

# Human-Automation Task Allocation in Lunar Landing: Simulation and Experiments

by

Hui Ying Wen

B.S. Aerospace Engineering with Information Technology  
B.S. Humanities

Massachusetts Institute of Technology, 2008

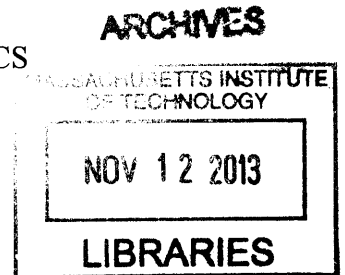
SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS IN  
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN AERONAUTICS AND ASTRONAUTICS  
AT THE  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2011

© 2011 Hui Ying Wen

The author hereby grants to MIT and Draper Laboratory permission to reproduce and to  
distribute publicly paper and electronic copies of this thesis document in whole or in part.



Signature of Author: \_\_\_\_\_  
Department of Aeronautics and Astronautics  
August 18, 2011

Certified by: \_\_\_\_\_  
Charles M. Oman  
Director, Man Vehicle Laboratory  
Senior Research Engineer  
Senior Lecturer

Certified by: \_\_\_\_\_  
Kevin R. Duda  
Senior Member of the Technical Staff  
The Charles Stark Draper Laboratory, Inc.

Accepted by: \_\_\_\_\_  
Eytan H. Modiano  
Professor of Aeronautics and Astronautics  
Chair, Committee on Graduate Studies



## Abstract

Task allocation, or how tasks are assigned to the human operator(s) versus to automation, is an important aspect of designing a complex vehicle or system for use in human space exploration. The performance implications of alternative task allocations between human and automation can be simulated, allowing high-level analysis of a large task allocation design space. Human subject experiments can then be conducted to uncover human behaviors not modeled in simulation but need to be considered in making the task allocation decision. These methods were applied here to the case scenario of lunar landing with a single human pilot.

A task analysis was performed on a hypothetical generic lunar landing mission, focusing on decisions and actions that could be assigned to the pilot or to automation during the braking, approach, and touchdown phases. Models of human and automation task completion behavior were implemented within a closed-loop pilot-vehicle simulation for three subtasks within the landing point designation (LPD) and final approach tasks, creating a simulation framework tailored for the analysis of a task allocation design space. Results from 160 simulation runs showed that system performance, measured by fuel usage and landing accuracy, was predicted to be optimized if the human performs decision making tasks, and manual tasks such as flying the vehicle are automated. Variability in fuel use can be attributed to human performance of the flying task. Variability in landing accuracy appears to result from human performance of the LPD and flying tasks.

Next, a human subject experiment (11 subjects, 68 trials per subject) was conducted to study subjects' risk-taking strategy in designating the landing point. Results showed that subjects designated landing points that compensated for estimated touchdown dispersions and system-level knowledge of the probabilities of manual versus automated flight. Also, subjects made more complete LPD compensations when estimating touchdown dispersion from graphical plots rather than from memories of previous simulated landings. The way in which dispersion information is presented affected the consistency with which subjects adhered to a risk level in making landing point selections. These effects could then be incorporated in future human performance models and task allocation simulations.

Thesis Supervisor: Charles M. Oman  
Title: Senior Lecturer, Department of Aeronautics and Astronautics,  
Massachusetts Institute of Technology

Thesis Supervisor: Kevin R. Duda  
Title: Senior Member of the Technical Staff, Draper Laboratory

## Acknowledgments

I would like to thank Charles “Chuck” Oman for inviting me to join this project. His inspired and gentle guidance, born of his formidable knowledge of the field, gave this body of work its contours and nuances. Kevin Duda’s mentorship at Draper Laboratory was indispensable: without his vigilant attention to the innumerable scientific, technical, and managerial details of conducting this project, this thesis would not have been possible. Alan Natapoff gave much of his time and patience in helping with data analysis and experiment design, in the process enormously expanding my knowledge of statistics and infecting me with a bit of his passion for the field.

To Linda Fuhrman and the Draper Education Office, I am grateful for their capable and attentive leadership of the Draper Fellowship program. I would also like to thank Justin Vican (Draper Lab) for his technical support with the Draper lunar landing cockpit simulator. Numerous volunteers at Draper Laboratory gave their time as subjects in the human subject experiment.

C. J. Hainley and Alexander Stimpson implemented many of the simulation models and graphics that were used in this thesis project. Justin Kaderka, a fellow member of this project, gave his help and insight into many aspects of the work. I would also like to thank the NASA Education Associates Program who hosted my stay at the Ames Research Center, and Jessica Marquez who served as my mentor and sponsor during my time there.

The students of the Man Vehicle Laboratory (MVL) kept me uplifted with friendship, fun, and moral support. Also, no matter what sundry tasks and requests we brought to the door of Liz Zotos, the MVL’s administrative assistant, she has always handled them with an enthusiasm I have never seen flag.

Finally, I would like to thank the MIT Theater Arts Department, and its astonishing faculty and students, for cultivating my artistic life alongside my engineering one. Last but certainly not least, I thank my father, Yat Sun Wen, and mother, Jian Li Wen, for helping me get to this point in life – it is already beyond my wildest childhood dreams. This thesis is dedicated to the memory of my cousin Victor.

This work was supported by the National Space Biomedical Research Institute through NASA NCC 9-58, Project HFP02001. It was performed at The Charles Stark Draper Laboratory, Inc., and any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Aeronautics and Space Administration.



# Table of Contents

- 1 Introduction ..... 9
  - 1.1 Motivation and Problem Statement ..... 9
  - 1.2 Research Aims and Contributions ..... 9
- 2 Lunar Landing Task Analysis ..... 10
- 3 Evaluation of Human-Automation Task Allocations in Lunar Landing Using Task Network Simulation of Human and Automation Performance ..... 14
  - 3.1 Introduction ..... 14
  - 3.2 Background ..... 14
    - 3.2.1 Task Allocation Methods ..... 14
    - 3.2.2 Human Performance Modeling ..... 16
    - 3.2.3 Human Information Processing ..... 17
  - 3.3 Method ..... 18
    - 3.3.1 Modeled Tasks ..... 18
    - 3.3.2 Simulation Implementation ..... 19
    - 3.3.3 Simulation Procedure ..... 25
  - 3.4 Results ..... 27
  - 3.5 Discussion ..... 31
  - 3.6 Conclusion ..... 31
- 4 Effect of Perceived Risk on Lunar Landing Point Selection ..... 32
  - 4.1 Introduction ..... 32
    - 4.1.1 Motivation ..... 32
    - 4.1.2 Problem Statement ..... 33
  - 4.2 Background ..... 34
    - 4.2.1 Memory Factors ..... 34
    - 4.2.2 Biases in Perception of Risk and Reward ..... 35
  - 4.3 Methods ..... 36
    - 4.3.1 Independent Variables and Trial Schedule ..... 36
    - 4.3.2 Hypotheses ..... 38
    - 4.3.3 Scenario Design ..... 39
    - 4.3.4 Equipment and Displays ..... 41

4.3.5	Subjects .....	44
4.3.6	Measurement Collection and Data Calculation .....	44
4.4	Results.....	47
4.4.1	Overview of Flying and Landing Performance in Part 1.....	47
4.4.2	Landing Aim Point $t$ 's .....	48
4.4.3	Subjective Responses.....	53
4.5	Conclusion and Discussion.....	54
5	Conclusion.....	55
6	Future Work.....	55
6.1	Future Work for Task Allocation Simulation.....	56
6.2	Future Work for Human Subject Experiment .....	56
7	References .....	57
Appendix 1	Preliminary Lunar Landing Task Analysis .....	61
Appendix 2	Ranges for Human Model Parameters.....	63
Appendix 3	Subject Training Slides .....	64
Appendix 4	Schedule of Trials.....	67
Appendix 5	Verbal instructions for scatter plot portion of experiment .....	69
Appendix 6	Verbal questions between parts of experiment .....	70
Appendix 7	COUHES Forms.....	71
Appendix 8	Landings in Experiment that Violated Safety Limits at Contact.....	73
Appendix 9	False Positives in Human Experiment Results.....	73
Appendix 10	Subjects' Self Risk Ratings.....	74

## List of Figures

FIGURE 1: DIAGRAM OF APOLLO LUNAR LANDING PHASES. P63 THROUGH P67 REFER TO APOLLO GUIDANCE COMPUTER SOFTWARE PROGRAMS GOVERNING SPECIFIC LANDING PHASES. FROM (DUDA, JOHNSON, & FILL, 2009).....	10
FIGURE 2: TOP LEVEL OF HIERARCHICAL TASK ANALYSIS FOR LUNAR LANDING. LIGHT BLUE BOXES WITH SOLID BORDERS INDICATE SUBGOALS THAT CAN BE FURTHER RE-DESCRIBED INTO LOWER-LEVEL GOALS. CLEAR BOXES WITH NON-SOLID BORDERS INDICATE TASKS AT THE LOWEST LEVEL OF RE-DESCRIPTION.....	12
FIGURE 3: HIERARCHICAL TASK ANALYSIS FOR THE BRAKING PHASE OF LUNAR LANDING .....	12
FIGURE 4: HIERARCHICAL TASK ANALYSIS FOR THE APPROACH PHASE OF LUNAR LANDING .....	13
FIGURE 5: HIERARCHICAL TASK ANALYSIS FOR THE TERMINAL DESCENT PHASE OF LUNAR LANDING .....	13
FIGURE 6: FOUR-STAGE MODEL OF HUMAN INFORMATION PROCESSING .....	17

FIGURE 7: TASK NETWORK, CONTAINING LPD AND TOUCHDOWN SUBTASKS, OF THE CASE EXAMPLE TO BE MODELED. .... 19

FIGURE 8: ARCHITECTURE OF TASK ALLOCATION SIMULATION FRAMEWORK..... 20

FIGURE 9: HUMAN PERCEPTION AND DECISION MODELS FOR SUBTASK 1 (DECIDE WHETHER OR NOT TO USE AUTOMATED SYSTEM IN LPD) ..... 22

FIGURE 10: HUMAN PERCEPTION AND DECISION MODELS FOR SUBTASK 2 (SELECT LAP SUGGESTED BY AUTOMATED SYSTEM) ..... 22

FIGURE 11: HUMAN PERCEPTION AND DECISION MODELS FOR SUBTASK 3 (SELECT LAP USING OTW VIEW AND AUTOMATED SYSTEM’S SCAN OF LANDING AREA) ..... 23

FIGURE 12: HUMAN PERCEPTION AND DECISION MODELS FOR SUBTASK 4 (FLY TO SELECTED LAP) ..... 23

FIGURE 13: CLOSED FEEDBACK CONTROL LOOPS MODELING HUMAN MANUAL CONTROL OF VEHICLE MOVEMENT VERTICALLY (TOP FIGURE) AND HORIZONTALLY (BOTTOM FIGURE) ..... 24

FIGURE 14: DEM OF 591 FT. X 591 FT. (180X180M) TERRAIN AREA USED IN SIMULATION SCENARIO. THE SCENARIO INCLUDES 3 RANKED LAPs SUGGESTED BY AUTOMATED SYSTEM (“1” = MOST HIGHLY SUGGESTED), A POI, AND THE STARTING LOCATION OF THE VEHICLE..... 26

FIGURE 15: SUMMARY OF PARAMETERS IN HUMAN PERFORMANCE MODELS WHICH WERE VARIED FROM ONE SIMULATION RUN TO THE NEXT ..... 27

FIGURE 16: DEM OF TERRAIN WITH LAPs SUGGESTED BY AUTOMATED SYSTEM (SQUARES), A POI (GREEN STAR), AND LAPs INDEPENDENTLY SELECTED BY HUMAN MODEL (BLUE ASTERISKS) OVER ALL SIMULATION RUNS ..... 28

FIGURE 17: SIMULATED PERFORMANCE OF EACH TASK ALLOCATION, BASED ON MEAN REMAINING FUEL AND RANGE FROM SELECTED LAP AT MOMENT OF TOUCHDOWN. EACH ALLOCATION IS DENOTED BY 4 DIGITS, WITH EACH DIGIT SIGNIFYING THE ALLOCATION OF A SUBTASK: “<SUBTASK 1> <SUBTASK 2> <SUBTASK 3> <SUBTASK 4>.” “1” = ASSIGNED TO HUMAN, “0” = ASSIGNED TO AUTOMATION. ERROR BARS INDICATE STANDARD ERROR IN THE MEAN.. 28

FIGURE 18: PLOT HIGHLIGHTING STANDARD ERROR IN THE MEAN OF REMAINING FUEL AMOUNTS FOR EACH TASK ALLOCATION (SEE FIGURE 17 CAPTION FOR EXPLANATION OF ALLOCATION NOTATION) ..... 29

FIGURE 19: PLOT HIGHLIGHTING STANDARD ERROR IN THE MEAN OF FINAL RANGE FROM SELECTED LAP FOR EACH TASK ALLOCATION (SEE FIGURE 17 CAPTION FOR EXPLANATION OF ALLOCATION NOTATION) ..... 30

FIGURE 20: TRENDS OF RELATIVE VARIABILITY IN LANDING ACCURACY FOR DIFFERENT TASK ALLOCATIONS (SEE FIGURE 17 CAPTION FOR EXPLANATION OF ALLOCATION NOTATION). NOTE: IF SUBTASK 1, “DECIDE WHETHER OR NOT TO USE AUTOMATED SYSTEM IN LPD,” IS ALLOCATED TO AUTOMATION, THE AUTOMATION WILL ALWAYS DECIDE “YES.” ..... 31

FIGURE 21: SUBSET OF TASKS MODELED IN SIMULATION THAT WAS SELECTED FOR HUMAN SUBJECT EXPERIMENTATION. SUBTASK 3 IS IMPLEMENTED IN THE EXPERIMENT WITHOUT THE USE OF AN OTW VIEW (I.E., ONLY AUTOMATED SYSTEM SCAN). SUBTASKS 1 AND 2 (GRAYED OUT) WERE NOT USED IN THE EXPERIMENT..... 33

