the key features of Unity that make it an excellent robot simulation engine.

- **Documentation** - The Unity engine comes with complete documentation with examples for its entire API. This is the biggest benefit of Unity and leads to increased productivity when compared to other engines such as Unreal or Source which only provide partial documentation for non-paying customers (mod developers).

- **Developer Community** - There is an active on-line developer community which can often provide assistance for new users. The Unity Technologies developers also are very willing to add features to the engine upon a user’s request, which seldom occurs when using free versions of popular commercial code. Several of the features existing in the Unity API are a result of the author’s requests during the development of the SARGE application.
Component Oriented Programming - Unity's editor is by far the easiest to use when compared to Unreal, Source, or Torque. Content is listed in a tree and is added to an environment in a drag-n-drop manner. Objects in the environment are listed in a separate tree, each of which can be assigned multiple scripts written in either C#, Javascript, or Boo (a dialect of the Python language for Mono) as well as physics and rendering properties. Script developers have access to the complete Mono API. Scripts can give objects interactive behaviors, create user interfaces, or simply manage information. Figure 1 shows a screenshot of the Unity Editor being used to develop a user interface for the iRobot Packbot.

Physics and Rendering - By using physics properties, objects can be given mass, drag, springiness, bounciness, and collision detection as well as be assembled using a variety of joints. The physics properties are simulated by nVidia's PhysX engine, which is used in many AAA commercial games. The rendering properties include shader and texture assignment which affect the appearance of visible objects. Unity's custom rendering engine uses a simplified shader language which is compiled into DirectX 9 or OpenGL 2.0 shaders depending on the target platform.

Multiplatform Distribution - The Unity engine's editor runs on OSX and Windows, applications created using Unity can be compiled for OSX,
Windows, or as a Web-Player (which runs in a web browser via a plugin, similar to Adobe Flash). Additionally the Windows builds can run on Linux through the WINE emulator. There are no restrictions on distribution of applications created with Unity and because applications created with Unity are not mods of existing games the end user does not need to own a copy of anything. Complete binaries can simply be distributed as the developer wishes.

- Low Cost - The Unity engine has a relatively low cost for a complete game engine (although more expensive than the free open source engines). The Indie version of the engine is US $199 while the Pro version, which is required for real-time shadowing and other similar effects, is US $750 for an Academic license or US $1499 otherwise. This pricing is comparable to Torque, yet Unity's Editor is in the author’s opinion much easier to use. While this is more expensive than modifying an existing game, which usually only requires the purchase of a game, it provides far more developmental freedom. If one were to try to license the Unreal or Source engines for use in a small project, the cost would be prohibitive. DevMaster.net estimates the cost for a source code license for the Unreal 3 engine to be over US $700,000.
Figure 1. The Unity IDE.
2.3 Games for Education

Games have long been used for education and computers have provided an opportunity for simple classroom games and experiments to be replaced or augmented using virtual environments. These computer based games allow the instructor to present material in new ways and give the students the ability to experiment in the virtual environment in a way not possible with traditional classroom instruction. The instructor can monitor individual students and provide instant feedback on their performance.

Cooper, Dann, and Pausch [Cooper2003, Cooper2003a] use the Alice environment to teach novice computer science students the basics of object-oriented programming. Alice presents the user with a 3D world in which they can place objects. Each object can have one or more scripts associated with it which provide a means of interacting with other objects in the virtual world. The scripts are created through a drag-n-drop interface which allows the students to learn the concepts without worrying about language syntax. On the downside they note that the students do not gain experience creating code using proper syntax in this manner. Cooper et al’s results show that Alice users earned a higher graded than the control group as well as the group of all students in an introductory computer science class.

Schell, et.al. of Carnegie Mellon University's Entertainment Technology Center created the multiplayer Hazmat:Hotzone as a training tool for firefighters.
Hazmat:Hotzone allows an instructor to set up a hazmat scenario, such as a chlorine leak, in one of several virtual worlds. A team of firefighters will play out the scenario many times exploring various methods of solving its challenges as part of classroom based instruction. Hazmat:Hotzone has been commercialized at Code3D from SimOps Studios. There has been no published work quantifying the effectiveness of using Code3D.

The Institute for Creative Technologies (ICT) at the University of Southern California under the supervision of James Korris designed the Xbox title Full Spectrum Warrior (FSW). The development of FSW was contracted out to an existing Xbox developer, Pandemic Studios, with the requirement that the game must be publishable to the general public, but have special Army features protected by an unlock code. FSW is used to give fire squad members an idea of the tactics used by the squad leader. ICT presumed that “By taking the “boss's” job, soldiers might deepen their appreciation for the correct execution of dismounted battle drills in the urban context.” [Korris 2004]. This is termed cross-training in the psychology literature [Agnihothri 2004].

This section shows the few games without tutoring systems that have been used for adult education in various fields. None of these focus on robot operation; however, two are closer to the desired goal. Code3D and Full Spectrum Warrior are games that take place in complex environments. Code3D is multiplayer, but runs in a networked classroom under the supervision of an instructor. The special Army version of Full Spectrum Warrior is also used in a classroom with an instructor but is single player only.
This shows that while training in complex environments is possible, there has been no published work investigating the effects of distributed, multiplayer training on teams.
2.4 Video Game Score Keeping

Video games have almost always included some scoring mechanism with which players could compare scores. Players would attempt to beat their own scores as well as the scores of others. There is no standard scorekeeping method used in the video games industry. Each game has a unique design that influences how points are gained. Scoring can be generalized to a reward for competing one or more goals. Goals in video games are typically one of the following three: find an object, destroy an object or reach some point in the environment as fast as possible. There has been no academic research into scorekeeping in video games. What has been the topic of research is player ranking. Player ranking is the process of comparing the abilities of players for a given game. The most naive method is simply to compare scores and assume that the players with higher average scores are better than those with lower average scores. In general this naive method works well, but a problem arises when attempting to match random players in multiplayer games. Novice players do not usually want to play with or against an expert player because the expert will easily best the novice every time. This type of mismatched pairing also causes the expert to lose interest because the game is not challenging. One of the first more complex ranking systems, Elo, was created to rank world chess players. The Elo system (named for its creator Arpad Elo) is a probabilistic method for comparing player skill, based on the players' performance in past games the probability that one player will win versus another is computed. The Elo system is briefly described in [Herbrich 2006] which introduced a new Bayesian method for ranking players, called
TrueSkill, which works for players of team-based games in addition to two player paired games. TrueSkill was developed by Microsoft Research for the Xbox Live! system to rank players of online games. Player ranking and pairing is outside the scope of this work as during the experiment, players will not play long enough to build up a skill ranking. However it is possible that an implementation of TrueSkill system may be integrated with a later version of SARGE.
Maximal transfer of training is the goal of any training program. Transfer of training is the ability of the trainee to generalize the knowledge gained in the training environment to their work environment. Considerable research has been done investigating how transfer is best achieved. The studies typically involve office or manufacturing environments. Two reviews of training literature which cover many of the significant works of the 20th century [Baldwin 1988, Cheng 2001]. Cheng et al. notes that “Training has been regarded as an expensive investment.” This statement applies doubly to the military, law enforcement, and robotics communities. Not only does training incur costs for paid man-hours, training using expensive robots in potentially hazardous environments can lead to tens to hundreds-of-thousands of dollars worth of damage to the robots. Thus a training system that minimizes these costs while maximizing transfer is vital.

Cheng [Cheng 2001] proposed a conceptual framework for measuring transfer of training based on Kirkpatrick’s often cited article on evaluating training [Kirkpatrick 1996]. The framework lists four stages of the transfer process:

- Pretraining motivation - refers to the intended effort towards mastering the content of a training program
- Learning - the process of mastering the content of a training program
• Training performance - the measurement of the extent of what a trainee has achieved in a training context.

• Transfer outcomes - those attainments made by the trainees when they apply what they have acquired in a training context back to the job, which can benefit both trainees and the organization.

This dissertation will focus on steps three and four, that is measuring training performance and transfer outcomes. The Cheng article also identifies nine additional factors which have been identified to affect transfer and suggests that some or all these factors should be measured along with performance evaluation by observation and surveys. The factors identified are:

• Locus of control
• Self-efficacy
• Career/job attitudes
• Organizational commitment
• Decision/reaction to training
• Post-training interventions
• Supports in organization
• Continuous-learning culture
• Task constraints
These works have focused on real world training, typically for corporate and management jobs. More recent works have focused on identifying transfer from training in virtual environments to the job. Rose et al. [Rose 2000] show that there are no significant differences between training for a simple sensorimotor task in the real world or a virtual environment. The task involved moving a metal ring around a curved rod without making contact, the number of contacts were counted as errors. Training in either the real world or the virtual environment significantly reduced the number of errors. This obviously is a contrived problem; however, on a basic level the sensorimotor actions may be similar to that needed when tele-operating a robot.

Cromby et al. [Cromby 1996] conducted an experiment in which mentally handicapped students were trained in a supermarket shopping task using a virtual environment. The students were split into two groups. Both groups participated in a baseline experiment. One group then received training in a virtual environment that was designed to look like two aisles of a supermarket while the control group played with other virtual environments. For the second experiment both groups returned to the supermarket, which had rearranged for holiday sales. The group trained in the supermarket virtual environment performed significantly better than the control group; however, neither group performed better than own their baseline due to the increased difficulty of the followup task. This study shows the potential benefit of training in virtual environments, by training in the virtual environment the experimental group was able to complete a more complex task significantly faster than the untrained group.
2.6 Fractal Path Analysis

The term fractal as it appears in this work is defined as an abbreviation of “fractional dimension”, coined by Benoit Mandelbrot in his 1967 article ``How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension” [Mandelbrot 1967]. A fractional dimension is used to describe the variation in a surface or line when examined at varying magnification levels. A seemingly smooth surface may be very rough when magnified or a rough surface may also appear similarly rough on a smaller scale (self similarity). For instance, Mandlebrot's article posits that the coastline of Britain will be significantly longer if measured in millimeters versus kilometers due to the self-similar nature of coastline. This intuitively makes sense, by using millimeters as the unit of measure significantly more of the variation in the surface can be included in the measurement. The fractal dimension is computed by comparing the measurement of a line or surface at multiple scales. A truly straight line has a fractal dimension of 1, while a line that crosses a plane enough times to completely cover the surface (i.e. follows a Brownian motion) has a fractal dimension of 2, thus when describing the fractal dimension D of a line $1 \leq D \leq 2$. Fractal dimensions can also be applied to the three dimensional analysis of surfaces such that the fractal dimension $2 \leq D \leq 3$. Figure 2 shows an example of several lines and the associated fractal dimensions.

Fractal dimension has been used recently to study the behavior of animals [Nams 2006]. By classifying their path tortuosity (how convoluted the animal's movement is
over time) with a fractal dimension it becomes easier to compare the territorial behaviors of different species and individual animals. Voshell, Phillips, and Woods [Voshell 2005, Voshell 2007] used path tortuosity measured by fractal dimension to compare the effectiveness of two robot user interfaces. In their study, operators with better situational awareness smoothly navigated the robot through the environment resulting in lower fractal dimension values.

This dissertation will implement a real-time fractal path analysis (RTFPA) algorithm in an attempt to determine the relationship of fractal dimension to an operator’s skill level. It is hypothesized that a more skilled operator will have higher situational awareness and a better understanding of how to navigate the robot in a given environment. If successful, the skill level determined using the RTFPA could conceivably be used in the future as input to an intelligent tutoring system to adjust the complexity and difficulty of the training environment. This will be the first real-time implementation of a fractal path analysis algorithm.
Figure 2. Fractal Dimension Example. This figure shows three paths a robot could take and the associated fractal dimensions. A perfectly straight line has a fractal dimension (D) equal to 1.0. As the tortuosity of a line increases its fractal dimension approaches 2.0.
2.7 Chapter Summary

This chapter has presented a review of related literature in several fields. The beginning of the chapter reviewed many available robot simulators and simulation engines and discussed why the Unity engine was chosen to implement the simulator for this dissertation. The main reasons for the choice were usability and simulation fidelity. Unity provides simple well documented tools which allow the creation of a robot simulator with higher fidelity than the alternatives and in a very short period of time. The preceeding sections discussed training and educational games and training theories. Published work on games which have been used to teach complex skills from computer programming to firefighting was presented and demonstrated that while teaching complex tasks to individuals with games is an active area of research, examining the effects of distributed, game-based training on teams is not. Finally this chapter presented a short discussion of the prior use of fractal path analysis in animal research and robotics research.
Chapter 3

Approach

The approach taken in this work is to extend the theories and results from Rose, Cromby, Suebnukarn and Nams into a new environment with more complex tasks. As mentioned in Chapter 2 Rose, et al. showed that there is no difference in training for a sensorimotor task in the real world or a virtual environment. Driving a robot via teleoperation is a sensorimotor task in which the operator has reduced situational awareness, thus it would seem logical that training robot operators using a high fidelity simulator would be as effective as training using a real robot. This was demonstrated by Cromby, et al. in the virtual supermarket experiment. By providing a similar environment to a real super market, the participants learned to focus on the given task and performed better than the untrained control group when placed in a more complex real world supermarket. As well, the aviation industry has long demonstrated that simulators are effective training environments for complex tasks performed by individual operators.

Suebnukarn, et al. in [Suebnukarn 2007] show, it is possible for computer based training applications to be effective in a multi-user scenario; however, their work is limited to solving word problems as a group in a chat room. Yamnill [Yamnill 2001]
states that the more a training program reflects the actual work environment, the more
effective the near transfer of that training. The Yamnill article defines near transfer as
follows “near transfer would be the objective of short-term skill development that can be
applied immediately to improve performance in one’s present position.” We can leverage
the fact that the distributed, multi-user training worked for the word problems as the basis
for the hypothesis that distributed, multiplayer virtual environments can be used to train
operators to perform coordinated physical tasks with robots. The SARGE simulation
environment was created to provide a high fidelity simulation of a robot and a search
environment similar to the environments encountered in real world urban search and
rescue operations. The simulator was designed leveraging past work showing the
effectiveness of high-fidelity simulations for near-transfer training.
Chapter 4

SARGE Implementation

This chapter describes the implementation of the portions of SARGE that will be used for this dissertation. The SARGE simulation environment already included a GUI menu system for selecting the play mode, the environment, and setting options; as well as multiplayer support which had been tested with up to twenty simultaneous players. The main additions for this dissertation were the enhancement of the VGTV model to include polymorphic capabilities and GUI elements similar to the real OCU GUI, the addition of the real-time fractal path analysis system, the creation of four environments that were used during training, a score keeping system for the new environments, and a data logging system.
4.1 Inuktun Extreme VGTV

Tracked vehicles are difficult to simulate. Typically tracks are simulated using many wheels or using a sliding surface. SARGE uses a sliding surface to simulate the tracks of tracked vehicles. At each contact point the physics system automatically applies the force necessary to keep the vehicle resting on the surface. To move the vehicle, a force is applied at each contact point in the direction the track would move while the texture of the track is shifted to give the illusion that the track is turning. This is the typical approach used in video games which include tracked vehicles such as tanks. Simulating individual track links is too computationally demanding and often produces inferior results. By applying a force at the contact points for each track, the vehicle responds similarly to the real VGTV. Drag forces are applied automatically by the physics engine to each track depending on the surface properties of the object the track is in contact with. Figure 3 shows a screenshot of the VGTV model in SARGE. The camera tilt mechanism of the VGTV was implemented by modeling the “head” of the vehicle as a separate object which can be rotated independently of the main body, but the head is parented to the body so that it moves along with it. The polymorphic implementation required the tracks to be constructed in small, rectangular segments which can be animated individually.
Figure 3. Simulated Inuktun VGTV.
4.2 Real-Time Fractal Path Analysis

The real time fractal path analysis (RTFPA) component uses a dynamic programming solution based on the averaging method used by With [With 1994]. This method averages the dimension (D) value found by successive evaluations of a path from various starting points near the beginning of the path. Nams [Nams 2006] showed that this method produces a more accurate estimate of D than using a single evaluation of a path. Nams also showed that a D value calculated by averaging the D values estimated using a forward and reverse traversal of a path generally provides an even better estimate; however, a reverse traversal of a path is not possible when using a dynamic programming solution.

In a standard fractal path analysis two or more spatial scales are used to measure the path length. The slope of the log-log plot of estimated lengths versus spatial scales determines the D value for the path where \( D = 1 - \text{slope} \). As the spatial scale used increases so does the underestimation of the path length, thus the slope is always less than or equal to zero leading to a dimension greater than or equal to 1. A conventional algorithm would remeasure the entire path length at each spatial scale for every new reading that comes in. Not only would this waste energy by recomputing the entire path for each update, but does so with \( O(n^2) \) time complexity. To eliminate this problem the RTFPA algorithm is based on a dynamic programming solution. For each position update that comes in the
total estimated path length is updated for each spatial scale used only from the end of the previous calculation. This allows the RTFPA algorithm to run with $O(n)$ time complexity.

The path length is estimated by performing successive line/sphere intersection calculations with the sphere radius equal to the spatial scale and the sphere center starting at the beginning of the path. For each intersection calculation it is possible to have zero, one, or two intersection points. In the case of zero or one intersection points, the radius of the sphere extends past the current segment being evaluated. The evaluation at a given spatial scale is skipped until a new reading arrives at which point the segment end point is moved to the position of the new reading. This process repeats until a valid two-point intersection is computed. If two intersection points are found then the sphere center is moved to the point closest to the segment end point, the radius is added to the path length sum and the process is repeated until only intersection point is found.
4.3 Environments

SARGE included six environments which were used in prior research activities. For this dissertation, four more environments were created. Each environment is implemented similarly. Environment features are created in an external 3D modeling application, textured using several 2D and 3D texturing applications and finally placed into a Unity scene where all objects are arranged to create the final environment. For a detailed description of the creation of previous environments see [Craighead 2008b]. The layout of the four environments that were used for this dissertation are shown in below. The first two challenge environments are shown in Figures 4 and 5. These environments consist of a simple room mesh and cup objects. The cup objects are assigned properties that allow them to be controlled by the physics engine and thus interact with the robot while the room mesh simple is a static collider object, which allows physics system objects (such as the cups and robot) to collide with floor and walls. The search environment, shown in Figure 6, is modeled after a crawl space in a collapsed building similar to one used during a training exercise in which CRASAR participated at the end of 2005 [Craighead 2006].
Figure 4. Slalom Course. The left image shows the concept layout for the level and the right image shows the final design as implemented. The player must avoid hitting the cups while hitting the markers on the path.

Figure 5. Step Course. The left image shows the concept layout for the level and the right image shows the final design as implemented. The player use the robot to push each of the cups into the goal area.
Figure 6. Search Course. The left image shows the concept layout for the level and the right image shows the final design as implemented. The player use the robot to push each of the cups into the goal area.
4.4 Data Logging

SARGE automatically records score, robot pose (position and orientation of the robot), and fractal-D data to an SQLite database at runtime. This is handled via a logging class which any object in the environment can reference. This logging class uses a thread pool to handle incoming log requests. The VGTV control script is configured to log environment, position, rotation, raise, tilt, light intensity, and camera zoom every 1/2 second. The fractal-D script which is also attached to the robot logs environment, position, D, mean step size, path length, minimum estimated path length, maximum estimated path length, number of steps in the current path and mean step velocity. The pose logging allows a fine grained recreation of the path taken by the player at a later date while the fractal-D logging allows for the analysis of the evolution of the D value over the course of play for each level.
This chapter discussed the features that have been implemented in SARGE for this dissertation. SARGE had been in development for nearly two years, was used successfully in previous work and provided a good multi-user framework for conducting the team training experiments. Several additions to SARGE were required for this dissertation. Four new environments were created specifically for training. The Inuktun VGTV model will be enhanced to add support for polymorphism. A real-time fractal path analysis algorithm was created and integrated into a position tracking and logging system which was used to record and archive data automatically for later analysis
Chapter 5
Experiments

The experiment for this work was conducted at both the main Texas A&M University (TAMU) campus in College Station, Texas as well as the adjacent Disaster City training facility over a four day period in May and June 2009. The experiment was divided into two parts. Part One used the SARGE multiplayer robot simulation game [Craighead 2008a, Craighead 2008b] to train the participants in the operation of an Inuktun Extreme VGTV for search tasks. Part One took place in a computer lab at Texas A&M in College Station, TX. During Part One, one-half of the participants received individual operator training while the remaining half will received both individual operator training as well as team training in pairs. The training sessions took on average 1 hour per group of participants, regardless of whether the group was a team training group or a individual training group. Part Two measured the proficiency of ad-hoc teams of participants using a real Inuktun Extreme VGTV to conduct a search task in a collapsed building. Part Two took place at the “pancake house” building at the TEEX Disaster City training facility in College Station, TX. The teams were scored on the number of relevant items found in the building. Relevant items included anything that may have indicated the presence of a human (clothing, equipment, and body parts).
A power analysis using the G*Power program [Faul 2007] showed the experiment required a minimum of 40 teams (80 participants) to achieve a power of 0.9 with alpha = 0.1 if we assumed an effect size of 0.5. It was estimated that ten teams (20 participants) could be processed in one day. However, Disaster City was only available on weekends. It was estimated that a power of 0.7 could be achieved with 20 teams (40 participants) in the case that 80 participants were not available. Over the four day period forty-one volunteers participated in the experiment. Due to an adjustment made to the complexity of the field trial following the participation of the first two groups, thirty-seven of the participants produced usable data. This adjustment was made after the two groups failed to find all but one of the items placed in the collapsed building. Part One was run in two three-hour blocks, with up to ten participants per block, over four days. Part Two was run in ten thirty-minute blocks with one team per block over a five hour period each day for four days. Part Two ran concurrently with the second block of Part One.