FIGURE 22: SUBJECTIVE VERSUS TRUE PROBABILITY AND UTILITY (KAHNEMAN & TVERSKY, 1984) ..... 36

FIGURE 23: MAPS OF LANDING AREAS USED IN HUMAN SUBJECT EXPERIMENT. SHORT MAP AT LEFT, LONG MAP AT RIGHT. .... 39

FIGURE 24: ADDITIONAL MAPS INCLUDED IN TRAINING SESSION. NORTHERN POI MAP AT LEFT, SOUTHERN POI MAP AT RIGHT. .... 40

FIGURE 25: EXPERIMENTER SEATED IN LUNAR LANDING SIMULATOR. .... 41

FIGURE 26: DISPLAYS FOR TRAINING AND PART 1 OF EXPERIMENT. FROM LEFT TO RIGHT: FLIGHT DISPLAY, LANDING AREA MAP, AND LANDING RATING SCREEN. .... 42

FIGURE 27: DISPLAYS FOR PART 2 OF EXPERIMENT. FROM LEFT TO RIGHT: SCATTER PLOT OF SYNTHESIZED LANDING ERRORS AND LANDING AREA MAP. .... 43

FIGURE 28: SCATTER PLOTS OF SYNTHESIZED LANDING ERRORS. BLACK SQUARES REPRESENT WHERE LANDINGS ARE LOCATED RELATIVE TO A TARGET LAP (MAGENTA SQUARE WITH CROSSHAIRS). FROM LEFT TO RIGHT: MEAN LANDING DEVIATION = -2, 0, AND 2 VEHICLE WIDTHS IN BOTH NORTH-SOUTH AND EAST-WEST DIRECTIONS. TOP ROW: SD OF LANDING DEVIATIONS = 1 VEHICLE WIDTH. BOTTOM ROW: SD OF LANDING DEVIATIONS = 3 VEHICLE WIDTHS. .... 44

FIGURE 29: LAP LOCATION (CROSSED MAGENTA SQUARE) AND LANDING ERROR (BLUE VECTOR) ..... 45

FIGURE 30: RISK  $T$  DERIVED FROM THE MEAN OF A SET OF LANDINGS ON A SINGLE MAP ..... 46

FIGURE 31: MEAN LANDING ERRORS FOR EACH SUBJECT. ERROR BARS INDICATE SD IN EW AND NS DIRECTIONS. UNITS = VEHICLE LENGTHS. EACH DATA POINT CONTAINS LANDINGS FOR BOTH LANDING AREA MAPS, AS NO SIGNIFICANT EFFECT OF MAP WAS FOUND ON LANDING ERROR. .... 48

FIGURE 32: EFFECT OF FLYING TASK CONDITION ON  $T$  PLOTTED AGAINST SUBJECTS' MEAN LANDING ERRORS IN PART 1. A POSITIVE EFFECT (DIFFERENCE IN  $T$ ) MEANS THAT LAPs SELECTED FOR THE 25% MANUAL CONDITION WERE SAFER THAN THOSE SELECTED FOR THE 100% MANUAL CONDITION. EACH DATA POINT CONTAINS 10 LAP SELECTIONS (5 FOR EACH ALLOCATION) BY ONE SUBJECT FOR ONE MAP TYPE. .... 50

FIGURE 33: THE EFFECT ON THE T-RISK PARAMETER OF FLYING TASK CONDITION AND MEAN SYNTHESIZED LANDING ERROR IN PART 2. A POSITIVE EFFECT (DIFFERENCE IN  $T$  VALUES) MEANS THAT LAPs SELECTED FOR THE 25% MANUAL FLYING TASK CONDITION WERE SAFER THAN THOSE SELECTED FOR THE 100% MANUAL CONDITION. IN THE TWO UPPER CHARTS, EACH DATA POINT CONTAINS 8 LAP SELECTIONS (4 FOR EACH FLYING TASK CONDITION) BY ONE SUBJECT FOR ONE MAP TYPE. THE LOWER TWO CHARTS CONTAIN BOX PLOTS OF THE SAME DATA. \*\*=THE EFFECT OF FLYING TASK CONDITION (THE DIFFERENCE IN  $T$  VALUES) IS SIGNIFICANT ( $p < 0.05$ ) FOR BOTH MAPS. .... 51

FIGURE 34: VARIANCE IN  $T$ 'S OF TARGET SELECTIONS VS. LANDING ERROR INFORMATION. HORIZONTAL BARS SHOW SIGNIFICANT DIFFERENCE BY PAIRED T-TEST,  $p < 0.05$ . EACH BOX PLOT CONTAINS 11 DATA POINTS, ONE FOR EACH SUBJECT (20 DATA POINTS PER SUBJECT FOR PART 1, 24 DATA POINTS PER SUBJECT FOR EACH PART 2 BOX PLOT). NOTE THAT THE EW PLOT EXCLUDES AN OUTLIER SUBJECT FOR THE SHORT MAP. \*=VARIANCES ARE SIGNIFICANTLY DIFFERENT ( $p < 0.05$ ) ONLY FOR THE LONG MAP. \*\*=VARIANCES ARE SIGNIFICANTLY DIFFERENT ( $p < 0.05$ ) FOR BOTH MAPS..... 52

## List of Tables

TABLE 1: TASK ALLOCATIONS FOR SIMULATION ..... 26

TABLE 2: NUMBER OF TRIALS IN EACH PART OF EXPERIMENT ..... 38

TABLE 3: FLYING AND LANDING PERFORMANCE OVER ALL SUBJECTS FOR PART 1 ..... 47

# 1 Introduction

## 1.1 Motivation and Problem Statement

In designing a system in which the human is an integral component, one essential question is, “who does what?” What tasks or functions should be performed by automation, and what tasks are best left in the domain of a human operator? Instead, the answer is often determined by project constraints on budget, available technology, desires of the human operators, and current societal attitudes towards automation. This is especially the case in the resource-intensive, technologically risky, and high-stakes world of space exploration (Mindell, 2008). There is also the tendency to do things as they have been done in the past – not only is change expensive, but a system design that has worked sufficiently well in the past, and the human operators that were trained on it, carry their own resistance to change.

The problem addressed in this thesis is the development of quantitative methods for human-automation task allocation. The case scenario used in this project is a lunar landing system. The complexity of such a system allows many decisions to be made on how the human crew and automation interact, and what tasks are allocated to each. For example, who decides where to land? Who flies the vehicle? The case scenario of lunar landing is directly applicable to landing systems for other celestial bodies such as the Earth, Mars, or asteroids. Analogous human vs. automation task allocation decisions also arise during the design of other complex systems involving humans and automation: aircraft and air traffic control, ships, trains, nuclear power plants, and patient monitoring systems (Sheridan, 2002).

## 1.2 Research Aims and Contributions

An objective way to decide the human-automation task allocation question is by running experiments with different allocations; however, this approach is often too costly to use to explore any given system’s full design space. An alternative is to create simulations of such experiments. Modeling how a system would respond to human operators and automation performing a given set of tasks can, in principle, generate high-level evaluations of many different task allocations. Such results can be used to suggest directions in which to narrow the system design space and, further along the development process, indicate areas requiring special attention in the testing and validation of chosen designs.

Early methods for evaluating task allocation relied on guiding principles. This project is an application of the latest trend in task allocation, which is system-level simulation of human performance in collaboration with automation and the surrounding environment. An overview of tools developed over the past few decades to aid in determining task allocation is provided in Section 3.2.1.

Three main bodies of work are presented: 1) task analysis of the chosen case scenario, lunar landing, 2) creation of a computer-based task allocation simulation for selected lunar landing tasks, and 3) experimental evaluation of human performance on the same modeled tasks. Task analysis is required to determine the high level tasks, states, and goals before a scenario can be modeled in simulation. The modeling work demonstrates the advantages of simulation in evaluating the optimality of varying task allocations for a given system and scenario. Finally, experimentation provides empirical data on human performance that may be used to partially validate and enrich the simulation.



## 2 Lunar Landing Task Analysis\*

Lunar landing is the case scenario used in this project, although the methods of simulation and experimentation used here can be applied to exploring human-automation task allocations for any other case scenario that includes a complex system with a human component.

In the Apollo missions, lunar landing was divided into three phases: braking, approach, and terminal descent, as shown in Figure 1 (Bennett, 1972).

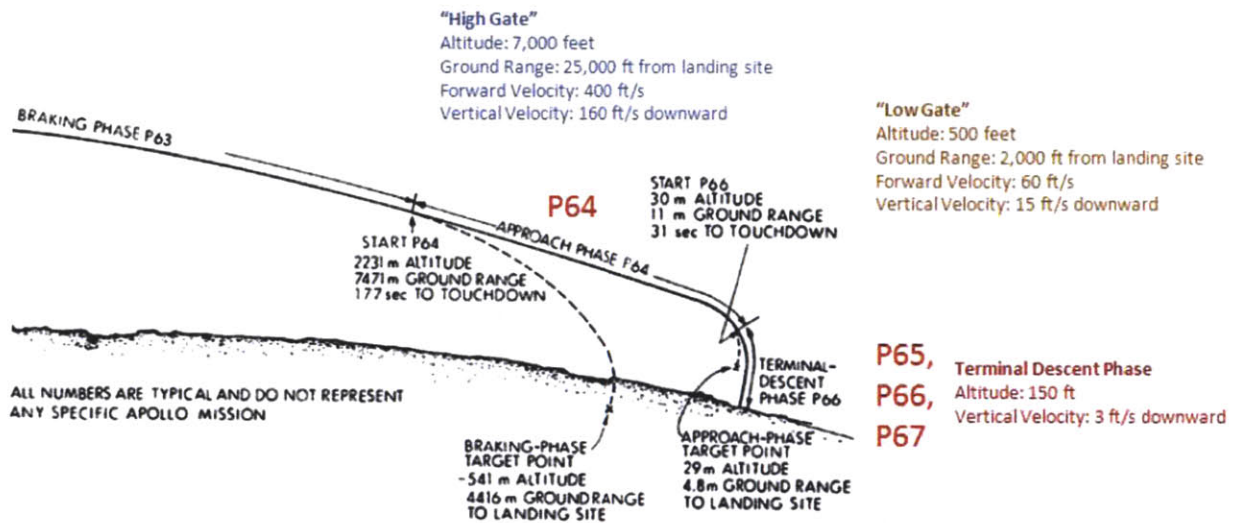


Figure 1: Diagram of Apollo lunar landing phases. P63 through P67 refer to Apollo Guidance Computer software programs governing specific landing phases. From (Duda, Johnson, & Fill, 2009).

The purpose of the braking phase was to bring the Lunar Module (LM) spacecraft down from a lunar orbit to a set of guidance target conditions known as "High Gate". Then, in the approach phase, the LM pitched nearly upright so that the astronauts had a view of the lunar surface and could designate a final landing aim point (LAP). The vehicle was then navigated close to the landing site to meet a set of guidance target conditions known as "Low Gate" (Klumpp, 1974). In the terminal descent phase, the vehicle was navigated to a safe area that appeared level and free of hazards, and its horizontal and vertical velocities were nulled to bring it to a touchdown on a chosen landing site.

Before exploring task allocations, an analysis was performed of the tasks required in piloted lunar landing, from the beginning of the braking phase to the end of terminal descent. Normally, a task analysis would be based on the specifications of a given system and mission under design. However, the goal of this analysis was to produce a profile of basic command and control tasks that would likely be common to any piloted lunar landing mission.

The task profile documented for the Apollo LM (Mindell, 2008; NASA, 1971) was used as the baseline for this task analysis. Tasks and portions of tasks that were deemed specific to the Apollo landing system were removed. To reflect technology that would likely be included in a modern landing system, tasks

\* Adapted from (Wen, Duda, Slesnick, & Oman, 2011)

were included for the use of an automated landing and hazard avoidance that scans the terrain, detects hazards, and identifies and prioritizes possible LAPs, allowing the vehicle to land even when visibility is poor for a human pilot (Epp, Robertson, & Brady, 2008; Forest, Kessler, & Homer, 2007). A cockpit system that allows human crew to perform supervisory actions, such as selecting a LAP from among those recommended by the automated hazard avoidance system, was also assumed (Forest et al., 2007). Finally, Apollo's rate-control attitude hold (RCAH) mode, in which the pilot had direct control of the vehicle's attitude and descent rate, was assumed as the mode of manual flight control (Hackler, Brickel, Smith, & Cheatham, 1968). A preliminary version of this task analysis, showing the separate sources drawn from Apollo documentation and modern landing technology, can be found in Appendix 1.

In addition, this task analysis had to produce a description of tasks at a level of detail appropriate for the modeling in this project. In the absence of specifications for a physical system, low-level human perceptual-motor primitives (originally defined by Jenkins, Mataric, and Weber (Jenkins, Mataric, & Weber, 2000)) such as interactions with a cockpit display and button presses were omitted. The remaining task profile consists of high-level decision making tasks and tasks required by the dynamics of a lunar landing, regardless of the specific landing system. Hierarchical task analysis (HTA) was found to be particularly suitable for this purpose (Stanton, 2006). HTA starts with identifying the main goal that the system is meant to accomplish, and that goal is then "re-described" into a tree of sub-goals that are necessary to accomplish the parent goal. The re-description of sub-goals continues until, at the lowest leaves of the tree, tasks are obtained at the desired level of detail. At each node in this tree of goal and sub-goal re-descriptions, a "plan" describes the temporal ordering of sub-goals and the completion criteria for the parent goal.

The tasks in the resulting task network are identified as one of three types, to guide the human performance modeling of these tasks in a simulation. These three types describe typical piloted spacecraft command and control tasks:

Navigation Task: Involves monitoring and changing the current dynamic state (position, velocity) of the vehicle. Note: although navigation conventionally means planning the trajectory and determining vehicle position relative to the planned trajectory, the term is used here to include the flying portion of the task.

Subsystem Supervisory Task: Involves monitoring an automated function or vehicle state, and taking action to cause a change in the system if necessary.

Decision Making Task: Involves utilizing information presented by automation, or from an out-the-window view, to make decisions that affect the mission trajectory at a high level.

The root goal and highest level of HTA for the case of lunar landing is shown in Figure 2. The main goal is to land at a desirable LAP while avoiding hazardous areas. In this case scenario, the LAP may be selected by an automated system or manually designated to be as close to as possible to a point of interest (POI) near which it is desirable to land.

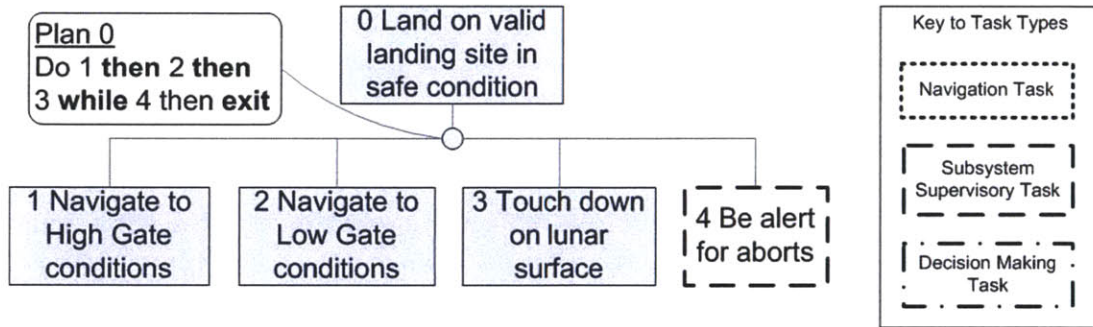


Figure 2: Top level of hierarchical task analysis for lunar landing. Light blue boxes with solid borders indicate subgoals that can be further re-described into lower-level goals. Clear boxes with non-solid borders indicate tasks at the lowest level of re-description.

The root goal can be accomplished only if navigation to High Gate conditions (subgoal 1), then Low Gate conditions (subgoal 2), and finally to a landing site (subgoal 3) is accomplished. Plan 0 specifies that while subgoals 1-3 are being accomplished in sequence, there is also a parallel task of monitoring for situations that call for a landing abort (subgoal 4). More detailed descriptions of subgoals 1-3 are shown below in Figure 3 through Figure 5.

The main result of subgoal 1 (

Figure 3) is to perform a descent-engine burn to decelerate the vehicle from lunar orbit to High Gate conditions. Meanwhile, radar data of the lunar surface below is received for the first time. Subgoal 2 (Figure 4) involves the simultaneous tasks of approaching the final landing area and designating a LAP that is safe for landing within the landing area. Subgoal 3 (Figure 5) consists of the vehicle's final descent onto the selected LAP.

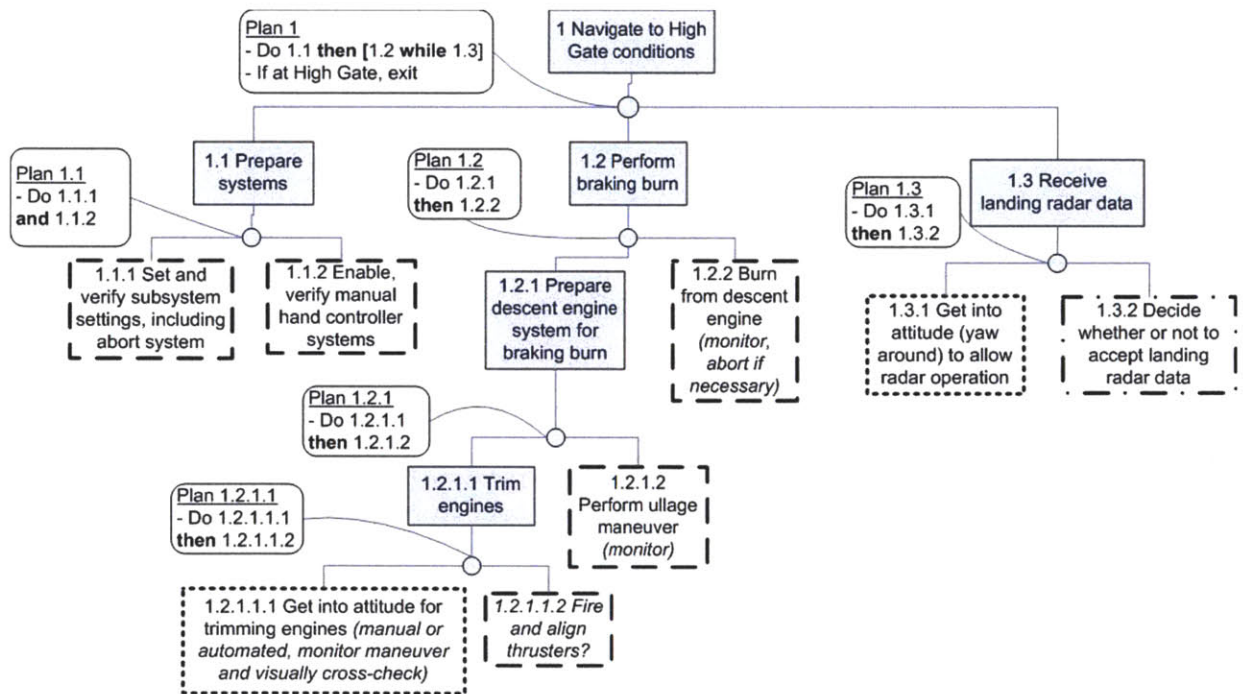




Figure 3: Hierarchical task analysis for the braking phase of lunar landing

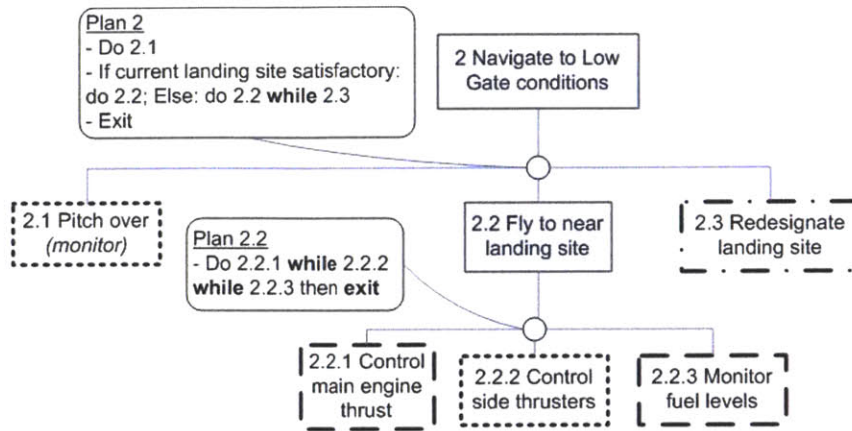


Figure 4: Hierarchical task analysis for the approach phase of lunar landing

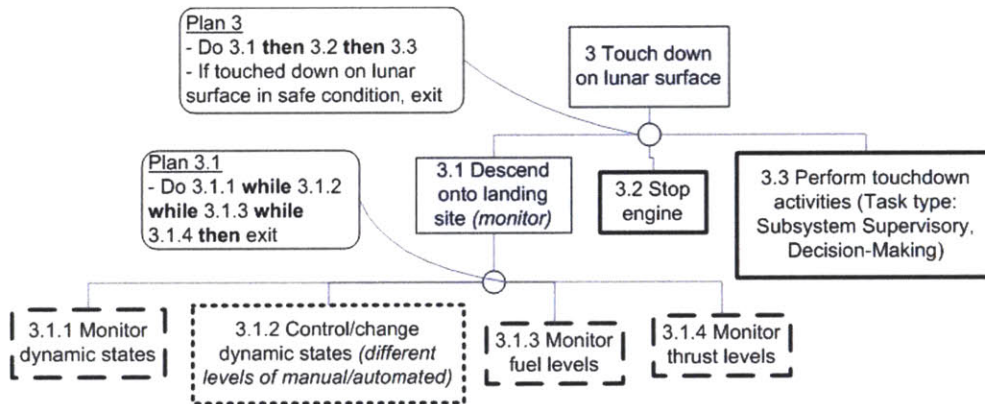


Figure 5: Hierarchical task analysis for the terminal descent phase of lunar landing

Identifying the tasks to be performed in a system is the first step in evaluating task allocations. The task analysis shown here was used as the master plan from which tasks were selected for modeling in a task allocation simulation and for human subject experimentation, as well as to keep track of the context in which the selected tasks are performed. Since this analysis contains tasks that should be common to many piloted lunar landing systems, it is an applicable starting point from which to analyze the task allocation of any specific future designs.

## **3 Evaluation of Human-Automation Task Allocations in Lunar Landing Using Task Network Simulation of Human and Automation Performance\***

### **3.1 Introduction**

The purpose of simulating different human-automation task allocations is to obtain an early understanding of the full human-automation task allocation design trade space, which will guide the overall system design process. Early in the design process, it may be premature or too costly to evaluate different system design options using other methods such as building mock-ups and running human subject experiments. Therefore, simulation can be a useful early-stage design tool that provides preliminary, high-level evaluations of all possible task allocation designs. Further along the design process, simulation can be used to discover areas requiring special attention when testing and validating the chosen system designs. The lunar landing simulation example considered here is intended to demonstrate the feasibility of this approach.

Although a simulation which contains detailed models of human performance and is used to evaluate human-automation interactions does exist (see the description of the Air Man-machine Integration Design and Analysis System (Air MIDAS) in Section 3.2.2.2), its complexity and high level of detail may make results difficult to trace and understand. Pritchett provides the following advice: “When critical design decisions are to be made based on their results, a coarser or sparser human performance model representing simple, well-understood phenomena may be more useful than a more detailed model based on tentative models of behavior” (Byrne et al., 2008). Therefore, another motivation for this project is to build a more transparent simulation containing simple, preliminary models of human behavior.

### **3.2 Background**

#### **3.2.1 Task Allocation Methods**

The tools used to tackle the question of human-automation task allocation has evolved through the past decades from general guiding principles and rules of thumb (Fitts, 1962; Jordan, 1963) to more concrete quantitative methodologies that prescribe processes by which to answer the question (Marsden & Kirby, 2005; Parasuraman, Sheridan, & Wickens, 2000; Sheridan & Verplank, 1978) and finally to analytical models that simulate and evaluate different task allocations (Chua & Major, 2009; Connelly & Willis, 1969; Madni, 1988; Sheridan & Parasuraman, 2000).

Early guiding principles came in the 1960's. In general, machines excel at performing any given task with consistency, while human operators can provide flexibility in response to dynamic task conditions (Jordan, 1963). Specifically, Fitts pointed out that a human operator can trade off speed for precision as needed: if a task needs to be done more quickly, a human can deliver speed by sacrificing precision in the execution of the task, and vice versa (Fitts, 1962). A machine is not likely to exhibit that flexibility. On the other hand, Jordan warned against falling into the obvious trap of comparing human and

---

\* Adapted from (Wen et al., 2011)

machine performance by the same criteria (Jordan, 1963). Most of the time, this leads to the deduction that if a task can be modeled, a machine can always be built to do it in a superior fashion, which in turn leads to the (somewhat unhelpful) conclusion that there is no role for a human in the system. As Jordan put it, “men and machines are not comparable, they are *complementary*” (Jordan, 1963).

However, guidelines alone are not sufficient to bring the problem of task allocation from the realm of art into that of science. The original and simplest quantitative methodology is what Marsden and Kirby called “tables of relative merit” (Marsden & Kirby, 2005). The most well-known example is Fitts’ table of general task types for which humans and machines are best suited to perform (Chapanis et al., 1951). Although these guidelines make logical sense, it is difficult to apply them in practice. Real tasks, performed in disparate environments, often defy categorization into one of the types described by the list.

Another lens through which to view task allocation is through levels of automation. Sheridan and Verplank identified a spectrum of ten levels for characterizing the degree to which the decision making in a task is performed by automation versus by the human operator (Sheridan & Verplank, 1978). Parasuraman, Sheridan, and Wickens took this further by breaking a task down into four information processing stages – sensory processing, perception / working memory, decision making, and response selection – and applied the ten levels to each of the stages (Parasuraman et al., 2000). Thus, the degree to which a task, and the separate information processing components of the task, was performed by a human or the automation can be characterized according to a scale.

Finally, analytical models and simulations were introduced that can potentially search a large task allocation design space and provide wide-ranging characterizations of system performance at a high level. The simplest models take the form of databases that do not model how tasks are performed but store knowledge that affects the choice of a task allocation, such as human task performance parameters and the technical costs, benefits, and feasibility of implementing automation to perform tasks (Connelly & Willis, 1969; Madni, 1988). Although these databases help organize information needed to make task allocation decisions, they do not contain descriptive models of human and system behavior or answer how system performance is impacted by task allocation. Other models take the form of algorithmic descriptions, such as expected-value calculations to decide whether a task that involves failure detection should be allocated to a human or automation (Sheridan & Parasuraman, 2000) and task time analyses (Chua & Major, 2009). Such algorithmic descriptions, however, are not descriptive behavioral models and can provide only rough evaluations of task allocations according to one metric (such as failure detection probability or task time).

The latest genre of task allocation models attempt to simulate the internal behavior of individual elements within a system, such as the human operator, any automated components, and the environment in which tasks are performed. The Function Allocation Methods (FAME) tool models the human operator as using a cyclical information processing model that continually guides future actions. The model also includes an automation component and allows for easy transfer of tasks between the human or automation models to evaluate different task allocations (Hollnagel & Bye, 2000). The task allocation simulation work in this project continues this tradition of modeling.



### **3.2.2 Human Performance Modeling**

NASA conducted a Human Performance Modeling (HPM) project over the course of 6 years, ending in 2008, in which five different human cognitive modeling efforts were applied to analyze system design impacts and human error in aviation (Byrne et al., 2008; Foyle et al., 2005). Ideas and lessons learned were drawn from these efforts to guide the simulation work in this project. This HPM project was used for inspiration because it too uses a piloted flight system as the case scenario, although for aviation and not planetary landing. Also, it is a convenient collection of diverse architectures and modeling approaches applied to the same type of system. In particular, lessons learned about architecture and level of detail in modeling were drawn from the Improved Performance Research Integration Tool / Adaptive Control of Thought-Rational (IMPRINT/ACT-R), Air Man-machine Integration Design and Analysis System (Air MIDAS), and Attention-Situation Awareness (A-SA).