Finally it must be emphasized to the reader that the sample size used in this dissertation is not large enough to claim statistical significance. The sample size was limited in this case due to the short span of time which was available to recruit participants. While the ANOVA and regression analyses reported in Chapter 6 show that trends do exist, without further experimentation using a much larger sample it is not possible to know if these effects will generalize to the population.
Participants were recruited by Disaster City from their existing volunteer pool via e-mail. They were directed to use Disaster City’s volunteer website to register for their preferred day and time (9:00 am or 12:00 pm). About half of the registered volunteers actually participated in the experiment, the remaining volunteers failed to arrive at their scheduled times. In total 41 volunteers participated in the experiment. As participants arrived at the TAMU computer lab they were given an ID number and the informed consent documents, they were then seated at a computer running SARGE. Seating was assigned based on ID number such that team training partners were not visible to each other. The computers were equipped with stereo headsets with a microphone and a Logitech Dual Action game pad (a Sony PlayStation style joystick).

Once the entire group was present the participants were instructed on the login procedure for SARGE, which required them to enter their assigned ID number. ID numbers were created in groups of ten and determined if the player was to participate in a team training exercise or an individual training exercise. If the ID corresponded to a team training exercise, the ID was also used to automatically connect them to their predetermined partner. To ensure that participants were not likely to partner with a person whom they knew outside the experiment ID numbers were not given out sequentially when multiple participants arrived together. During the training session SARGE recorded the completion time for each training game as well as the robot's movement, scoring
events, and the real-time Fractal-D. This information was uploaded to a remote server automatically after the end of each training session.

The participants received 40 minutes of individual training in SARGE. This training was divided into two parts. The first part introduced the player to basic operation of the robot through a short scripted tutorial. The tasks included basic robot control using the joystick, operation of the raise and camera tilt mechanisms, obstacle traversal, and the use of the robots headlights lights and camera zoom. The tutorial covered these topics in a series of mini-missions that presented both internal and external views of the robot.

Following the scripted tutorial the participants played two games that required them to one, maneuver the robot towards waypoint markers and and avoid plastic cups while following a zig-zag path, and two, to push scattered plastic cups into a goal area. The games were timed and the score and penalties were recorded for each game. The participants played each of the games for a maximum of 15 minutes or until they chose to continue to the next game. After a participant completed the first two games they proceeded to play a search game. See Figures 4 and 5 for diagrams of the training games.

The search game required the participants to locate ten objects related to a human presence in a confined space within 20 minutes. The goal was to maneuver the robot near each object, select it using the mouse, and then provide a description using the keyboard. If a participant was assigned to the individual training group they completed the search
task alone. If the participant was assigned to the team training group they were randomly paired with a teammate from the team training group. One of the teammates drove the robot while the other took on the role of observer. The observer was instructed to provide support to the operator and concentrate on locating victims or objects that might indicate the presence of a victim nearby. Additionally the observer was instructed to draw a map of the environment and mark the location of identified objects. The driver and observer communicated using the Mumble VoIP application. See Figure 6 for a layout of the search game.

The participants were scheduled to run simultaneously in groups of ten (five teams). This scheduling allowed the participants to remain physically separated from each other while in one room. This schedule was run two times per day, allowing twenty participants (ten teams) to be trained per day in a six hour time period. The teams participated in Part Two immediately following completion of their training. This schedule allowed the experiment to be completed in four days. Part One of the experiment required one research assistant to give directions to participants and set up the computer workstations before each group arrived.
Following the training session the participants were provided a map and instructed to drive to Disaster City, which is adjacent to the TAMU campus. Participants who were in the individual training groups were paired with the person assigned the following ID number; for instance ID 11 was paired with ID 12, 13 with 14, 15 with 16, etc. Participants who were in the team training groups were paired in a slightly more complicated manner. An attempt was made to stratify the participants within the teams in such a way that any effect of role swapping on the search score would be eliminated. There were four types of team groupings that could occur in this case: Driver-Observer Same Roles, Driver-Observer Swap Roles, Driver-Driver, and Observer-Observer. The Driver-Observer Same Roles case pairs a driver and an observer who performed those same roles during their team training in the video game. The Driver-Observer Swap Roles case pairs a driver and an observer who performed opposite roles during their team training in the video game, the driver was an operator during training and the observer was a driver during training. The Driver-Driver case pairs two participants that were drivers during training, one of them became an observer for the field trial. The Observer-Observer case pairs two participants that were observers during training, one of them became a driver for the field trial.

Participants were taken, one team at a time, to the collapsed building where the experiment was set up while the remaining participants in the group were given a tour of
the Disaster City facility while they waited. The participants were shown the robot and
the operator control unit (OCU) and given a one minute overview of where the various
controls were located on the OCU. The teams were directed to communicate with each
other via a hand-held radios during the search. The observer was instructed to draw a map
of the environment and assist the driver in identifying potential victims. The driver
remained at the OCU station while the observer was taken to the opposite side of the
collapsed building so that they were out of visual and audible range of each other. This
ensured the team was forced to communicate strictly via the radio. The observer could
see the same video from the robot that was available to the operator. This was transmitted
wirelessly using a small analog transmitter attached to the robot OCU to a PC monitor at
the observer’s station.

Once the participants were in place they were given twenty minutes to search the
building for victims. Following the search each participant was given a short
demographic survey. A team’s score was kept with pen and paper by the author as the
teammates confirmed to each other that they were viewing a victim or victim related
object. An additional factor that was noted for each team was a subjective amount of
communication between the pair. This was rated at the end of the search by the author as
“little”, “some”, and “a lot”. These ratings correspond to only talking when an object
identification was necessary, talking for object identification and some navigation, and
talking almost continuously about the search. Surprisingly, there was no off topic
communication by any of the teams. The building layout and pictures of the items in the surrounding area can be seen in Figure 7 below.
Figure 7. Disaster City Collapse Layout. This diagram shows a picture of each item placed in the collapsed building for the field trial and their relative locations.
5.3 Analysis

The analysis of the data compared the mean field scores of the team training group to the individual training group using an ANOVA analysis. The hypothesis was that ad-hoc teams consisting of members that received team training would have significantly higher scores than the ad-hoc teams consisting of members that received only individual training. A regression analysis was used to test for a relationship between the Fractal-D values to the search scores obtained during training and during the field trial. The hypothesis associated with the fractal dimension was that more skilled (higher scoring) operators would have significantly lower fractal-D values than less skilled (lower scoring) operators. Furthermore, correlation, ANOVA and regression analyses were used to identify relationships among participant age, technology experience, gender, occupation, education level, fractal-D values, and search scores. For all analyses $\alpha=0.1$ was used, thus the results will be significant at the 90% confidence level.
5.4 Demographics

There were thirty-seven participants that provided usable data for this experiment. At the end of the field trial, each participant filled out a demographic survey. This section shows the results of that survey. Each table and figure below show the breakdown of responses to the survey questions. Table 1 and Figure 8 show the statistics for the gender of the participants. Table 2 and Figure 9 show the statistics related to the age of the participants. Table 3 and Figure 10 show the statistics for the occupation of the participants. Table 4 and Figure 11 show the statistics for the highest completed level of education for the participants. Table 5 and Figures 12-17 show the statistics for the participants’ level of experience with various technology items.

Each demographic variable was evaluated using a t-test to determine if there was a difference between individual training and team training groups. As Table 6 shows, there was a difference in group means for gender and for experience with video games. The table shows that the mean gender was 0.60 for the team training group and 0.24 for the individual training group. Gender was coded with Male=0 and Female=1, thus the means show that more females were assigned to the team training group than the individual training group. The means for experience with video games are 3.33 for the individual training group and 2.55 for the team training group indicating that more experienced gamers were assigned to the individual training group.
Table 1. Participant Gender. This table shows the number of participants of each gender.

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>21</td>
<td>56.8</td>
</tr>
<tr>
<td>Female</td>
<td>16</td>
<td>43.2</td>
</tr>
</tbody>
</table>

Figure 8. Participant Gender. This graph shows the total number of participants of each gender.
Table 2. Participant Age. This table lists the minimum, mean, maximum, and standard deviation of the ages of the participants.

<table>
<thead>
<tr>
<th>Min</th>
<th>Mean</th>
<th>Median</th>
<th>Max</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>34.17</td>
<td>31</td>
<td>63</td>
<td>13.377</td>
</tr>
</tbody>
</table>

Figure 9. Participant Age. This graph shows the distribution of the ages of the participants.
Table 3. Participant Occupation. This table shows the total number of participants in each occupational field as marked on the survey.

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emergency Response</td>
<td>6</td>
<td>16.2</td>
</tr>
<tr>
<td>Law Enforcement</td>
<td>1</td>
<td>2.7</td>
</tr>
<tr>
<td>Military</td>
<td>2</td>
<td>5.4</td>
</tr>
<tr>
<td>Student</td>
<td>18</td>
<td>48.6</td>
</tr>
<tr>
<td>Other</td>
<td>10</td>
<td>27.0</td>
</tr>
</tbody>
</table>
Figure 10. Participant Occupation. This graph shows the total number of participants in each occupational field as marked on the survey.
Table 4. Participant Education. This table shows a count of the highest completed educational level of the participants as marked on the survey.

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Some HS</td>
<td>2</td>
<td>5.4</td>
</tr>
<tr>
<td>HS Grad</td>
<td>1</td>
<td>2.7</td>
</tr>
<tr>
<td>Some College</td>
<td>10</td>
<td>27.0</td>
</tr>
<tr>
<td>AA/AS</td>
<td>1</td>
<td>2.7</td>
</tr>
<tr>
<td>Bachelor</td>
<td>9</td>
<td>24.3</td>
</tr>
<tr>
<td>Some Grad.</td>
<td>7</td>
<td>18.9</td>
</tr>
<tr>
<td>Masters</td>
<td>6</td>
<td>16.2</td>
</tr>
<tr>
<td>Phd/MD/etc</td>
<td>1</td>
<td>2.7</td>
</tr>
</tbody>
</table>
Figure 11. Participant Education. This graph shows the distribution of the highest completed educational level of the participants as marked on the survey.
Table 5. Participant Technology Experience. This table shows the minimum, mean, and maximum responses for each of the five technology items listed on the survey. The standard deviation is listed for each item as well. A score of 1 indicates little or no experience with a particular item, a score of 5 indicates an expert level of experience with that particular item.