#### **3.2.2.1 Architecture**

There are two human performance modeling architectures: reductionist and first-principle modeling (Laughery, Lebiere, & Archer, 2006). The former uses a task network drawn from a task analysis as its overarching structure. The core of the latter architecture is a model of human cognition; all other components of the system, such as a task network or parts of the system external to the human operator, are treated as peripheral interfaces. The organizing features of both architectures – the task network and the human cognitive model – are equally important to the task allocation simulation in this project.

One of the HPM efforts, “IMPRINT/ACT-R” (Foyle & Hoey, 2007), is a hybrid of two previously existing models, each of which is based on one of the two architectures: Improved Performance Research Integration Tool (IMPRINT) (Archer, Lebiere, Warwick, Schunk, & Biefeld, 2002) and Adaptive Control of Thought-Rational (ACT-R) (Anderson, 1996). IMPRINT retains the network of tasks, which includes the ordering and conditions of task transitions, required for operations in a given system (Archer et al., 2002). ACT-R is a first-principle model of a human operator’s visual and motor interaction with the world, declarative and procedural memory, and actions chosen by pattern-matching information processed from the outside world with procedural memory (Anderson, 1996).

Together, the IMPRINT/ACT-R hybrid is a tool in which a task network is backed with models of human cognitive behavior. This structure is the foundation of the architecture of the simulation presented in this thesis. Its advantage for this project is that by making tasks the primary organizing feature, neither human or automation performance is implicitly assumed in the task network, allowing an objective study of human-automation task allocations.

#### **3.2.2.2 Level of Detail**

Another question that any simulation effort encounters at some point is, to what level of detail should the modeling be. One of the NASA HPM efforts notable for its modeling detail is Air MIDAS (Pisanich & Corker, 1995). Unlike ACT-R, which is based on a “unified theory of cognition,” Air MIDAS contains many separate detailed modules on various aspects of human behavior. A sampling of its modules includes visual and auditory perception and attention, working memory, domain knowledge, physical motor and anthropometric models, mental goal and task queues, and activity generation and scheduling which

feeds into the queues (Foyle & Hooey, 2007). At the lowest level, human behavior is represented as different types of “primitives”: motor (e.g., button push), visual (e.g., fixation on object), cognitive (e.g., recall), and auditory (e.g., monitor audio signal) (Tyler, Neukom, Logan, & Shively, 1998). Air MIDAS also includes a detailed simulation of the external world (Foyle & Hooey, 2007), and the outputs of this human behavior model are the effects of primitive actions on the external world model (Tyler et al., 1998). The downside to this model’s rich level of detail, however, is that outputs can be difficult to trace and validate.

At the other end of the spectrum, the Attention-Situation Awareness (A-SA) model is a simpler effort that focuses on a small subset of human cognition: attention and situation awareness (Wickens, McCarley, & Thomas, 2003). It is based on an algebraic formulation of the probability that an external event will be attended to by a human operator. This is back-ended by a model of situation awareness (SA) in which SA is rated by a numerical value from 0 (no awareness) to 1 (perfect SA). SA can be improved or degraded by correct or incorrect information. Its value also decays when there is old or irrelevant/distracting information.

Following the example of A-SA’s simplicity, the human behavior models in this project are based on basic and reasonable general assumptions of human cognition that are easy to understand and trace as the complexity of the simulation grows.

### 3.2.3 Human Information Processing

Models of human task performance behavior in this simulation were structured on a four-stage information processing model put forth by Parasuraman, Sheridan, and Wickens, as shown in Figure 6 (Parasuraman et al., 2000). It describes how information is attended to and perceived by a human operator and then used to select and generate appropriate actions on the external environment. This open-loop linear model is a simplification of Wickens’ closed-loop model in which information on the effects of an executed action is attended and perceived, forming a feedback loop (Wickens & Carswell, 2006). Wickens’ model also includes working and long-term memory and attention resources, which are not included in this simplified model.

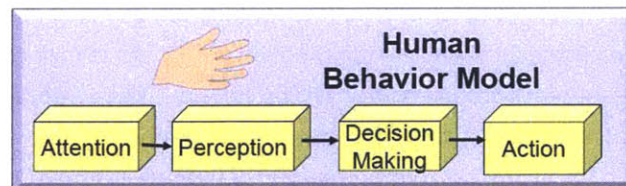


Figure 6: Four-stage model of human information processing

Assumed models of the four information processing stages are as follows:

**Attention:** Attention determines when and whether the human picks up an information input signal. An information input is assumed to be received, or it is not. This model is of focused attention; temporal changes in attention, as in visual search or guided attention, were not assumed (Wickens & Carswell, 2006).



Perception: Metadata and parameters needed for the next stage, decision-making, are extracted from information inputs that were successfully attended to in the previous stage. In addition, perceptual errors, namely errors in the calculation of needed decision-making parameters, are assumed.

Decision: Two types of decision making blocks are implemented to accommodate the different types of subtasks: rule-based decision making, in which actions are selected based on criteria applied to perceived information, and multi-attribute utility theory (MAUT) (Lehto, 1997). MAUT is a rational decision making model that takes account of the multiple objectives that the decision maker may pursue in making a decision (such as choosing a LAP that is far from hazards, close to a POI, as well as fuel-efficient). Each objective is given a weighting, or priority. Of the alternatives available to the decision maker, the alternative with the greatest expected value summed over the different weighted objectives is chosen.

Action: Since human behavior is being modeled at the abstract cognitive level rather than at the level of physical primitives, the action block is not needed for subtasks in which the action resulting from a decision is, for instance, a button press.

### **3.3 Method**

Unlike previously developed human performance models, the central focus of the simulation presented here is not to model human performance; rather, the focus of this approach is analysis of a system's human-automation task allocation design space. Equal importance is given to human and automated components, and system performance, rather than just human performance, in response to task allocations is evaluated.

The simulation here is also an extension of the latest trend in modeling tools: system-level simulations containing models of human operators, automated systems, and the surrounding system dynamics. Its models of human behavior, however, are simpler than those of MIDAS to enable understandability and traceability of results. Also, it draws on the architectural lessons of IMPRINT/ACT-R, as described in Section 3.2.2.1.

The approach was as follows: a subset of tasks from the lunar landing task analysis performed in Section 2 was selected and further elaborated for modeling (Section 3.3.1). A task allocation simulation was constructed based on a task network architecture. Within the simulation, models of human (Section 3.3.2.1) and automation (Section 3.3.2.2) behavior were implemented for each selected lunar landing task. The simulation was run for all possible combinations of human-automation task allocations (Section 3.3.3).

#### **3.3.1 Modeled Tasks**

From the lunar landing task analysis described in Section 2, landing point designation (LPD) (task 2.3 in Figure 4) and final vehicle touchdown (task 3.1 in Figure 5) were selected for modeling, as shown in Figure 7. The "automated system" listed in Figure 7 refers to automated technology that scans a landing area, creates a digital terrain map to identify hazards to the vehicle, and generates prioritized LAPs within the scan area, as assumed in the task analysis in Section 2. This system presents this

information to the user, who can choose to select one of the recommended points or free-select from other areas on the map (Epp et al., 2008; Forest et al., 2007).

LPD was elaborated into 3 subtasks incorporating the use of such an automated system. First, a decision has to be made on whether or not to use the automated system in LPD. If the decision is “yes,” then one of the LAPs suggested by the automation is selected. If decision is “no” (if faults are detected in the automated system, for example), then the alternative is for the human operator to free-select a LAP using the out-the-window (OTW) view or, if judged trustworthy, use the highest ranked selection based on the automated system’s terrain scan. The four resulting subtasks in Figure 7 were selected so that all three task types – Navigation, Supervisory, and Decision Making – can be demonstrated in human performance modeling in simulation.

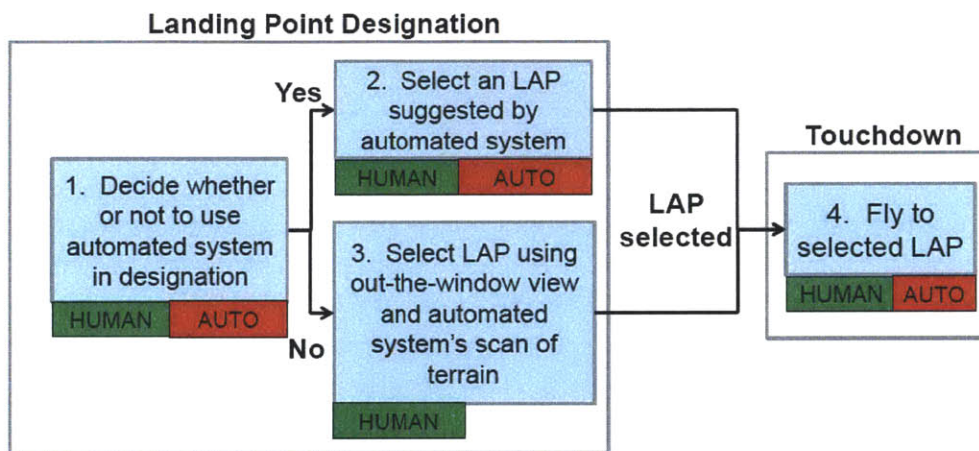


Figure 7: Task network, containing LPD and Touchdown subtasks, of the case example to be modeled.

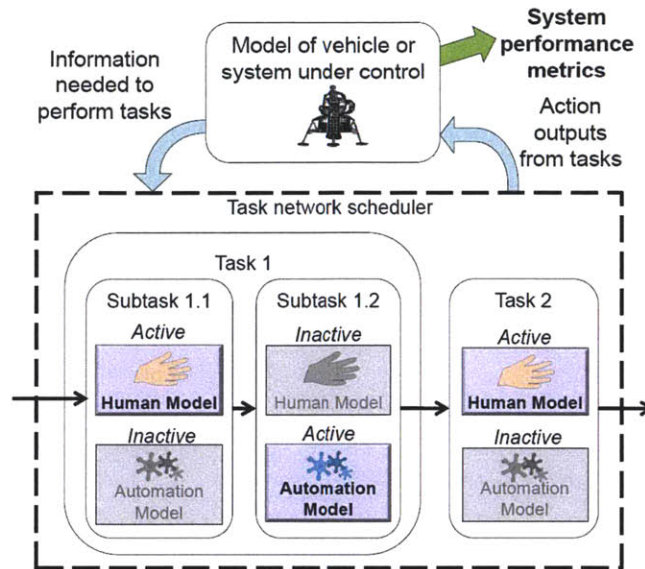
Each subtask may potentially be allocated to a human operator or to automation (the approach for developing these models is in Sections 3.3.2.1 and 3.3.2.2, respectively). Note that free selection of a LAP, subtask 3, can only be performed by the human.

Some of these subtasks modeled in simulation were later selected for human subject experimentation (Section 4). The models of human task performance in this simulation were implemented at a preliminary level of detail; therefore, the purpose of experiments is to uncover behaviors from human subjects which can potentially enrich these human performance models.

### 3.3.2 Simulation Implementation

The overall architecture of the task allocation simulation framework is depicted in Figure 8.





**Figure 8: Architecture of task allocation simulation framework**

Two models, one that describes human behavior and another one for automation behavior, were implemented for each subtask in Figure 7 (subtask 3 only has one model of how it is carried out by a human operator). For each simulation run, either the human or automation behavior model can be switched on for each subtask. By running all possible permutations of subtask allocations, system performance outputs can be obtained for every possible task allocation. This is possible since this scenario only has 8 task allocation combinations (Section 3.3.3, Table 1); scenarios with a larger number of combinations should consider segmented runs.

The task network scheduler, depicted in Figure 8 as a dotted-line box, is structured on the task network shown in Figure 7. The scheduler is the timing engine that drives the model. It triggers the appropriate performance model - human or automation, depending on the assigned allocation - for each subtask and directs the model to transition to the next task once a valid output is obtained from a previous task. For instance, the simulation transitions from subtask 1 in Figure 7 and begins executing subtask 3 or 4 (depending on the output of subtask 1) as soon as a non-null output, "0" for "no" or "1" for "yes," is read from the model of subtask 1. Likewise, subtask 4 begins executing as soon as a valid LAP location is output from subtask 3 or 4.

The model was implemented in MATLAB Simulink, with the task network scheduler implemented as a state machine in Simulink Stateflow.

Inputs to the human and automation behavior models include an out-the-window (OTW) 1-meter resolution view of a 180m x 180m lunar terrain area in which the vehicle may land, and its corresponding digital elevation map (DEM) (Cohanim, Fill, Paschall, Major, & Brady, 2009) generated by an automated landing and hazard avoidance system's scan of the terrain. Both are represented as 180x180 matrices of elevation values. There is also a digital hazard map (Forest, Cohanim, & Brady, 2008) calculated by an automated system based on the DEM and represented as 180x180 matrix of "1"



(hazard) and “0” (non-hazard) values. In addition, inputs include the coordinates of three LAPs generated by the automated system and ranked by desirability (Forest et al., 2008), as well as coordinates of a point of interest (POI). The POI is a location in the landing area to which it is desirable to land close, such as a lunar base camp, a geological feature of interest, or a grounded vehicle requiring assistance (Needham, 2008). Finally, inputs include current vehicle position, velocity, and attitude. The ways in which these inputs are used by the human and automation task performance models are detailed in 3.3.2.1 and 3.3.2.2, respectively.

Outputs from the simulation include the selected LAP location and fuel usage over the touchdown trajectory. The following vehicle dynamics are also output over the touchdown trajectory: position (altitude and location), velocity (descent rate and horizontal velocity), attitude (pitch and roll), and attitude rates.

### ***3.3.2.1 Human Behavior Models***

Perception and decision models for the subset of the task network selected for simulation (Figure 7) are shown below. To simulate errors in and the variability of human performance, the human task performance models contain numerical parameters that can be varied between simulation runs (see Section 3.3.3).

Errors in perception are modeled using variable gains and biases on information signals. For example, a perceived scalar distance between two points can be multiplied by a gain or have an added bias to simulate human misperception of that distance. Two simple visual perception behaviors are modeled here: calculating distances between points on terrain (such as between a potential LAP and the POI) and extracting characteristics of a visual field (such as slopes between elevation values in terrain).

Variability in human performance is simulated by varying decision weights (such as the relative importance of being far from a hazard versus being close to a POI while performing LPD), decision limits (such as the maximum percentage of a DEM display that can be deemed “blacked out” and still be used for LPD), and gains and time delays in the control loops used to model manual flying.

These variable parameters are shown in red in Figure 9 through Figure 13 and are listed in Appendix 2.

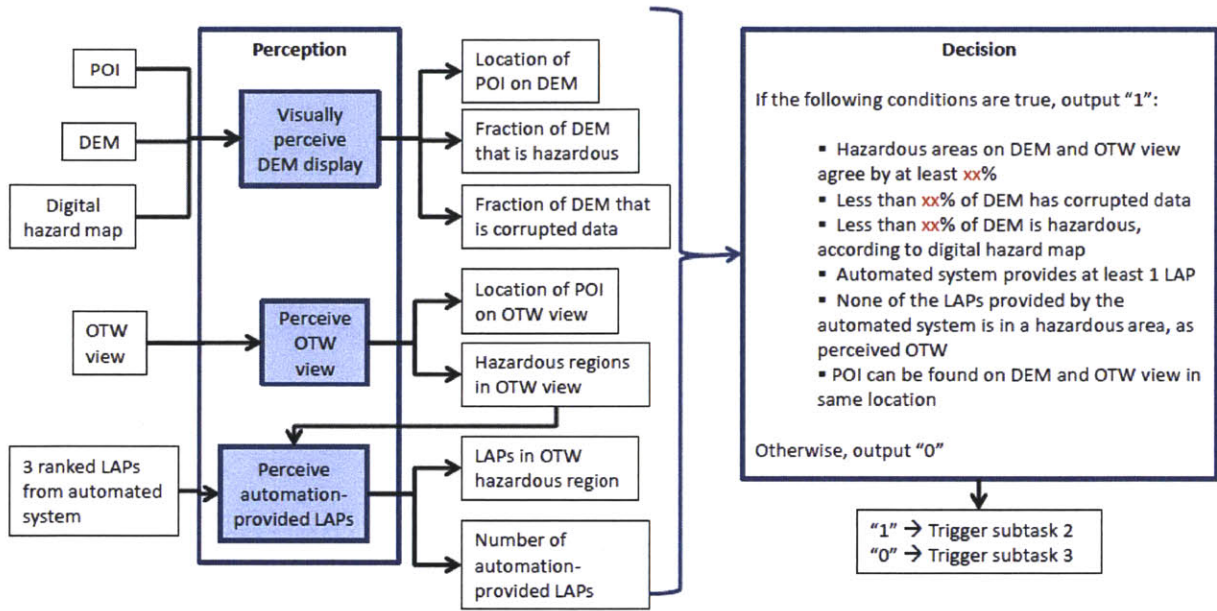


Figure 9: Human perception and decision models for subtask 1 (decide whether or not to use automated system in LPD)

In the two models of LPD (Figure 10 and Figure 11), a LAP is determined to be more fuel-optimal the more it lies in the direction of the vehicle's initial horizontal velocity and the farther it is from the vehicle's initial position (close to where the vehicle would land if all the pilot made no other maneuvers than to gradually decrease descent rate. See Section 3.3.3 for map of landing area and vehicle initial conditions).

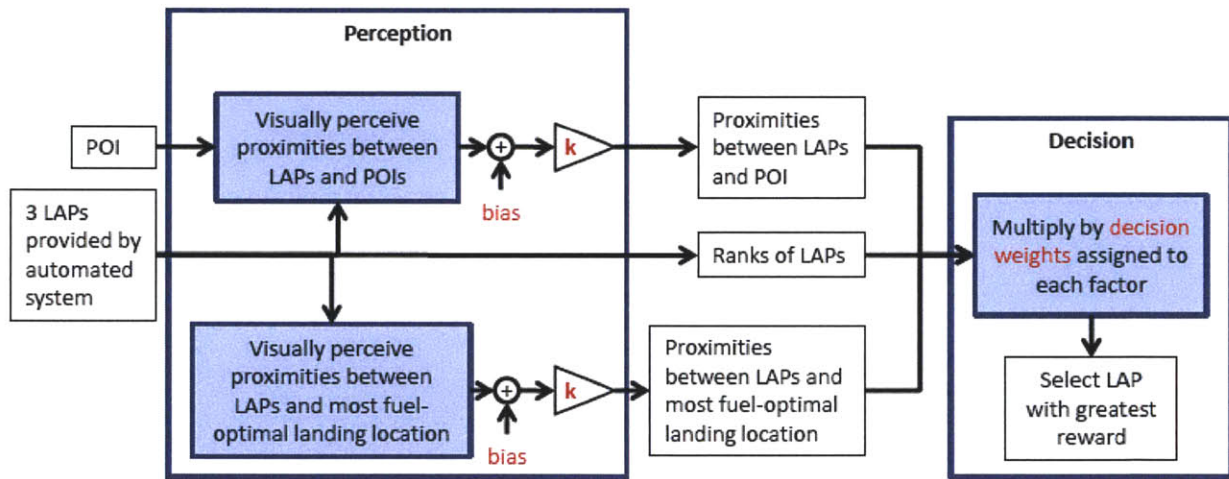


Figure 10: Human perception and decision models for subtask 2 (select LAP suggested by automated system)

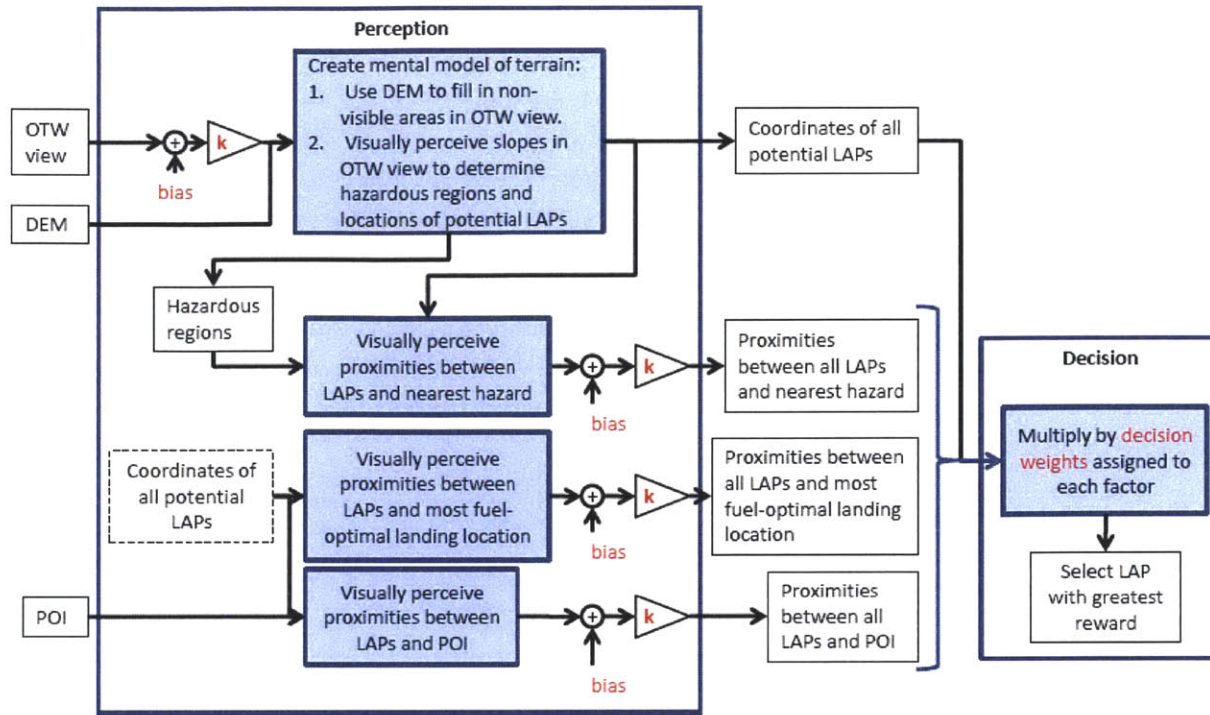


Figure 11: Human perception and decision models for subtask 3 (select LAP using OTW view and automated system's scan of landing area)

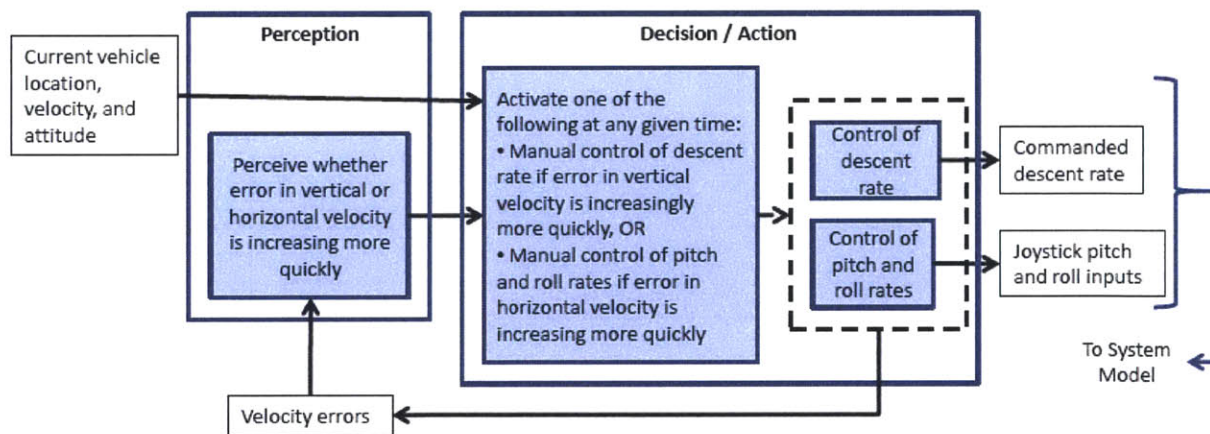


Figure 12: Human perception and decision models for subtask 4 (fly to selected LAP)

Physical actions resulting from a decision were not modeled. An exception is subtask 4, flying the vehicle to a selected LAP, in which the manual action of controlling the vehicle's dynamics substantially affects system performance. Human control of vehicle dynamics is divided into the vertical and horizontal axes of movement, as shown in Figure 12. To approximate human attention switching and workload limitations, only one control loop – descent rate or attitude rate – is active at any given time. This switching between control loops is governed by whether the perceived error in the vehicle's current



velocity is increasing more quickly in the vertical or horizontal direction. The control loops are based on the “crossover model,” which models human manual tracking behavior as a transfer function in a closed feedback control loop, as shown in Figure 13 (McRuer & Krendel, 1974). For the preliminary level of modeling used in this project, gains and time delays in the control loops were chosen so that a stable flight trajectory can be completed to LAPs selected almost anywhere within the landing area. The numerical ranges from which the gains and time delays were drawn are listed in Appendix 2.

Note that, here, the pilot is modeled as closing the loop on vehicle position and velocity. In future-generation vehicles, the pilot is likely to be aided by a flight director, which allows him to close the loop on attitude errors instead. The manual control modeled here is more analogous to how flying was performed in the Apollo program, with only an OTW view and instrument readings or callouts of position and velocity (Mindell, 2008).

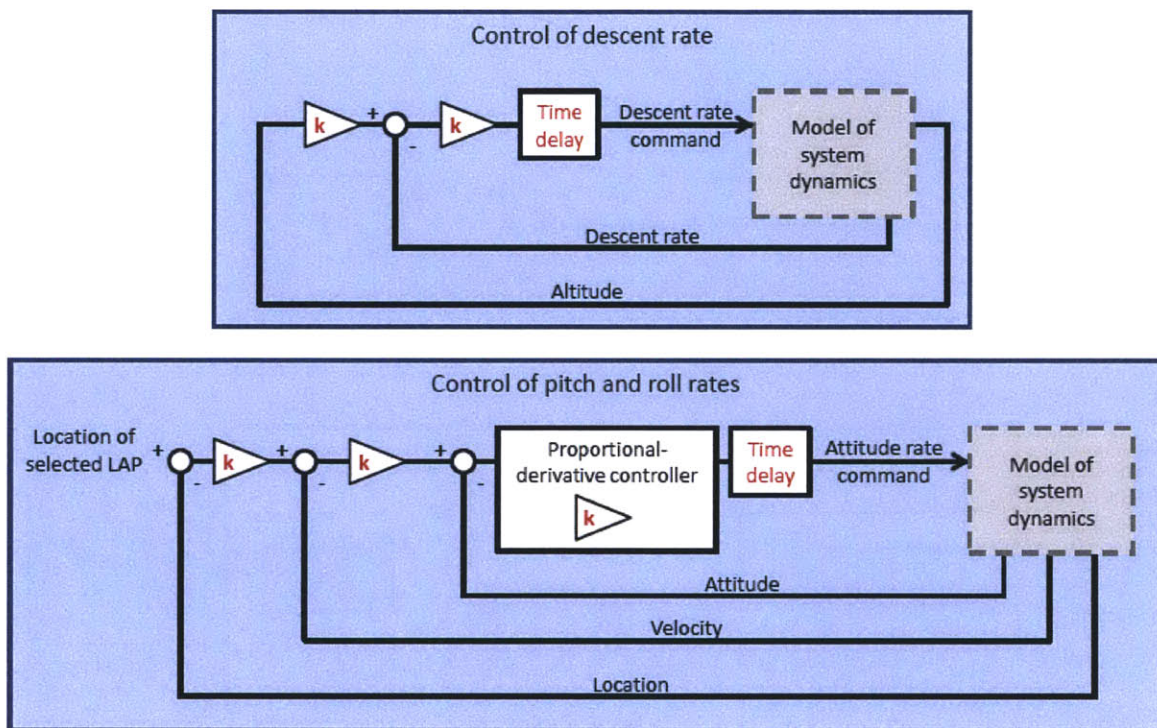


Figure 13: Closed feedback control loops modeling human manual control of vehicle movement vertically (top figure) and horizontally (bottom figure)

### 3.3.2.2 Automation and System Behavior Models

Models of automation behavior provide the automated counterparts to the human task performance models. They are based on, or incorporate previously generated outputs of, two existing models: an automated system for detecting hazards and selecting prioritized LAPs (Epp et al., 2008) and a control and dynamics model of a lunar landing vehicle (Hainley, 2011). The automation behavior models implemented for each of the subtasks in Figure 7 are as follows:

Subtask 1 (decide whether or not to use automated system in LPD): In the current implementation of the simulation, the automated system always decides to use itself in LPD.