<table>
<thead>
<tr>
<th>Item</th>
<th>Min</th>
<th>Mean</th>
<th>Median</th>
<th>Max</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>R/C Vehicles</td>
<td>1</td>
<td>2.23</td>
<td>2</td>
<td>4</td>
<td>1.114</td>
</tr>
<tr>
<td>Robots</td>
<td>1</td>
<td>1.83</td>
<td>1</td>
<td>5</td>
<td>1.2</td>
</tr>
<tr>
<td>Video Games</td>
<td>1</td>
<td>2.89</td>
<td>3</td>
<td>5</td>
<td>1.278</td>
</tr>
<tr>
<td>Video Cameras</td>
<td>1</td>
<td>3.03</td>
<td>3</td>
<td>5</td>
<td>1.043</td>
</tr>
<tr>
<td>Search Equipment</td>
<td>1</td>
<td>1.31</td>
<td>1</td>
<td>4</td>
<td>0.758</td>
</tr>
<tr>
<td>Computers</td>
<td>1</td>
<td>4.17</td>
<td>4</td>
<td>5</td>
<td>0.954</td>
</tr>
</tbody>
</table>
Figure 12. Experience Level with R/C Vehicles. This graph shows the number of responses for each experience level. 1 is little to no experience, 5 is expert level experience. Notice there were no responses at the 5 level for experience R/C vehicles.
Figure 13. Experience Level with Robots. This graph shows the number of responses for each experience level. 1 is little to no experience, 5 is expert level experience. Notice that the majority of the participants had no prior experience with robots.
Figure 14. Experience Level with Video Games. This graph shows the number of responses for each experience level. 1 is little to no experience, 5 is expert level experience. The distribution for experience with video games is relatively normal.
Figure 15. Experience Level with Video Cameras. This graph shows the number of responses for each experience level. 1 is little to no experience, 5 is expert level experience. Notice that a majority of the participants had at least some experience with video cameras; however, almost no one considered themselves to be an expert.
Figure 16. Experience Level with Search Equipment. This graph shows the number of responses for each experience level. 1 is little to no experience, 5 is expert level experience. The majority of participants had no prior experience with search and rescue equipment.
Figure 17. Experience Level with Computers. This graph shows the number of responses for each experience level. 1 is little to no experience, 5 is expert level experience. Nearly all participants considered themselves to be very proficient or expert computer users.
Table 6. Differences in Demographic Means by Team Training Group. This table shows the means for the demographic variables in which a significant difference in means existed between training groups. Group means were compared using a t-test with $\alpha=0.1$. Gender was coded with Male=0, Female=1.

<table>
<thead>
<tr>
<th></th>
<th>Individual Training Group</th>
<th>Team Training Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>$\mu=0.24 \ \sigma=0.44$</td>
<td>$\mu=0.6 \ \sigma=0.50$</td>
</tr>
<tr>
<td></td>
<td>$p=0.024$</td>
<td>$p=0.024$</td>
</tr>
<tr>
<td>Experience with Video</td>
<td>$\mu=3.33 \ \sigma=1.40$</td>
<td>$\mu=2.55 \ \sigma=1.10$</td>
</tr>
<tr>
<td>Games</td>
<td>$p=0.072$</td>
<td>$p=0.072$</td>
</tr>
</tbody>
</table>
5.5 Dependent Measures

Four of the variables measured during the experiment proved to have relationships with one or more of the independent variables or with each other. These variables are in-game search score, mean fractal dimension, field search score, and field communication frequency. In-game search score is the score an individual received during the final training level. It should be noted that for the team training group the observers received the same score as their drivers. Figure 18 and Table 7 show the statistics for the in-game search score. The mean fractal dimension for an individual was calculated from the average fractal dimension value of the three training games that followed the training introduction. It should be noted that for the team training group the observers received the same fractal dimension as their drivers in the final search level. Figure 19 and Table 7 show the mean fractal dimension statistics. Field score is the number of items found by a team during the field trial. Both members of the team were given the same score at the end of the trial. Figure 20 and Table 7 show the statistics for field score. Communication frequency is a three level variable that was measured subjectively by the author throughout the field trials. The rating groups intra-team communication frequency into “little”, “some”, and “a lot” groups. Figure 21 and Table 7 show the statistics for communication frequency. It should be noted that for each of these measures, only the driver’s score was used in the analyses to prevent biasing towards the team training group since the observers were assigned the same score, fractal dimension, and communication frequency as their partner.
Table 7. Descriptive Statistics of Dependent Measures.

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Mean</th>
<th>Median</th>
<th>Max</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search Score</td>
<td>0</td>
<td>8.259</td>
<td>9</td>
<td>10</td>
<td>2.443</td>
</tr>
<tr>
<td>Fractal Dimension</td>
<td>1.05</td>
<td>1.104</td>
<td>1.10</td>
<td>1.15</td>
<td>0.026</td>
</tr>
<tr>
<td>Field Score</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>1.856</td>
</tr>
<tr>
<td>Communication Frequency</td>
<td>0</td>
<td>1.06</td>
<td>1</td>
<td>2</td>
<td>0.854</td>
</tr>
</tbody>
</table>

Figure 18. Search Score Histogram. This figure shows the frequency distribution of in-game search score. The minimum possible score was 0 and the maximum possible score was 10.
Figure 19. Mean Fractal Dimension Histogram. This figure shows the frequency distribution of the participants' mean fractal dimension rounded to two decimal places.
Figure 20. Field Score Histogram. This figure shows the frequency distribution of field trail scores. The minimum possible score was 0 and the maximum possible score was 6.
Figure 21. Communication Frequency Histogram. This figure shows the number of participants in each of the communication frequency groups. 0 represents the “low” group, 1 represents the “some” group, and 2 represents the “a lot” group. Communication frequency was not recorded until a difference between groups was noticed by the author during the experiment, thus data for only 16 of the 19 teams was recorded.
This chapter has presented the experimental design that was used to test the effectiveness of a distributed, multiplayer video game for training ad-hoc teams using a series of robot mini-games and a search task in SARGE, a robot simulation game. In these mini-games participants received operator training for an Inuktun Extreme VGTV robot. Half of the participants additionally received team training with a randomly selected partner. During operator training a real-time fractal path analysis tool was used to estimate a fractal-D value for each operator. Following training the teams participated in a field trial in which they searched for items related to victims in a collapsed building. The field trial teams consisted of randomly paired participants to create ad-hoc teams. Teams were scored on the number of items located during the search. An ANOVA analysis was used to determine the mean difference in field scores between teams that received team training and teams that received only individual training. A regression analysis was used to identify the nature of the relationship between an operator’s fractal-D value and their score. Correlation, regression, and ANOVA analyses were used to identify relationships between fractal-D values, in-game and field scores and the collected demographic data. Section 5.4 presented the demographics of this studies 37 participants that were used in the analyses and Section 5.5 explained the dependent measures.
Chapter 6

Results

This chapter discusses the results of the data analysis as well as the implications of the findings. There were eight major findings and one observation that resulted from the analysis. Each finding, and a discussion of the finding is described in one of the following sections of this chapter. Below is a summary of the findings. A discussion of the implications of these findings as well as suggestions for improving the experiment in the future appear in Section 6.11 and 6.12 respectively. Section 6.13 provides a summary of this chapter.

• Finding: Team training has no effect on field score. \([F(1,17)=0.980 \ p=0.336]\)

• Finding: Fractal dimension values have a non-linear relationship \((\beta_0 + \beta_1 \ D + \beta_2 \ D^2)\) with a driver’s in-game search score. Score increases with D up to a peak, then decreases as D continues to increase. \([F(2,23)=4.125 \ p=0.029]\)

• Finding: Fractal dimension has a positive linear relationship with a driver’s prior gaming experience. \([F(1,22)=9.140 \ p=0.006]\)

• Finding: A driver’s prior video game experience has a positive linear relationship with field score. \([F(1,17)=8.425 \ p=0.010]\)
• Finding: A driver’s prior video game experience has a positive linear relationship with in-game search score. [F(1,23)=3.637 p=0.069] Driver age is a predictor of in-game search score in this model. [F(1,18)=5.125 p=0.036]

• Finding: Fractal dimension has a negative linear relationship with driver age. [F(1,23)=9.865 p=0.005]

• Finding: A driver’s age has a negative linear relationship with in-game search score. [F(1,24)=13.775 p=0.001]

• Finding: A driver’s age has no effect on field score. [F(1,16)=0.101 p=0.754]

• Observation: Team training appears to affect intra-team communication frequency for ad-hoc teams. [F(1,14)=12.250 p=0.004]
6.1 Correlations

For the first analysis of the collected data a simple correlation analysis was performed to determine if there were any linear relationships between the independent measures and team training. The correlations found at the $\alpha=0.05$ significance level in the collected data are shown in Table 8 highlighted. The third column shows that there is a positive correlation between communication frequency and team training. The fourth column indicates that there is a positive correlation between gender and team training. This means that there were more females that received team training compared to the number of males that received team training as a result of the random assignment which ignored gender. The fifth column indicates that there is a negative relationship between a participants self assessed experience with video games and both age and gender. This indicates that more of the younger and more of the male participants considered themselves to have a high degree of experience with video games, which is unsurprising. The sixth column indicates that there is a negative correlation between the mean fractal dimension recorded during training and age as well as a positive correlation between the fractal dimension and a participants experience with video games. The seventh column indicates that a positive correlation exists between field score and a participants experience with video games. The final column indicates that there is a negative relationship between a participants age and search score, which is the score a participant earned during the last level of the training game. Any relationships identified at this step
were further analyzed using ANOVA, ANCOVA, and regression analyses. The results for these analyses are presented in the following sections.

The most interesting of these correlations are the correlation between team training (TeamTrain) and communication frequency (Comms), the correlation between a experience with video games (ExperienceVideoGames) and mean fractal dimension from training (AltMeanD), and finally the correlation between experience with video games and field score. What is also interesting is that there is not a correlation between experience with video games and training search score (SearchScore). If we were to consider correlations at $\alpha=0.1$ significance level then the correlation between experience with video games and training search score is consistent with the correlation between experience with video games and search score. The relationship between team training and experience with video games indicates that participants assigned to team training group had less prior gaming experience than the participants assigned to the individual training group. This may be important considering that there is no correlation between team training and a field score. This is further discussed in Section 6.2.
Table 8. Correlations with the Dependent Measures. This table lists all the correlations between measured variables. The correlations that are significant at the $\alpha=0.05$ level are highlighted.