Subtask 2 (select LAP suggested by automated system): After compiling a DEM from a scan of the terrain, the automated system may identify areas that are hazardous and generate candidate prioritized LAPs (Forest et al., 2008). The automation always selects the highest-ranked (prioritized) LAP. If this subtask is allocated to the human behavior model, these prioritized LAPs are the same ones presented to the human for selection.

Subtask 3 (select LAP using OTW view and automated system's scan of landing area): This subtask can only be allocated to a human behavior model.

Subtask 4 (fly to selected LAP): Automated behavior for the Flying subtask is governed by guidance equations developed by Bilimoria for a lunar landing trajectory to the selected LAP (Bilimoria, 2008) and implemented by Hainley for human subject cockpit experiments (Hainley, 2011). The vehicle was controlled along a reference trajectory in which the horizontal velocity was proportional to the range to the selected landing point and the descent rate decreased linearly until the vehicle was at a near hover within a certain altitude and range of the LAP.

The system model, labeled as “model of vehicle or system under control” in Figure 8, simulates the dynamics of a lunar landing vehicle. It was developed by Hainley (Hainley, 2011) and was based on an early lunar vehicle design for the Constellation program (Duda et al., 2009). The model takes as input the starting location of the vehicle and commanded descent, roll, pitch, and yaw rates from the human or automation models of the Flying subtask (subtask 4). The descent engine was modeled to be fixed-gimbal. When the vehicle was commanded at a non-vertical attitude, the thrust was proportionally increased to maintain the commanded descent rate. Therefore, larger attitudes resulted in larger horizontal accelerations. Fuel usage was calculated based on the descent engine's thrust output and specific impulse. The model's outputs – vehicle fuel consumption, location, velocity, and attitude over time – are used as system performance metrics, and also as feedback information to aid the human and automation performance models for the Flying subtask.

### **3.3.3 Simulation Procedure**

Figure 14 shows the sample lunar terrain area, in which a LAP can be selected and the lunar landing vehicle can touch down, used for all simulation runs. The terrain includes a POI near which a human pilot may desire to land.

The vehicle starts at an altitude of 500 ft., descending at 16 fps (feet per second). It is initially located at the center bottom edge of the map as shown in Figure 14, and has a starting constant horizontal velocity of 30 fps forward (towards the top of the map) and an initial vertical upright attitude. The locations of the LAPs and POI relative to the vehicle's initial position are as follows: POI – 262 ft. (50 m) to the north, 98 ft. (30 m) to the east; LAP 1 – 335 ft. (102 m) north, 180 ft. (55 m) west; LAP 2 – 427 ft. (130 m) north, 20 ft. (6 m) east; LAP 3 – 233 ft. (71 m) north, 190 ft. (58 m) east.



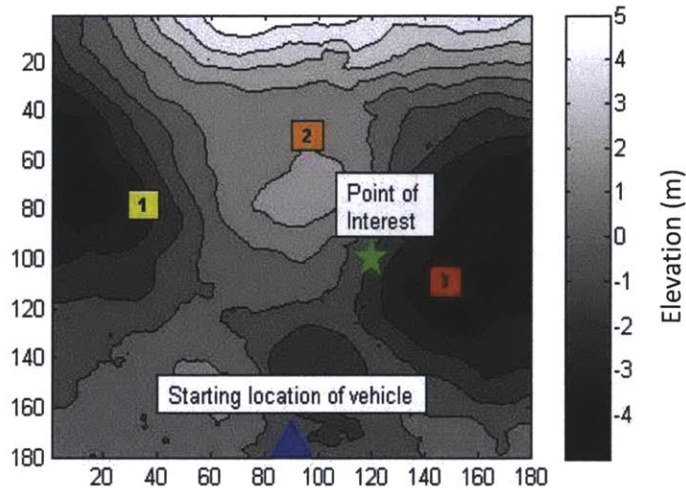


Figure 14: DEM of 591 ft. x 591 ft. (180x180m) terrain area used in simulation scenario. The scenario includes 3 ranked LAPs suggested by automated system (“1” = most highly suggested), a POI, and the starting location of the vehicle.

Since three out of the four subtasks selected for modeling can be allocated to either the human or the automation (see Figure 7), there are  $2^3 = 8$  possible different task allocations as shown in Table 1.

Table 1: Task allocations for simulation

1. Decide to use automated system or not	2. LPD using automated system	3. LPD without automated system	4. Flying
Human	Human	Human	Human
Human	Human	Human	Auto
Human	Auto	Human	Human
Human	Auto	Human	Auto
Auto	Human	Human	Human
Auto	Human	Human	Auto
Auto	Auto	Human	Human
Auto	Auto	Human	Auto

For each allocation, the simulation models were run 20 times, for a total of  $8 \times 20 = 160$  simulation runs on the one DEM show in Figure 14 (the number of runs was chosen based on limitations on computer processing time). For each of the 160 runs, all human model parameters were randomly generated afresh from a uniform distribution over predefined ranges to simulate the stochastic nature of human performance. Figure 15 summarizes the parameters varied in the human task performance models, and the numerical ranges from which parameters were generated can be found in Appendix 2.

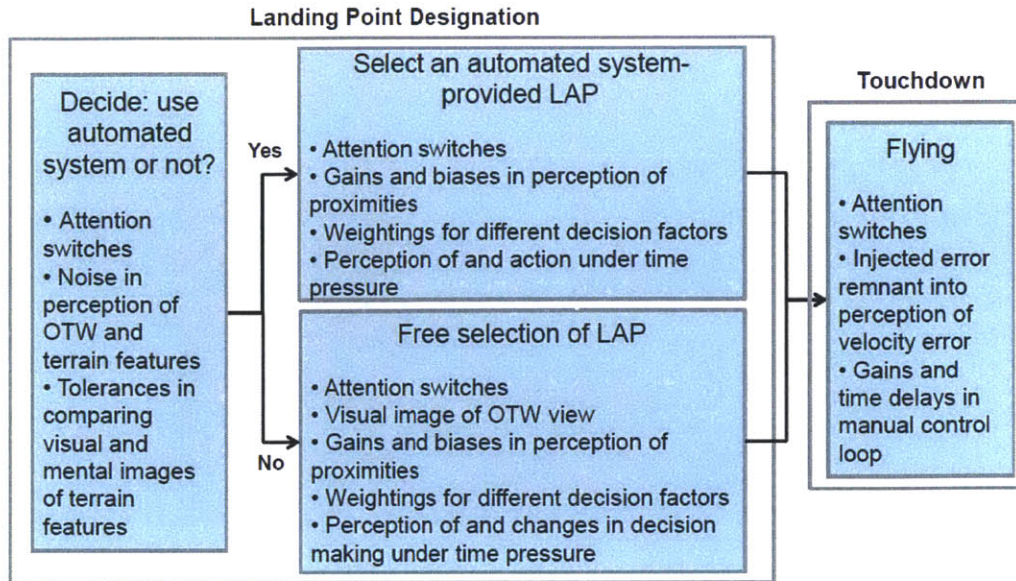


Figure 15: Summary of parameters in human performance models which were varied from one simulation run to the next

### 3.4 Results

The models developed here are a concept demonstration for the use of simulation in task allocation design. The resulting data analysis should be viewed likewise – not as an end in themselves, but as an example of how such simulation results may be useful in informing human-automation task allocation during early-stage system design.

Figure 16 below shows, for a given terrain map, the locations of LAPs (blue asterisks) selected by the human model.

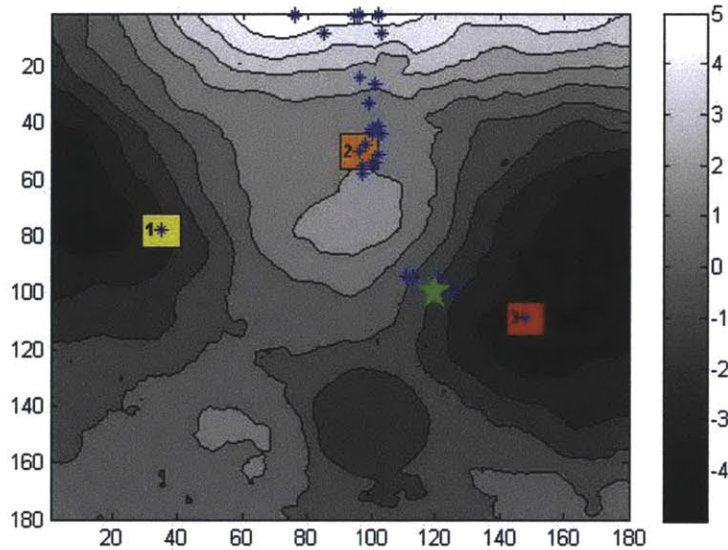


Figure 16: DEM of terrain with LAPs suggested by automated system (squares), a POI (green star), and LAPs independently selected by human model (blue asterisks) over all simulation runs

Figure 17 shows how each possible task allocation fares on fuel savings and landing accuracy based on simulation outputs.

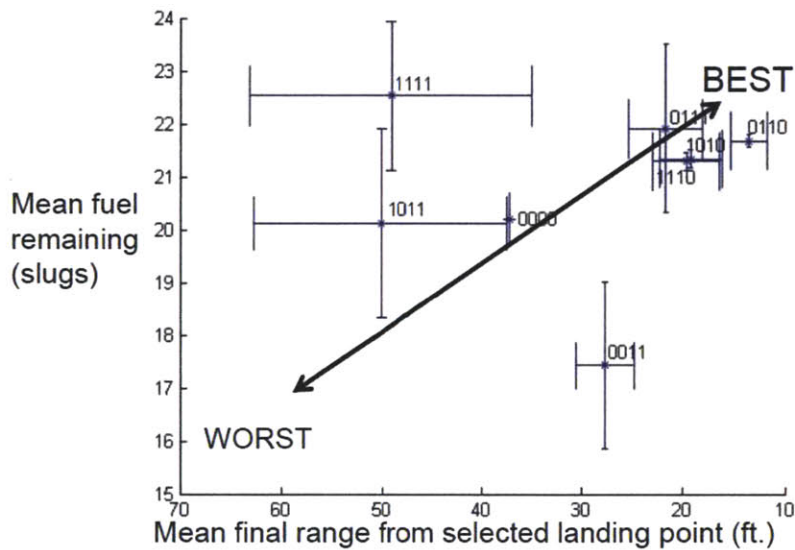


Figure 17: Simulated performance of each task allocation, based on mean remaining fuel and range from selected LAP at moment of touchdown. Each allocation is denoted by 4 digits, with each digit signifying the allocation of a subtask: “<subtask 1> <subtask 2> <subtask 3> <subtask 4>.” “1” = assigned to human, “0” = assigned to automation. Error bars indicate standard error in the mean.

The task allocation that appears to yield the best fuel and accuracy performance is allocation “0110,” in which the human does the decision-making but aided by automation, and automation does the flying. Judging by the standard error bars in Figure 17, it appears to be significantly superior in final landing accuracy, even if it is not significantly superior to its neighboring allocations – “0111,” “1010,” and



“1110” – in fuel remaining (allocations with overlapping standard error bars are not likely to be significantly different from each other. Note that the standard error bars for “0110” along the final range axis do not overlap those of neighboring allocations; hence, this allocation is likely to be significantly different from the others). This is to be expected: the human decision-making models take into account the location of a LAP for fuel savings, and automated flying is more capable than a human pilot would be of following an optimal flight trajectory. Note, from the spread of the data points in Figure 17, that the allocation of the flying subtask 4 is not by itself a strong predictor of fuel and accuracy performance; automated flight (the allocations ending in “0”) do not necessarily all produce the most fuel-optimal and accurate landings. This effect is further explored below, by looking at the variability in fuel and landing accuracy due to different task allocations.

The standard deviations in remaining fuel and final range, as shown in Figure 18 and Figure 19, provide a more nuanced view of the results above. The variability in fuel savings (Figure 18) obtained when the human model is performing the subtask of flying is greater than that obtained from automated flying – this is to be expected. This was verified by performing a Levene’s test between allocations that differ only in the human-automation allocation of the flying task (last digit). In all cases, the test was significant with  $p \leq 0.001$  (Levene’s test was not performed between “0000” and its counterpart allocation, as a fully automated allocation has no variability). Note, however, that not all allocations with human manual flight use more fuel than automated flight – only that there is more variability in fuel use in task allocations in which flying is allocated to the human. For instance, there was no variability in fuel use when the allocation is fully automated (“0000”) because the automation always selects the same LAP location and executes an identical flight trajectory for every simulation run.

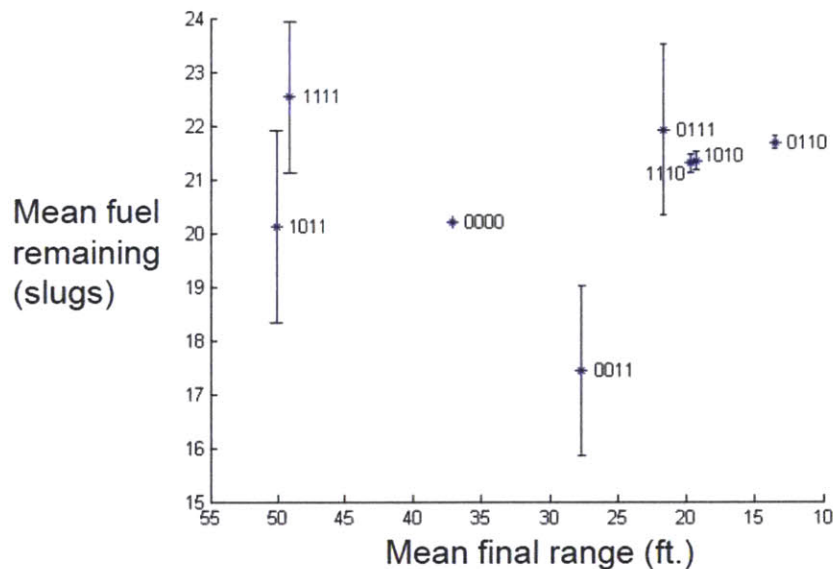


Figure 18: Plot highlighting standard error in the mean of remaining fuel amounts for each task allocation (see Figure 17 caption for explanation of allocation notation)

Although the allocation of the Flying subtask is the sole determining factor of variability in fuel use, this is not the case for landing accuracy (Figure 19). In the allocations “1110,” “1010,” and “0110” the

human performs much of the LPD decision making and the automation flies the vehicle. The opposite is true for allocations “0111” and “0011”: the automation has a share in LPD, but the human model controls the vehicle (in this model, the automation will always make the decision to use the automated system in LPD). Interesting, the variability in landing accuracy is similar for these five allocations, as seen in Figure 19. Bonferroni-corrected t-tests were performed between a group containing the allocations “1111” and “1011” and another group containing the allocations “1110,” “1010,” “0110,” “0111,” and “0011.” The two groups were significantly different in their means ( $p = 0.042$ ) and their variances ( $p < 0.0005$ ) estimated from the variances inferred from the variances of the different allocations within a group.

This suggests that human involvement in LPD can cause just as much variability in landing accuracy as human involvement in flying does. One possible reason is that human-performed LPD results in a greater variety of selected LAP locations, not all of which are as easy to navigate to as LAPs selected by the automated system, given the location and velocity of the vehicle. The automated guidance and control system has its own landing dispersions; in other words, its landing accuracy depends on how far the targeted LAP is located from the initial vehicle position and whether the LAP lies in line with the vehicle’s initial horizontal velocity vector. This would account for the fact that allocations ending in “0,” in which flying is automated, do not necessarily produce landings with zero range error (this includes the fully automated allocation “0000” – an automated flight to an automated-selected LAP will still have a landing dispersion, even if the dispersion does not vary because the LAP selection and flight trajectory does not vary). Therefore, the variability in human LAP selection is likely to cause variability in landing accuracy, even if flying is automated. Figure 20 summarizes the general trends in landing accuracy variability as a result of task allocation.

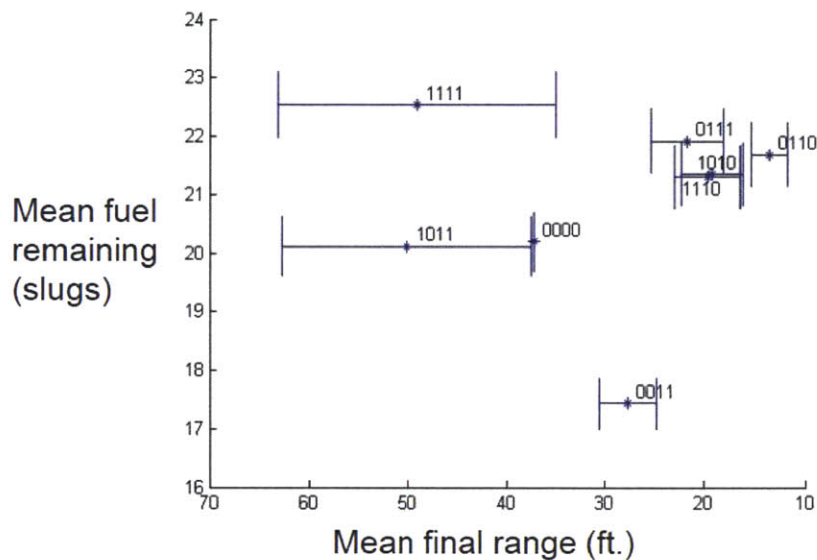


Figure 19: Plot highlighting standard error in the mean of final range from selected LAP for each task allocation (see Figure 17 caption for explanation of allocation notation)

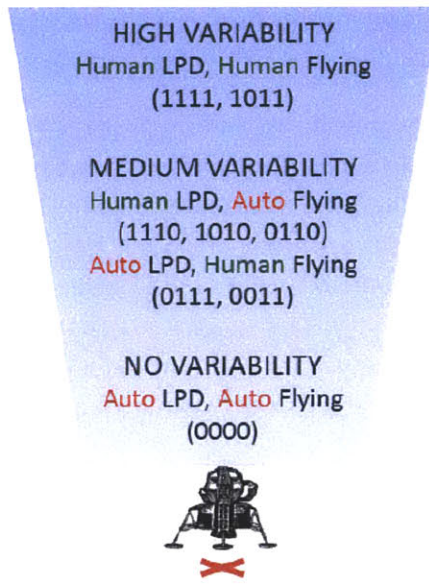


Figure 20: Trends of relative variability in landing accuracy for different task allocations (see Figure 17 caption for explanation of allocation notation). Note: if subtask 1, “Decide whether or not to use automated system in LPD,” is allocated to automation, the automation will always decide “yes.”

### 3.5 Discussion

One drawback of simulation is that it is impossible to fully model all human behaviors that would affect system performance under any given task allocation. Human behavior at the cognitive level is complex and difficult to model; research is ongoing in understanding and modeling the true mechanics of human visual processing, attention, and decision making. These behaviors are currently implemented in the model as idealized, simple algorithms that attempt to capture the essence of how a task would be completed at the abstract level, but will require validation against human experimental data. For example, numerical values that were chosen for the human model parameters and varied between simulation runs (Appendix 2) would realistically be highly specific to a given pilot, system, and lunar landing scenario. Here, representative values were chosen for a hypothesized system.

Aspects of human behaviors not currently represented in this model are long-term memory, situational awareness, workload, and fatigue. Also, ranges of system performance are lost when aspects of human performance simply cannot be modeled. Humans are often put in supervisory control of systems for their judgment, flexibility to a range of off-nominal situations, and high-level pattern recognition. These intangible traits help keep system performance high when off-nominal situations are encountered; however, such situations are difficult to predict and model. Therefore, it is all the more important to perform human subject experiments as a next step to uncover such un-modeled human behaviors that, within the context of task allocation, affect system performance.

### 3.6 Conclusion

An approach was introduced by which to generate high-level performance evaluations for a system given different allocations of tasks between a human operator and automation. The tasks of Landing



Point Designation (LPD) and Touchdown in lunar landing were selected for modeling. Models of human and automation behavior were implemented for all subtasks included in these two tasks.

Results from the simulation show the magnitudes and variability of system performance in response to different allocations of the subtasks to human and automation behavior models. For instance, system performance was predicted to be optimized if the human performs decision making tasks, and manual tasks such as control of vehicle dynamics are automated (as in the allocation “0110” in Figure 17, modeled in this simulation scenario). Variability in fuel use can be directly attributed to human performance of the flying task. Variability in landing accuracy appears to result not only from variability in human manual flight but also from the variability in human performance of the LPD task (even when flying is automated). Results indicate that a mixed human-automation allocation increases the chances of optimized system performance.

The work presented here shows the value of simulation in quantitatively evaluating system performance in response to many possible task allocations in a system’s design space; such broad and high-level evaluations would have been difficult to obtain by other means, such as human experiments. Simulation results indicate which types of allocations are likely to be superior and how sensitive system performance is to the allocation of a given task for a given set of human and automation model parameters. In addition, the simulation framework presented here differs from previous developed human performance models in that it is specifically tailored to evaluate ranges of human-automation task allocations.

## **4 Effect of Perceived Risk on Lunar Landing Point Selection**

### **4.1 Introduction**

#### **4.1.1 Motivation**

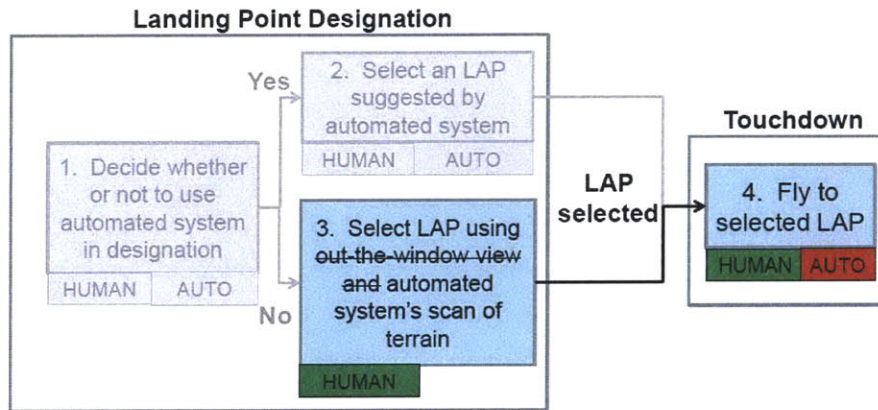
No matter how tasks are allocated in a crewed lunar landing system, a human operator will still be in the role of supervisor. The task allocation simulation described in Section 0 does not capture the adjustments in decision making and manual flying that a pilot with high-level knowledge of the system performance is capable of making. Such behaviors are best understood by studying the performance of human subjects in a lunar landing simulator. The resulting evidence could then eventually be used to enrich a task allocation model.

One such issue is the case in which the POI is located very near hazardous terrain. If the human pilot decides to re-designate the currently selected LAP, the selection of that new aim point will likely depend on whether the pilot anticipates the terminal descent will have to be flown manually or automatically. In addition, the dispersion of actual touchdown points relative to the LAP when either flying manually or automatically will likely also be factored in to the decision. The pilot knows that his manual flying accuracy when following guidance cues to touchdown will likely not be as good as that of the automatic system. Placing the LAP close to hazardous terrain requires estimation of the risk of touching down in the hazardous area. LAP placement thus requires the pilot to make trade-offs among the perceived

benefit of landing close to the desired POI, the perceived risk that the actual touchdown may occur in hazardous terrain, and the likelihood of having to take over and fly manually. It may also depend on the individual pilot’s personal willingness to take risks. A human subject experiment was performed to explore how human pilots tailor LAP placements in response to different terminal descent task allocations.

#### 4.1.2 Problem Statement

From the subtasks modeled in the simulation (Section 0), subtasks 3 and 4 – “Select LAP using OTW view and automated system’s scan of terrain” and “Fly to selected LAP” – were chosen for human subject experimentation (Figure 21). Due to the difficulty of simulating a realistic lunar OTW view, subtask 3 was modified so that a LAP would be chosen using only an automated system’s scan-based map of the landing area, i.e., simulating a landing in darkness at the lunar pole or dark side, for example (Epp et al., 2008).



**Figure 21: Subset of tasks modeled in simulation that was selected for human subject experimentation. Subtask 3 is implemented in the experiment without the use of an OTW view (i.e., only automated system scan). Subtasks 1 and 2 (grayed out) were not used in the experiment.**

The landing area map included hazardous and nonhazardous regions, with a POI located on the edge of a hazardous region (Figure 23 and Figure 24). Subjects were asked to place the guidance LAP target so that, anticipating the likely touchdown dispersion, the actual touchdown would be as close as possible to the POI but never in the hazard region. Pilots were told they would be evaluated based on their average performance (whether vehicle was out of the hazard region, close to the POI, and within fuel and dynamic safety limits at contact) over multiple trials. On each run, the pilots were told the probability that they would have to take over and fly manually. They estimated the touchdown dispersion (scatter and bias) based on their experience in initial practice sessions, and later from plots shown to them depicting typical dispersion of simulated runs. This scenario of landing close to a POI located near hazard area is typical of several anticipated precision landing scenarios (e.g. land near a base camp but not on top of it, or landing near a geological site that, in itself, is hazardous for landing). It is also representative of an operational scenario in which repeated landings are made, such as supply/resupply missions to a lunar base camp.



Scientific question:

How does the allocation of the second task (manual vs automatic terminal descent flying) influence pilots' performance of the first task (manual placement of the LAP)?

i.e.,

Will pilots risk landing closer to a POI – and hence, closer to a hazard region – if the chances of a precise automated landing are greater? Will they choose LAPs that correctly account for landing accuracy errors in their own flying ability and in the automation, if they know who will, or may\*, be doing the flying?

\*In an operational scenario, the pilot may be aware of the (usually small) probability of a flying task allocation changing in the midst of a mission. For example, he may have to take over flying manually in case of an automated system failure or late-stage discovery of undetected hazards in the targeted landing point.

## 4.2 Background

Previous research on decision making in LPD studied how information display affected the choice of LAPs. For instance, Needham researched different levels of automation aid in selecting a LAP, and discovered that higher levels of aid (for example, presenting subjects with a winnowed selection of LAPs, or ranking the LAP selections) increased the safety of LAPs chosen by subjects (Needham, 2008). In her experiment scenario, subjects had the objective of selecting LAPs that were close to a POI; therefore, higher levels of automation allowed them to minimize the loss of safety in pursuing that objective.

This experiment, however, attempts to uncover how human decision making biases and factors affect LAP selection. Such biases and factors would cause decision making to deviate from rational or normative models, such as the rule-based and multi-attribute utility decision making modeled in the task allocation simulation in Section 3 (Wickens & Hollands, 2000). In this experiment's scenario, rational behavior would call for subjects to adjust the LAP location according to knowledge of the flying task allocation and manual vs. automated landing accuracy (presented to or achieved through experience by the subject) to achieve the goal of landing as close to a POI as possible without landing in a hazard region over several flights. In other words, subjects would maximize "expected long-term gain, which after all can only be realized following a long-term average of the outcome of several decisions (Lehto, 1997)" (Wickens & Hollands, 2000).

### 4.2.1 Memory Factors

Humans utilize shortcuts in their recall of past observations that may prevent them from making decisions based on the "long-term average of the outcome of several decisions" referred to above (Wickens & Hollands, 2000). The following factors could conceivably influence memory of previous touchdown accuracies if the subject observed them successively over time:

- Availability factor: Certain past observations may be easier to recall than others, especially those that are more recent, simpler to understand, or whose details are salient enough to aid their encoding into long-term memory (Tversky & Kahneman, 1974).
- Anchoring factor: Early observations are weighted more heavily than later observations (Tversky & Kahneman, 1974).
- Recency factor: The opposite of anchoring – the most recent observations are weighted the most heavily (Baddeley & Hitch, 1974).
- Limitations of working memory factor: Since not all past observations can be kept in short-term memory, the above factors dictate which ones are remembered (Baddeley & Hitch, 1974).