<table>
<thead>
<tr>
<th></th>
<th>Comms</th>
<th>Age</th>
<th>TeamTrain</th>
<th>Gender</th>
<th>ExperienceVideoGames</th>
<th>AltMeanD2</th>
<th>FieldScore</th>
<th>SearchScore</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>1</td>
<td>-0.086</td>
<td>0.665**</td>
<td>0.194</td>
<td>-0.305</td>
<td>0.140</td>
<td>-0.143</td>
<td>0.145</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>31</td>
</tr>
<tr>
<td>Age</td>
<td>-0.086</td>
<td>1</td>
<td>-0.129</td>
<td>-0.219</td>
<td>-0.418**</td>
<td>-0.595**</td>
<td>-0.144</td>
<td>-0.433**</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.646</td>
<td>0.455</td>
<td>0.199</td>
<td>0.014</td>
<td>0.403</td>
<td>0.088</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>35</td>
<td>36</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>TeamTrain</td>
<td>0.665**</td>
<td>-0.129</td>
<td>1</td>
<td>0.367</td>
<td>-0.308</td>
<td>0.248</td>
<td>-0.223</td>
<td>0.311</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.000</td>
<td>0.455</td>
<td>0.026</td>
<td>0.072</td>
<td>0.145</td>
<td>0.184</td>
<td>0.061</td>
<td>0.111</td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>36</td>
<td>37</td>
<td>37</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>Gender</td>
<td>0.194</td>
<td>-0.219</td>
<td>0.367</td>
<td>1</td>
<td>-0.463**</td>
<td>0.087</td>
<td>-0.265</td>
<td>0.235</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.297</td>
<td>0.199</td>
<td>0.026</td>
<td>0.005</td>
<td>0.614</td>
<td>0.113</td>
<td>0.161</td>
<td>0.161</td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>36</td>
<td>37</td>
<td>37</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>ExperienceVideoGames</td>
<td>-0.305</td>
<td>-0.418</td>
<td>-0.308</td>
<td>-0.463**</td>
<td>1</td>
<td>0.516**</td>
<td>0.371</td>
<td>0.291</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-0.101</td>
<td>0.014</td>
<td>0.072</td>
<td>0.005</td>
<td>0.002</td>
<td>0.028</td>
<td>0.090</td>
<td>0.090</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>34</td>
<td>35</td>
<td>35</td>
<td>34</td>
<td>35</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>AltMeanD2</td>
<td>0.140</td>
<td>-0.595**</td>
<td>0.248</td>
<td>-0.087</td>
<td>0.516**</td>
<td>1</td>
<td>0.191</td>
<td>0.279</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.459</td>
<td>0.000</td>
<td>0.145</td>
<td>0.614</td>
<td>0.002</td>
<td>0.264</td>
<td>0.099</td>
<td>0.099</td>
</tr>
<tr>
<td>N</td>
<td>30</td>
<td>35</td>
<td>36</td>
<td>36</td>
<td>34</td>
<td>36</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td>FieldScore</td>
<td>-0.143</td>
<td>-0.144</td>
<td>-0.223</td>
<td>-0.265</td>
<td>0.371</td>
<td>0.191</td>
<td>0.1</td>
<td>-0.180</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.444</td>
<td>0.403</td>
<td>0.184</td>
<td>0.113</td>
<td>0.028</td>
<td>0.264</td>
<td>0.286</td>
<td>0.286</td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>36</td>
<td>37</td>
<td>37</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>SearchScore</td>
<td>0.145</td>
<td>-0.433**</td>
<td>0.311</td>
<td>0.235</td>
<td>0.291</td>
<td>0.279</td>
<td>-0.180</td>
<td>1</td>
</tr>
<tr>
<td>Pearson Correlation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.438</td>
<td>0.008</td>
<td>0.061</td>
<td>0.161</td>
<td>0.090</td>
<td>0.099</td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>31</td>
<td>36</td>
<td>37</td>
<td>37</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>37</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).
In addition to experience with video games, the survey also asked participants about their prior experience with other technologies such as video camera, R/C vehicle, search equipment, and computers. The statistics associated with these questions can be seen in Table 5 and Figures 12-17. The survey also asked about gender and education level. These were compared against several dependent variables including field score, in-game search score, mean fractal dimension, and field search completion time as well as tested for their relationship with each other. Field search completion time did not correlate with anything except field search score, which was expected and obvious since those teams that found all the items in the field trial must have done so in less than the allotted time. Gender was shown to correlate only with experience with video games with men having more prior experience with video games than women. Gender was not correlated with any other survey question or dependent variable, thus gender was not included in any further analyses. Occupation was shown to have a slight correlation with training search score; however, the data is unreliable since the Law Enforcement and Military categories each had a single participant. Further investigation of Occupation showed it was not a factor in any of the dependent variables. Education level was only correlated with experience with computers however upon further investigation no relationship was, the majority of participants marked that they had a high degree of experience with computers. For the statistics associated with computer use, see Table 5. The remaining experience variables (R/C vehicles, robots, video cameras, and search equipment) were not correlated with the dependent variables although some were inter-correlated. Experience with R/C vehicles was correlated with all of the other experience
variables except experience with computers. Experience with robots was correlated with experience with search equipment and experience video games. Experience with video games was also correlated with experience with search equipment. Finally, experience with video cameras was correlated with experience with computers. Each of these correlations proved to have no relationships to any of the dependent variables and were not included in any of the follow up analyses. These correlations can be seen in Table 9.
Table 9. Inter-item Correlations from the Survey. The correlations that are significant at the $\alpha=0.05$ level are highlighted.

<table>
<thead>
<tr>
<th></th>
<th>ExperienceRC</th>
<th>ExperienceRobo</th>
<th>ExperienceVideoGames</th>
<th>ExperienceVideoCams</th>
<th>ExperienceSearchEquip</th>
<th>ExperienceComputers</th>
<th>Education</th>
<th>Gender</th>
<th>Occupation</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExperienceRC</td>
<td>1</td>
<td>0.046..</td>
<td>0.018..</td>
<td>0.001</td>
<td>0.160</td>
<td>0.014</td>
<td>-0.087</td>
<td>-0.044</td>
<td>-0.336</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.566</td>
<td>0.938</td>
<td>0.621</td>
<td>0.803</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>ExperienceRobo</td>
<td>0.646..</td>
<td>1</td>
<td>0.428..</td>
<td>0.286</td>
<td>0.772..</td>
<td>0.103</td>
<td>-0.021</td>
<td>-0.158</td>
<td>0.162</td>
<td>-293</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.000</td>
<td>0.010</td>
<td>0.096</td>
<td>0.000</td>
<td>0.554</td>
<td>0.907</td>
<td>0.365</td>
<td>0.354</td>
<td>0.092</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>ExperienceVideoGames</td>
<td>0.618..</td>
<td>0.016</td>
<td>0.025</td>
<td>0.069</td>
<td>0.712</td>
<td>0.168</td>
<td>0.005</td>
<td>0.308</td>
<td>0.014</td>
<td>-418</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.000</td>
<td>0.010</td>
<td>0.025</td>
<td>0.069</td>
<td>0.712</td>
<td>0.168</td>
<td>0.005</td>
<td>0.308</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>ExperienceVideoCams</td>
<td>0.551..</td>
<td>0.386</td>
<td>0.378..</td>
<td>1.000</td>
<td>0.438..</td>
<td>0.280</td>
<td>-0.193</td>
<td>-0.017</td>
<td>-0.124</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.001</td>
<td>0.096</td>
<td>0.025</td>
<td>0.568</td>
<td>0.008</td>
<td>0.103</td>
<td>0.267</td>
<td>0.924</td>
<td>0.486</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>ExperienceSearchEquip</td>
<td>0.535..</td>
<td>0.722</td>
<td>0.311</td>
<td>0.100</td>
<td>1</td>
<td>-0.199</td>
<td>-0.025</td>
<td>-0.079</td>
<td>-0.134</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.001</td>
<td>0.096</td>
<td>0.025</td>
<td>0.568</td>
<td>0.008</td>
<td>0.103</td>
<td>0.267</td>
<td>0.924</td>
<td>0.486</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>ExperienceComputers</td>
<td>0.100</td>
<td>0.103</td>
<td>0.065</td>
<td>0.438..</td>
<td>0.199</td>
<td>0.512..</td>
<td>0.045</td>
<td>0.103</td>
<td>-0.133</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.566</td>
<td>0.554</td>
<td>0.712</td>
<td>0.008</td>
<td>0.253</td>
<td>0.002</td>
<td>0.796</td>
<td>0.556</td>
<td>0.487</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>34</td>
</tr>
<tr>
<td>Education</td>
<td>-0.024</td>
<td>-0.021</td>
<td>-0.238</td>
<td>-0.280</td>
<td>-0.025</td>
<td>0.512..</td>
<td>0.010</td>
<td>-0.003</td>
<td>-0.271</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>-0.338</td>
<td>-0.168</td>
<td>-0.568</td>
<td>-0.193</td>
<td>0.103</td>
<td>0.079</td>
<td>-0.045</td>
<td>0.010</td>
<td>-0.218</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.087</td>
<td>-0.158</td>
<td>-0.463..</td>
<td>-0.193</td>
<td>-0.079</td>
<td>-0.045</td>
<td>0.101</td>
<td>0.049</td>
<td>-0.219</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.621</td>
<td>0.365</td>
<td>0.000</td>
<td>0.267</td>
<td>0.552</td>
<td>0.971</td>
<td>0.775</td>
<td>0.199</td>
<td>0.987</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>Occupation</td>
<td>-0.044</td>
<td>0.162</td>
<td>0.848</td>
<td>-0.017</td>
<td>0.154</td>
<td>0.014</td>
<td>-0.037</td>
<td>0.049</td>
<td>-0.152</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.683</td>
<td>0.354</td>
<td>0.308</td>
<td>0.924</td>
<td>0.443</td>
<td>0.556</td>
<td>0.987</td>
<td>0.775</td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>Age</td>
<td>-0.336</td>
<td>-0.293</td>
<td>-0.418..</td>
<td>-0.124</td>
<td>-0.193</td>
<td>0.271</td>
<td>-0.219</td>
<td>-0.152</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>0.652</td>
<td>0.092</td>
<td>0.014</td>
<td>0.486</td>
<td>0.356</td>
<td>0.847</td>
<td>0.109</td>
<td>0.199</td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>34</td>
<td>36</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).
* . Correlation is significant at the 0.05 level (2-tailed).
6.2 Team Training Has No Effect on Field Score

This section details the results of the experiment related to the support of the first research question posed in Chapter 1: *Does the performance of distributed, ad-hoc teams of robot operators improve if the team members have previously participated in online, multiplayer robot operator training?* Figure 22 shows box plots comparing field scores for the group that received team training versus the group that received individual training. The field score indicates the number of items a team found during the field trial. As can be seen in the figure there was clearly no difference between means given the large variance. This conclusion was also supported by an ANOVA test which did not support the presence of a reliable difference between training groups with $F=1.838$, $p=0.184$, Cohen’s $f=0.419$, and $1-\beta=0.804$.

This failure to reject the null hypothesis, that there is no difference in mean scores, indicates that 20 minutes of distributed, game-based team training is not sufficient to directly improve the performance of an ad-hoc team compared to an ad-hoc team consisting of individually trained members. It is possible that a longer training session may produce different results, however this is unlikely given the large variance in group means. Or it is possible that this lack of effect is due to an unbalanced distribution of participants based on their prior gaming experience. A t-test showed that there was a difference in the mean gaming experience level of the two training groups at the $\alpha=0.1$ significance level. The mean gaming experience level of the individual training group
was 3.33 while the team training group had a mean gaming experience level of 2.55. In the future it will be more useful to focus efforts game-based training efforts on other methods of improving team and operator performance.

Figure 22. Mean Field Score vs. Team Training. This graph shows a box plot comparing the field score for each training group. There is no statistical difference between groups.
6.3 Fractal D Has a Non-Linear Relationship with Search Score

The second research question posed in Chapter 1 was: *Does the fractal dimension of a tele-operated robot's path indicate the skill of its operator when conducting a search task?* The hypothesis was that the fractal dimension would have a negative relationship with an operator's skill, that is, D would decrease as an operator's skill increased. This hypothesis was based on the work of Voshell and Woods. Figure 23 shows the quadratic relationship between fractal dimension and the score earned in the search level of the training game (search score). This relationship was tested using a non-linear regression (\( \text{SearchScore} = \beta_0 + \beta_1 D + \beta_2 D^2 \)) with F(2,23)=4.125, p=0.029, \( R^2=0.264 \), Std. Err. = 2.206, Cohen's \( f^2=0.359 \), and 1-\( \beta=0.832 \). This curvilinear relationship supports both work by Voshell and Woods [Voshell 2005] that indicates that a higher situational awareness may be related to lower fractal dimensions (straighter paths) and supports findings by Clarke and Goldiez [Clarke 2007] which indicates that a chaotic path, measured using the Lyuoponov exponent, is related to higher search performance in a maze. At the low to mid end of the scale increasing fractal dimension is clearly related to an improvement in search score; however, as can be seen in the figure, a mean D value greater than 1.125 is related to decreasing search score.

The relationship between fractal dimension and search score suggest that the data supports the initial hypothesis and indicates that a lower fractal dimension is associated with a driver's skill in operating a robot but that this relationship is non-linear; at fractal
dimensions below a certain point, search score increases which may indicate that as
fractal D peaks a driver becomes proficient at operating the vehicle and they begin to
drive faster and search more of the area; this effect is consistent with the findings of
Clarke and Goldiez. At D > 1.12 perhaps D continues to increase as the operator drives in
an erratic pattern due to the inability to maintain situational awareness. This would most
likely cause their score decreases because of the decreased understanding of the
environment; this effect is consistent with the work of Voshell and Woods.
Figure 23. Mean Fractal Dimension vs. Training Search Score. This graph shows the non-linear relationship (Search Score = $\beta_0 + \beta_1 D + \beta_2 D^2$) between a participant’s mean fractal dimension (D) and their score in the training game. This relationship may indicate that there is a peak D that indicates a thorough search strategy. Values of D higher than the peak may indicate that an operator has decreased situational awareness and is struggling to localize themselves. Further experimentation is needed to verify this theory.
6.4 A Driver’s Prior Gaming Experience Increases Fractal Dimension

A regression analysis of participants that were drivers in the training game was used to test this relationship, and shows that a driver’s prior experience with games has a relationship with mean fractal dimension with $F(1,22)=9.140$, $p=0.006$, $R^2=0.294$, Std. Err.$=0.024$, Cohen’s $f^2=0.416$ and $1-\beta=0.911$. This is somewhat consistent with the findings from 6.3 that showed a relationship between search score and fractal dimension as well as the findings in Section 6.5 that show a positive relationship between experience with video games and field score. Figure 24 shows the relationship between mean fractal dimension and experience with video games. It is interesting to note the decrease in $D$ at the expert gaming level (5) corresponds to a decrease in field score at the same gaming experience level as shown in Figure 25. Upon further investigation it was found that all participants rating themselves as expert video game players were assigned to the individual training group; however, there were only four participants in this category which prevented further statistical analysis of this phenomenon.

During the experiment it was noticed that the teams that generally followed the Localize the robot, Observe the surrounding environment, look for Victims, Report findings (LOVR) [Burke 2004a] strategy scored the highest. This strategy was not explained or taught to the participants in any way, yet it appears that the more experienced game players may have already learned a similar strategy. Based on the relationship between prior gaming experience and in-game search score discussed in
Section 6.6 and the relationship between prior gaming experience and fractal dimension, it may be possible to use D to assess a driver’s use of the LOVR strategy. A successful use of the LOVR strategy would produce a D value near the peak of the curve shown in Figure 23.

Figure 24. Mean Fractal D vs. Prior Gaming Experience. This graph shows the relationship between a trainee’s prior experience with video games and the mean fractal dimension calculated during their training. The markers indicate the mean fractal dimension for each experience level.
6.5 A Driver’s Prior Gaming Experience Improves Field Score

Figure 25 shows the mean field score versus experience with video games. This analysis was performed only on the drivers of the robot. The figure indicates there is a positive relationship between a driver’s prior experience with video games and the team’s field score. A regression analysis suggests gaming experience is a predictor of field score with $F(1,17)=8.425$, $p=0.010$, $R^2=0.331$, Std. Err.=1.562, Cohen’s $f^2=0.495$ and $1-\beta=0.887$. This relationship is important because it shows that even though the team training introduced as part of this experiment did not have an effect on a team’s field score, there is some aspect of video game playing that improves a team’s performance in a search task. It should be noted when reading the following sections that in the sample of field trial drivers there was no relationship between age and prior gaming experience with $F(1,16)=0.388$ and $p=0.542$. The lack of a relationship with age is important because in Section 6.6 it is shown that age is a factor in predicting in-game search score.
Figure 25. Prior Gaming Experience vs. Field Score. This graph shows the relationship between field score and video game. The markers indicate the mean score at each experience level.
6.6 A Driver’s Prior Gaming Experience Improves Search Score

Figures 26 shows the relationship between the robot drivers’ in-game search score and experience with video games. A regression analysis indicates that prior gaming experience is a predictor of in-game search score at $F(1,23)=3.637$, $p=0.069$, $R^2=0.137$, Std. Err.=2.260, Cohen’s $f^2=0.159$ and $1-\beta=0.598$. However, as will be described in Section 6.8 a driver’s age also has a relationship with in-game search score and thus should be included in the regression model. When age is added to the model it becomes significant at the $\alpha=0.01$ level with $F(2,21)=5.980$, $p=0.009$, $R^2=0.363$, Std. Err.=2.011, Cohen’s $f^2=0.570$ and $1-\beta=0.927$. This age effect is most likely the cause of the spike at experience level 2 in Figure 26. Figure 27 (left) shows the median age of participants who were drivers during training for each gaming experience level, notice the dip at experience level 2. This can be compared to Figure 27 (right) that shows a more normal distribution of gaming experience across the age range. As discussed in Section 6.9 there is no relationship between age and field score. A regression analysis shows that there is in fact a relationship between age and prior gaming experience for the set of training drivers with $F(1,22)=7.894$ and $p=0.010$. This may explain why prior gaming experience is shown to have no relationship when age is added into the regression model predicting search score.
Figure 26. Prior Gaming Experience vs. In-Game Search Score. This graph shows the relationship between a driver’s prior experience with video games and their in-game search score. Age is a factor in this relationship.
Figure 27. Median Age vs. Gaming Experience Level. This figure shows the median age of each level of video game experience for the participants who were drivers during training on the left and for the participants who were drivers during the field trial on the right.
6.7 A Driver’s Increasing Age Decreases Fractal Dimension

Figure 28 shows the negative relationship between age and a driver’s fractal dimension that was confirmed with a regression analysis with $F(1,23)=9.865$, $p=0.005$, $R^2=0.300$, Std. Err.=0.023, Cohen’s $f^2=0.429$ and $1-\beta=0.928$. Because the spatial scale is automatically calculated in the RTFPA algorithm based on mean distance between each point in the record the lower D may indicate that a driver is more likely to operate the vehicle in a slower, more cautious manner as their age increases. This conclusion is reached because the slower the vehicle moves, the smaller the spatial scale used to calculate D and smaller spatial scales will produce a lower fractal dimension for a given path. This relationship between age and mean fractal dimension is important for later analyses that involve other variables that are also highly correlated to age or mean fractal dimension.
Figure 28. Mean Fractal Dimension vs. Age. This figure shows the relationship between a participant’s mean fractal dimension as measured in training and their age. An increase in age appears to lead to a decrease in mean fractal dimension. The trend line passes through the mean age for each fractal dimension value.
6.8 A Driver’s Age Has a Negative Relationship with Search Score

A regression analysis was used to test the relationship between a driver’s age and their in-game search score. The relationship was found with $F(1,24)=13.775$, $p=0.001$, $R^2=0.365$, Std. Err.=2.006, Cohen’s $f^2=0.575$ and $1-\beta=0.979$ and is shown Figure 29. This finding is important because it affects the findings in Section 6.3 that suggests increased fractal D leads to improved search performance, Section 6.6 that shows age is a factor in the relationship between prior gaming experience and in-game score, as well as Section 6.7 that shows that fractal D decreases as age increases. This leads to the question of how age affects performance? This link of age to reduced performance is not a new or unexpected finding, in fact there are many articles in the psychology and medical literature that show a clear reduction in motor control [Morgan 1994] and visual acuity [Klein 1995] which could cause difficulty in operating a robot. The loss of motor control and visual acuity may lead an older driver to drive more cautiously. A more cautious driver may drive the vehicle slower and less erratically out of fear of damaging the robot or loosing situational awareness. This type of movement would reduce the area searched and in turn lower the fractal dimension value and the number of items found. Additionally, the RTFPA algorithm uses mean step length as a base value for determining the spatial scale used in the analysis, therefore, slower movement would reduce the mean step length and reducing D even further.
If this finding can be validated the implication is that law enforcement, military, and US&R teams should recruit potential robot operators from a pool of young officers or find a way to overcome the negative effects of age. While this may seem discriminatory, there is clear evidence that operators in the younger demographic are more effective at performing the search task. This may generalize to all robot tele-operation tasks because the US&R task seems to be one of the more difficult robot operating tasks due to the nature of the confined space and the design of the robot which is constrained to fit in said space. That said, Section 6.9 contradicts the evidence presented here and shows that a driver’s age has no effect on field search score. The difference in the distribution of gaming experience versus age, shown in Figure 26, may partially contribute to the discrepancy between the effect of age on search score versus the effect of age on field score. It is also likely that a large portion of the difference comes from the different robot controllers used during training and the field trial. Additional experimentation which controls for age, gaming experience and robot controller type is necessary to identify which factor is the cause for the discrepancy in age effects.
Figure 29. Effect of Driver Age on Search Score. This figure shows a box plot comparing age of participants who were drivers during training to their search score. The negative effect of age on search score is clearly visible in this plot. Note the circle labeled 30 above the 10.0 value on the Search Score axis is an outlier that was included in the analysis. 30 is just an ID number and does not indicate the age value of the data point.
6.9 A Driver’s Age Has No Effect on Field Score

Contradicting the evidence that age is a factor in robot operator performance shown in Section 6.8, a regression analysis of the relationship between a driver’s age and field search score shows that there is in fact no relationship between age and field score as shown in Figure 30 and confirmed using a regression test with $F(1,16)=0.101, p=0.754$, $R^2=0.006$, Std. Err.=1.946, Cohen’s $f^2=0.006$ and $1-\beta=0.116$. The question is which model is correct? Does operator age have a relationship with their ability to perform a search task or not? One possible explanation for this is that the set of drivers for the field trial was different from the set of drivers during training. The distribution of driver ages was nearly identical between the training and field trial, thus there must be another factor contributing to the difference in the findings. The relationship between age and prior gaming experience differed between drivers during training and the field trial. In the field trial there was no relationship between age and prior gaming experience level with $F(1,16)=0.388, p=0.542$ while during training there was a relationship with $F(1,22)=7.894$ and $p=0.010$, these relationships can be seen in Figure 26. If we assume the discrepancy between the effect of age on search score and field score is in fact caused by the distribution of gaming experience instead of familiarity with the controller or random chance, then the data suggests that gaming experience can be used to overcome the negative effect of age on search score. Further study is needed to resolve this discrepancy and identify if the effect on search score exists due to age or controller fidelity.
Figure 30. Effect of Driver Age on Field Score. This figure shows a box plot comparing age of participants who were drivers during the field trial to their field score. The lack of a effect of age on search score is visible in this plot.

$F(1,16)=0.101 \ p=0.754$
6.10 Observation: Team Training Increases Intra-Team Communication

Section 6.2 showed that the data recorded in the experiment does not support the hypothesis that team training improves ad-hoc team performance in a search task; however, team training does provide some benefit to ad-hoc teams as this section will show. Figure 31 shows that there is a strong positive relationship between team training and the amount of communication between team members. This was verified with an ANOVA test that showed an effect of team training on communication with $F(1,14)=12.250$, $p<0.004$, $R^2=0.467$, Std. Err.=$0.645$, Cohen’s $f=0.566$ and $1-\beta=0.695$. This was an expected benefit of team training, but not initially hypothesized.

What is the significance of increased communication is on an ad-hoc team? As discussed in Chapter 2, Burke has demonstrated that a team of operators can have superior performance versus a single operator and communication is an important factor for a distributed team, for without communication there is effectively no team. Additionally, Fussell found that high communication frequency was an important factor in team effectiveness for other computer mediated team tasks regardless of the mediating technology [Fussell 1998]. Thus the author assumes that the increased communication frequency found in this dissertation will have some positive effect on a team, perhaps as a reduced mental workload for each teammate.
Figure 31. Communication Level vs. Team Training. This graph shows the increase in communication between team members that received team training versus those teams with members that only received individual training.
6.11 Discussion of Findings

Combined, these findings support the results of other studies showing a benefit to using video games for training. The findings that link an operator’s experience with playing video games to their team’s performance in a search task indicates game-based training or some form of video game playing is beneficial for robot operators. One implication is that there exists one or more genre of video game that teaches players the same techniques needed to successfully operate a robot, at least for a search task, and most likely for other robot operation tasks as well. It would be prudent to identify which genre of game impacts an operators performance the most and design a robot training game around it in addition to assigning other games of that type as homework for potential robot operators.

Additional evidence for the use of game-based training applications such as SARGE was gathered anecdotally from the participants following their field trial. The author asked about half (20) of the participants before they left Disaster City how they thought operating the simulated robot compared to operating the real robot. All of the participants stated that the game was a good introduction tool and was enjoyable to use. One interesting finding from this was that the participants were roughly divided on whether that the simulated robot or the real robot was more difficult to operate. From the author’s experience with robot simulators, real world vehicles are usually easier to operate than their simulated counterparts because simulated vehicle physics are usually
over responsive or more unstable than the real world. The group that believed the real robot was easier often cited that their belief was based on the fact that the real robot was less responsive than the simulated robot. The group that believed that the real robot was more difficult to operate appeared to base their belief more on the environment than the robot itself. In the simulation there was no debris on which the robot or its tether could get stuck (the simulation implemented the Inuktun VGTV as a wireless robot), nor was there any dust in the air to interfere with the camera’s autofocus. Additionally, the real robot often lost one or both tracks as the teams explored the building, but the simulated robot never experienced this issue. The discrepancies between the feel of the real robot versus the simulated robot present several areas for improvement in the simulated robot and training game.

The observation that the 20 minute experience in team training seemed to increase the amount of communication between teammates in the field teams as shown in Figure 31 is an important outcome because it shows that ad-hoc team members who have experienced a similar task in training are better at sharing information than ad-hoc teams comprised of individuals who have never been exposed to the task. High communication frequency was indicated to be an important factor in team effectiveness for other computer mediated team tasks regardless of the mediating technology [Fussell 1998]. From this standpoint there is an incentive to provide distributed team training to robot operators.
One thing to note however is that the participants expressed a significant dislike towards the manually activated radios used in the field trial, preferring instead the always-on headset used during their training. The major complaint was that they had to stop what they were doing, pick up the radio, and press a button to talk to their partner. This combined with the single duplex nature of the radios broke the their concentration. This indicates that for future field work it may be beneficial to employ a communications medium that operates similar to a conference call in which all participants are able to speak simultaneously. Even if this is approach is not taken, the use of a hands free system with standard radios is recommended.

The finding that the fractal dimension of a robot’s path does in fact indicate the skill of the operator within the game suggests that fractal D could be used as a qualifying score in future training applications. D could easily be used as an input to an intelligent tutoring system that teaches specific robot operating strategies. Each strategy would likely have an optimal execution path within an environment and this path would have an associated D value. The tutor could then use this D value as a target for the trainee to reach, adjusting the intermediate lessons to encourage the optimal operation in a specific sub-task.

Additionally the non-linear relationship of fractal dimension to performance may be indicative of the operating team’s cognitive ability if measured over time. This information could be used as a possible indicator of fatigue along with some threshold
value which would allow an incident commander to relieve a team who’s performance is deteriorating. This assumes that the position of the robot can be tracked with sufficient accuracy in the environment. In the case of unmanned aerial vehicles (UAVs) this is relatively simple using the Global Positioning System (GPS) or one of several ultra-wide band (UWB) tracking technologies. In the case of unmanned ground vehicles (UGVs) that operate in a collapsed structure these systems tend to be very inaccurate or completely unavailable, these robots will need to have sensors onboard that can localize them within the environment in order to calculate an accurate D value.

The relationships between score, age, and experience in simulation and the real world can perhaps be explained by layout of the training environments. It is possible that the layout of the real world environment and robot made it easier to conduct a proper LOVR search. The in-game search was very restrictive in that the robot was constrained to moving within a series of 1.5 meter wide hallways where as the real world environment was an open room filled with debris. Additionally, the real world environment was brightly lit versus the simulation which was generally dark and required the use of the vehicles headlights, thus it would have been easier to identify landmarks and objects of interest in the real world environment. Another factor that may have contributed to the difference in relationship between age, game experience, and score is the robot controller. The game used a PlayStation style controller which is probably more familiar to the younger participants given the relationship between age and experience with games. This controller is not labeled with functions, this required the participants to
either remember which buttons performed an action or try them all until the correct action happened. This takes time and could lead to decreased performance for the participants that had to take the trial-and-error approach. On the other hand the real robot control unit was unfamiliar to everyone, additionally the buttons are labeled with the function they perform. It is very likely that the controller was at least one factor in changing the relationship between age and score. The fidelity requirements for a robot controller used in training (similarity to the real robot controller) should be investigated in the future to determine if it is necessary to train using a mockup of the real robot controller or whether any commercially available joystick will suffice.
6.12 Changes for the Future

After analyzing the collected data, and even during the experiment, the author found several areas in which the experiment could be enhanced to provide more data which could be used to confirm or reject some of the weaker trends or strengthen the findings. The most important adjustment would be to increase the length of time over which the experiment is conducted. Over forty people participated in over a four day period, roughly ten per day. Had there been twenty people per day as planned, the experiment would have taken much longer each day as there was some lag time when swapping out groups in the field trial. Increasing the duration of the entire experiment from four days to sixteen days could give roughly 160 participants. This increase in participant count would provide significantly more power in the data analyses. This assumes that one-hundred sixty participants can be recruited; however, Disaster City’s volunteer pool lists over three-hundred individuals and there were several student groups at TAMU that were interested in participating. Providing this additional flexibility in scheduling would most likely increase the volunteer count. Additionally, many of Disaster City’s volunteers are minors (under 18 years of age). There were at least 15 minors that had initially signed up to participate in the experiment even though the recruiting advertisement clearly stated an age requirement. Minors were specifically excluded from the IRB application to reduce paperwork demands, this was clearly a mistake. Given the finding that age seems to have an effect on team score, if this study were to be rerun the author would include minors as part of the study.
The measure of fractal dimension was shown to be a predictor of operator skill and seemed to confirm that it is also an indicator of situational awareness as described by Voshell and Woods. However, the environmental designs used in training almost certainly had an effect on a participant’s mean D within each environment. The first training level that introduced the robot features was not included in the mean D value used for the analyses in Chapters 6, this level directed users to simply drive between a few waypoints in an open room while they read text descriptions of various features of the robot. This led to relatively low D values for this environment. The second level directed players to slalom between cups to pick up items artificially inducing a high D. The third level was an open room with objects scattered around which had to be collected. This level probably provided the least constrained measure of D. The final level was a series of narrow hallways which constrained the robot’s movement, this resulted in a narrow range of low D values across participants. When looking at the means across levels it appears that the individual player’s D values decrease slightly as they become familiar with the robot, which would be expected; however, the author does not believe that these results indicate learning. Instead it is believed that this effect is most likely due to the environmental design. To test whether increased familiarity with the robot decreases a player’s D value an experiment would need to have participants play in similar environments, presented in a random order to eliminate any biases that might occur based on ordering or environment layout.
While the data show that team training has a positive impact on team communication frequency, this measure could be improved to provide more useful information in future studies. The exact utterance count would enable a quantifiable analysis of the mean difference between the team training and individual training groups which would provide a better understanding of the improvement. After the first few groups finished the field trial the author noticed that the teams in the team training group tended to communicate more often than the teams in the individual training group. At that point a subjective estimate of utterance count was recorded on on paper in terms of “little”, “some”, and “a lot”. Towards the end of the experiment it was decided that a better method would be to actually count the total number of utterances made by the team members; however, at that point there were not enough groups remaining to perform any meaningful analysis on those numbers. For future studies a single video camera pointed at the operator should suffice for gathering the communication from both parties assuming that headsets are not used. If headsets are used then it may be possible to tap into the output line of the radio along side the headset and use that as input to the camera.

This dissertation showed that a robot operator’s prior experience with video games impacts their skill as a driver, but there was no information collected that identified what types of games each participant plays. A question on the survey that asked game players to rank their favorite genres of games would have provided a good clue as to what type of games should be used as training tools as would a question asking participants to indicate the number of hours played per week of each genre. Additionally,
adding gaming experience as a manipulated variable when assigning teams would have
allowed a better analysis of the relationship between age and prior gaming experience as
well as the relationship between fractal dimension and prior gaming experience and how
each affects a player’s skill. Future studies should survey participants about their gaming
habits before making decisions about team and role assignments.

Finally, during the field trial, the observers often complained about the quality of
the video available to them. The analog transmitter used operated in the open 2.4
gigahertz range. This frequency has problems penetrating the metal and concrete present
in a collapsed building. The signal degradation that resulted caused the video on the
observer’s monitor to be very noisy and often blank out for 1 to 2 seconds. This real
world problem often hampers the use of wireless devices for urban search and rescue use;
however, there are means of overcoming these issues. One solution to the video
degradation may be the use of a more costly digital video transmitter and receiver along
with directional antennas. Another solution may be to use a transmitter and receiver that
operate on a lower frequency that has better penetration capabilities.
This chapter presented the results of an experiment in which participants played a training video game and then operated a small US&R robot as part of a search task. There was participation by 41 volunteers in the experiment, of those, 37 produced usable data. There were eight findings in the collected that were data described in this chapter. Those eight findings can be combined into four overall findings. The first overall finding and most significant with respect to this dissertation is the lack of a difference in mean scores for the individually trained and team trained groups. In fact, not only is there no difference, but the means are almost identical. The second overall finding is that the fractal dimension of a robots path can be used to assess the driver’s skill. The third overall finding indicates that an operators prior experience playing video games has an effect on their team’s performance. The fourth overall finding shows that there is a positive effect of team training on the amount of communication between ad-hoc team members.

This chapter also discussed the implications of these findings. It was suggested that within existing teams or in newly formed ad-hoc teams the role of operator be assigned to members with a high degree of video gaming experience in order to maximize the effectiveness of the use of a tele-operated robot. A second suggestion made was to use fractal dimension as a measure of skill during training to provide a simple means of comparing the search strategies of trainees. Finally, the participants noted a
preference for the always-on, headset means of communication used in training versus the hand-held, hand-activated radios used in the field trial. Many of the operators thought that the radios were cumbersome to use and broke their concentration because they had to take their hands and eyes off the controller to find and use the radio.

Additionally, this chapter made several recommendations for improving the team-based training experiment in the future. The most important recommendation is to increase the length of time over which the study is run. The experiment was restricted to a four day period for this dissertation and was the primary cause for the small sample size. The small sample size in turn led to the inability to reach statistically significant conclusions for several of the relationships examined. The second recommendation relates to the measure of fractal dimension (D) and the ability to compare the evolution of a player’s D score over time. In the current experiment, the training stages were all very dissimilar in environmental layout; one was an open room, another was a series of hallways, and another directed the player to drive in a zig-zag pattern. The differences in these environments make a within subjects analysis of D impossible because of the constraints they placed on the robot’s path. For a future experiment the participants should either be forced to repeat operations in each environment several times instead of being given the option to play only once, or new levels should be designed that are similar to each other in layout. Finally, the intra-team communication frequency and the video game expertise rating, while providing useful information, could have been recorded in a way that was more meaningful. Intra-team communication should be
recorded by utterance count to allow an exact comparison of frequency instead of being grouped into three categories and the video game experience survey question should either break down gaming experience by genre or be listed along with a second question in which participants rank their favorite genres along with the number of hours per week each genre is played.
Chapter 7
Conclusions and Future Work

This dissertation has shown that there is a clear benefit to video game based training applications for robot operators that applies to both individual and team training. While the experiment failed to prove the main hypothesis that stated team training would have a significant impact on a team’s performance in a search task, the experiment did show team training has a benefit on team communication frequency. Communication frequency has been shown in the literature to be a critical factor in establishing situational awareness among teams. The second hypothesis of this dissertation stated that the fractal dimension (D) of a robot’s path could be used as a measure of the operator’s skill and was shown to be correct. The data show that D increases from a value of 1 as an operator or team is able to find more objects in a search. This increase peaks around 1.12 above which the number of objects found decrease. This decrease may indicate that the team or operator is thoroughly traversing the environment, but is unable to maintain situational awareness thus is driving the vehicle in an erratic pattern.

Additional findings from the collected data show that there is a three-way relationship between age, experience with video games, and the ability to operate a robot.
As an operator’s age increases, their ability to successfully conduct a search task within the training game decreases and as an operator’s prior experience with video games increases, their ability to operate a robot in a search task increases. However it may be that the age effect found in video game training was due to random assignment or unfamiliarity with the controller used as the field trial (which used a controller unfamiliar to everyone) data did not show the age relationship however the gaming experience relationship persisted. Thus the recommendation was made that robot operators (at least for tele-operated robots) have a high degree of prior gaming experience. Future studies will investigate if the age effect actually exists and if so whether the effect can be mitigated by gaining experience playing certain types of video games.

This dissertation provides five major contributions to the robotics, computer science, and robot user communities. The first is a contribution to all three communities. The SARGE game and robot simulator developed over the course of this work is available on-line for research, education, and training at [http://sourceforge.net/projects/sarge/](http://sourceforge.net/projects/sarge/). The use of this game was shown to have an impact on intra-team communication frequency and was viewed as a useful introductory tool to robot operation by the participants in the experiment. The second contribution is to the robotics community and is the demonstration of positive effects on team communication from the use of video games as a training tool. The third contribution is also to the robotics community and is the identification of the positive effect of prior gaming experience on performance. These two contributions suggest that the use of video games for training robot operators be
developed further. The fourth contribution is to the field of computer science and is the development of a new real-time fractal path analysis algorithm which can be used in future training applications as well as other domains in which a simple measure of path or signal tortuosity is desired in real time. Finally, the fifth contribution is to the robotics community and is the demonstration that fractal dimension can be used to assess a robot operators skill for training purposes.

There are several projects that could be considered as a continuation of this work. In this dissertation, the age minimum of the participants was 18 years. The data from this dissertation shows that age has an impact on a player’s in-game performance, it would be interesting to see how that trend continues with participants under 18 years of age and if the effect can be found in the field with a larger sample. Work by Green and Baveller [Green 2003] suggest that playing action games increases the visual selective attention of the players over time. Pairing participants who regularly play video games, including children 10-17, and comparing their results to teams consisting of non-game-players would allow us to identify if a particular age range or demographic group (game players) would be more suited to performing reach-back and robot operator tasks.

A second project would investigate how additional material can be added to the existing training game to teach specific search techniques to maximize the coverage area. During the experiment for this dissertation it was noticed that the participants that were very methodical about their search and generally followed the Localize, Observe the
surroundings, look for Victims, and Report (LOVR) [Burke 2004a] method found more items and searched more of the building. Note that the participants were never taught this strategy. Additionally, these participants also seemed to have a significant amount of prior experience playing video games. This suggests a benefit to integrating a method of teaching the LOVR strategy into the game. This work would attempt to identify a genre and game elements within that genre that lead to the player learning the LOVR strategy, then modify the SARGE training application to include a similar features.

Third, several articles which suggest that high fidelity is not always necessary for, and in some cases hinders, training. For example in an article discussing simulation fidelity concerns for the US Navy, Montague [Montague 1981] cited a *Human Factors* article [Johnson 1981] which showed that a low fidelity simulation (marking operations on photograph) produced more long-term retention of knowledge than operating a full reproduction of the real equipment. There have also been more recent suggestions that this type of abstraction would also benefit computer based training in some circumstances. Toups, et al. [Toups 2009] has shown that an abstract computer game in which a team players hunt for items in a 2D maze, guided by another player with overview information, can teach team coordination for fire fighting operations. It would be beneficial to look at various levels of simulation fidelity for robot operations to see if the effort by the community to create every detail of the real world is actually necessary.
List of References


Appendices
Appendix A: Demographic Survey

Demographic Survey

1. What is your age? ____________________

2. Please indicate your gender. (Circle only one.) [M] [F]

3. What is your occupational background? (Circle only one)
   a. Emergency Response
   b. Law Enforcement
   c. Military
   d. Student
   e. Other

4. How many years of experience do you have on the job? ____________________

5. What is your highest level of education attained? (Circle only one)
   • Some High School
   • High School Graduate
   • Some College
   • AA/AS
   • Bachelor
   • Some Graduate School
   • Masters
   • Ph.D./M.D./J.D./D.M.D/etc.
   • Postdoctorate
6. What was your major field of study?

7. Please rate your degree of skill in the following categories from 1 to 5.  
   *1 indicates very little experience while 5 indicates you are an expert.*
   
   a. Remote-controlled cars, planes, or boats.  
   
   b. Robots.  
   
   c. Video games.  
   
   d. Video cameras.  
   
   e. Search-cams or other technical rescue equipment.  
   
   f. Computers  

8. Do you have any other experience relevant to robot operation?  
   
   Y  N  

9. If you answered Yes to Question 6, please describe your experience in the box below.
10. Have you ever worked with these team members before? [ ] Y [ ] N

11. If you answered Yes to Question 8, how long have you worked with these team members? ____________________
Appendix B: IRB Approval Letter

May 21, 2009

Jeffrey David Craighead
Computer Science
ENB 342

RE: Expedited Approval for Initial Review
IRB#: 107956-I
Title: Increasing Ad-Hoc Team Performance Using Video Games
Study Approval Period: 05/20/2009 to 05/19/2010

Dear Mr. Craighead:

On May 20, 2009, Institutional Review Board (IRB) reviewed and APPROVED the above protocol for the period indicated above. It was the determination of the IRB that your study qualified for expedited review based on the federal expedited category number six (6) and seven (7).

Approval included with Waiver of Informed Consent Documentation on the Adult Minimal Risk Informed Consent Form.

Please note, if applicable, the enclosed informed consent/assent documents are valid during the period indicated by the official, IRB-Approval stamp located on page one of the form. Valid consent must be documented on a copy of the most recently IRB-approved consent form. Make copies from the enclosed original.

Please reference the above IRB protocol number in all correspondence regarding this protocol with the IRB or the Division of Research Integrity and Compliance. In addition, you can find the Institutional Review Board (IRB) Quick Reference Guide providing guidelines and resources to assist you in meeting your responsibilities in the conduction of human participant research on our website. Please read this guide carefully. It is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB.
We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-2036.

Sincerely,

Krista Kutash, Ph.D., Chairperson
USF Institutional Review Board

Cc: Various Menzel/cd, USF IRB Professional Staff
    William Kearns, PhD
About the Author

Jeffrey David Craighead graduated received his Bachelor of Science in Computer Science from the University of South Florida in 2004 and received his Master of Science in Computer Science from the University of South Florida in 2008. Jeff has participated in several urban search and rescue responses and training events. He has numerous publications in the areas of simulation, intelligent sensor systems, and mobile robotics. He currently lives in Temple Terrace with his wife, son, and English bulldog.