One way to mitigate the effects of these heuristics is to remove the time factor from the way in which landing error information is presented to the subject. For example, subjects could be asked to make LAP selections based on plots simply depicting touchdown dispersion data for all previous landings in a single plot.

#### 4.2.2 Biases in Perception of Risk and Reward

Rationally, the greater the difference in the landing accuracies of the subject and the automated system, the more the subject's LAP selections should compensate for the difference when presented with manual vs. automated flying task allocations. However, this may not be the case if a subject is risk-averse and biased towards planning for the worst case rather than to maximize long-term gain. To avoid any risk of landing in a hazardous area the subject may always designate LAPs far from a POI located on the edge of a hazard, no matter what the chances are of automated versus manual flight. In other words, the subject aims for a future in which all landings are safely away from hazards, rather than to maximize the number of landings that are close to the POI at the cost of a higher risk of landing in a hazardous region.

Kahneman and Tversky identified this bias as "loss aversion," in which losses are seen as worse than gains of the same magnitude (Kahneman & Tversky, 1984). In lunar landing, selecting a LAP close to a POI and having the automated system execute an accurate landing that avoids the hazardous region would yield a large reward. However, the loss-averse subject may consider the possibility that the flight may have to become manual in an off-nominal situation, in which case the decreased accuracy of a manual flight may result in touchdown in a hazardous region. This high cost is perceived to outweigh the reward, even if the cost and reward are of equal magnitudes. This is depicted by the plot of subjective vs. real-world value (utility) in Figure 22, left. The large slope on the left-half plane indicates that losses are perceived as worse than they really are; the relatively gentle slope on the right-half plane shows that gains are perceived as less than their actual value.

A corollary to this is the tendency to overestimate the frequency of low-probability events and underestimate higher-probability events (Kahneman & Tversky, 1984). Even when explicitly informed of the high chance of automated flight compared to the lower chance of manual flight, a cautious subject may overestimate the probability of the less desirable outcome (manual flight). This bias is illustrated by the hypothetical relationship between subjective and real-world probability in Figure 22, right. When probability is low, the perceived subjective probability (solid line) is greater than the real-world



probability (dotted line). At higher probabilities, the opposite is true. Therefore, low-probability events are overemphasized compared to high-probability events.

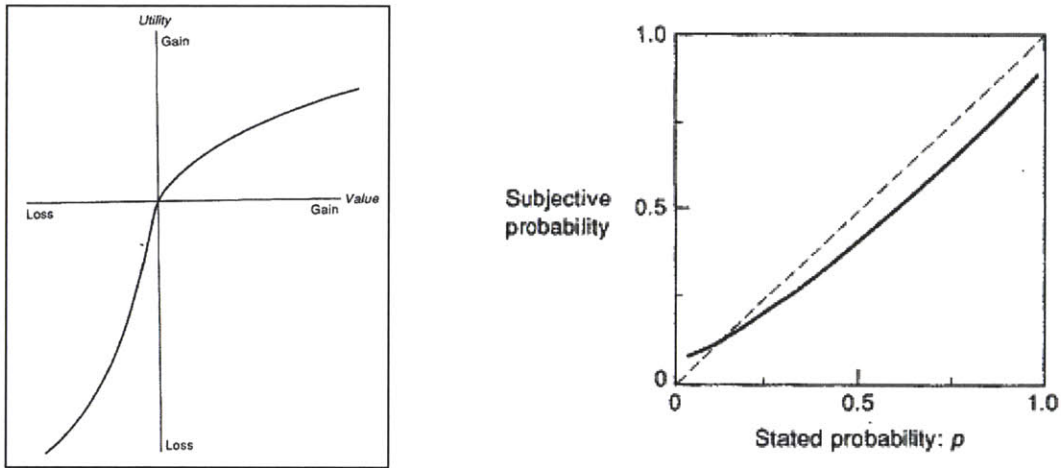


Figure 22: Subjective versus true probability and utility (Kahneman & Tversky, 1984)

### 4.3 Methods

#### 4.3.1 Independent Variables and Trial Schedule

Three independent variables in the LPD task were manipulated in this experiment: 1) flying task condition 2) landing dispersion information presentation condition, and 3) hazard area location.

1. Flying task condition: Subjects were told that either a) they would definitely be flying manually (100% probability), or b) that the descent would be initially flown automatically, but there was a 25% probability they would have to take over and fly manually once they have selected a LAP (representative of an automated system failure). It was expected that if pilots knew there was a chance they might fly manually, they would deliberately adjust the guidance LAP relative to the hazard area to compensate for manual touchdown dispersions, unlike if automatic flight was expected (this adjustment may put the LAP closer or farther to hazards, depending on the distribution of manual touchdown dispersions).

Based on preliminary runs of the experiment, a 25% chance of manual flight and 75% chance of automated flight were chosen to elicit expected-value decision-making in the LPD task: the chance of a precise automated landing should be high enough to encourage subjects to select LAPs differently from the way they would if they knew that flying would be manual, yet the chance of a possibly less accurate manual landing was also high enough to be considered.

(Since a manual LPD option was provided, and there is always some chance of a human pilot having to take over manual flight in case of off-nominal situations, the 0% probability of manual flight case was not studied.)

1. Landing Error Information condition: There are two ways in which the subject could acquire information on manual landing dispersion: a) previous direct experience, and b) scatter plot of



representative errors. The first way is for the subject to fly successive trials in manual mode and observe the consistency and bias in his landings relative to the guidance LAP. However, as discussed in Section 4.2.1, the subject's memory may be influenced by recall biases and the limitations of working memory. To avoid these biases, the subject could be shown a scatter plot of synthetically generated landing errors, and told that they are representative of a previous pilot's manual touchdown dispersion. The subject would then be asked to set the LAP based on the scatter plot data. In both instances, the subject was expected to bias the LAP in a direction towards or away from the hazardous area by a distance that depends on the anticipated touchdown dispersion and the personal loss aversion tendencies of the subject. To assess the effect of the character of the scatter plot data on the subject's performance, the scatter plot synthetic data was derived from a normal distribution, but with three different bias levels and two different standard deviations, for a total of six different scatter plots (Section 4.3.4.2).

2. Landing Area Map condition: Approaches were always flown north, but in one condition, the vehicle had to overfly the hazard area to reach the POI, and therefore had to be landed in the safe zone "long" of the POI. In a second condition, the vehicle overflowed safe terrain to the POI and had to land "short" of the hazardous region.

Table 2 shows the design of trials in this two-part experiment. After a PowerPoint briefing (Appendix 3), subjects were allowed a total of 12 trials (9 manual and 3 automated) to practice selecting a LAP and flying to the LAP when the flying task was allocated to them. This also allowed them to gain a preliminary impression of the character of their manual landing errors, prior to “Part 1” of the formal experiment.

Part 1 of the experiment allowed the subject to gain manual landing error information by performing successive manual flights. After subjects completed Part 1, they underwent “Part 2” in which landing error information was presented in scatter plots. In both parts of the experiment, the two flying task allocations were presented to the subject during LAP selection.

**Table 2: Number of trials in each part of experiment**

<b>Part 1 (Landing error information = subject's own manual flights)</b>	<b>Part 2 (Landing error information = scatter plots)</b>
100% chance of manual flying: 2 landing area maps x <b>5 repetitions</b> = 10 trials  25% chance of manual flying → but flying turns out to be automated (no manual flying performed, touchdown not simulated): 2 landing area maps x <b>3 repetitions</b> = 6 trials  25% chance of manual flying → but flying turns out to be manual (subject manually flies to LAP): 2 landing area maps x <b>2 repetitions</b> = 4 trials  = 20 trials	2 flying task conditions (100% or 25% chance of manual flying)  x 6 scatter plots (Section 4.3.4.2)  x 2 landing area maps  x <b>2 repetitions</b> per trial type  = 48 trials

All subjects completed Part 1 before Part 2, since it was necessary for the subject to gain hands-on understanding of the manual flying task before selecting LAPs based on scatter plot landing dispersions. Within each part of the experiment, trials were pseudo-randomized: the order of conditions was random, but balanced so that all distinct trial types had approximately the same average trial age so that there would be no confounding of possible learning effects with the effect of trial type. Each set of conditions had at least 2 repetitions, as seen in



Table 2. In Part 1, the “100% chance of manual flight” flying task conditions were repeated 5 times to allow more data on subjects’ manual flights to be collected, and the “25%” flying task conditions, in which flying turns out to be automated, were repeated 3 times to give subjects the feel that, when presented with that condition, there was indeed a greater chance of the flight turning out to be automated rather than manual. All subjects received the same schedule of trials, shown in Appendix 4.

#### 4.3.2 Hypotheses

1. The greater the individual subject’s manual landing error, the greater will be the effect of flying task condition on the risk level of selected LAPs.

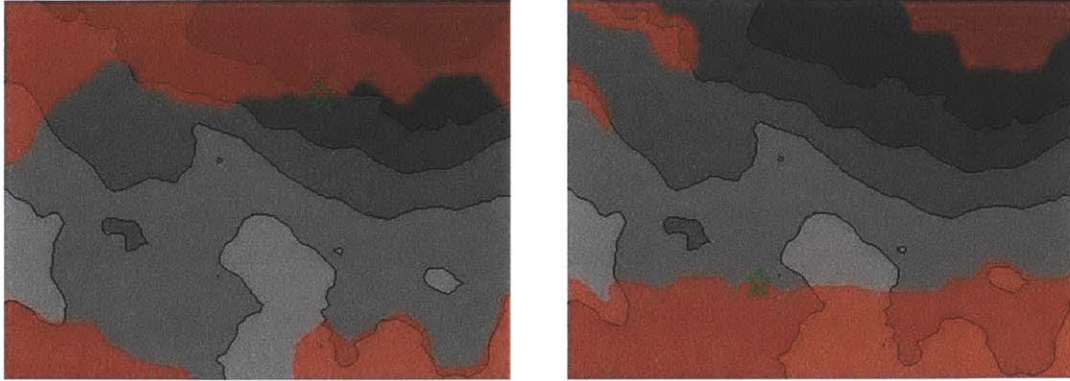
This hypothesis assumes that subjects use rational decision making in LPD – the greater the subject’s manual landing error, the more the subject should adjust LAPs to compensate for the error when there is 100% rather than 25% chance of manual flight.

2. The presentation of landing error information (successively observed landings vs. scatter plot of previous landings) affects the risk level of selected LAPs.

This hypothesis assumes that subjects’ recall of their manual landing errors are subject to memory factors described in Section 4.2.1. If this is not the case, then their recall would, in theory, be similar to their looking at a scatter plot of landing errors; the LAPs selected under both cases would not have significantly different risk levels (risk level given the landing accuracy of whoever the LAP is selected for, the subject or the hypothetical pilot represented by the scatter plot).

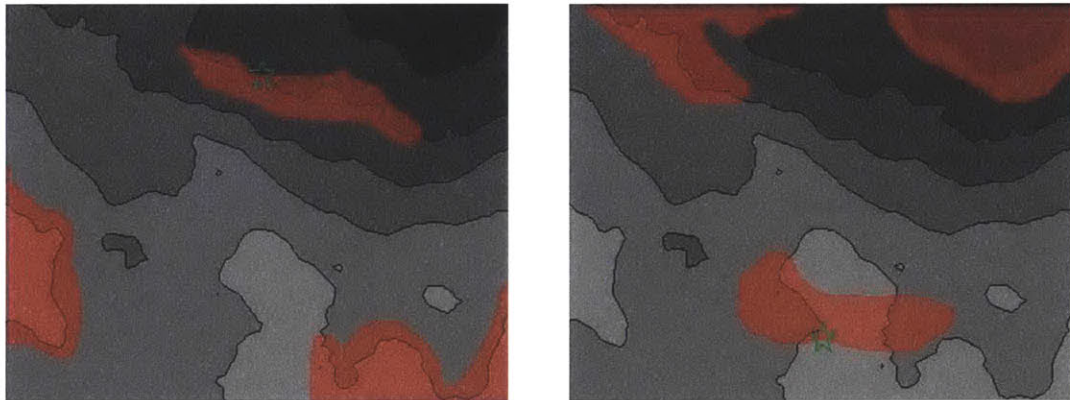
#### 4.3.3 Scenario Design

In each trial, one of the two different landing area maps shown in Figure 23 was presented to the subject for LPD. Each map represents an 800 ft. x 600 ft. area, and a LAP could be selected anywhere on the map. Regions that are hazardous for landing were colored red, and each map contained one POI (green star) that lay on the edge of a hazardous region. To avoid hazard regions, the vehicle must be landed short of the POI in one map (called the “short” map) and long of the POI in the other map (called the “long” map). The POI was offset about 100 ft. east of center in the short map and west of center in the long map. These maps are representative of the DEMs that would be generated by an automated system that scans a landing area on the lunar surface and calculates hazardous parts of the terrain.



**Figure 23: Maps of landing areas used in human subject experiment. Short map at left, long map at right.**

Two additional maps, called the “northern POI” and “southern POI” maps (Figure 24), were presented in which subjects may decide to land long or short of a hazardous region. In the northern POI map (Figure 24, left), flying over the hazardous region would allow a landing closer to the POI, but would consume more fuel (the vehicle always approaches from the south). In the southern POI map (Figure 24, right), landing short of the hazard region would allow a landing closer to the POI, but there is a risk of overshooting the POI and landing in hazards. These were presented during the second half of training when subjects had more familiarity with their landing errors.



**Figure 24: Additional maps included in training session. Northern POI map at left, southern POI map at right.**

Using a set of PowerPoint briefing charts (Appendix 3), subjects were instructed to land as close to the POI as possible without any part of the vehicle intersecting a hazardous region. To encourage this behavior, a landing scoring system was introduced which penalized subjects for landing in a hazardous region and rewarded them for landing close to the POI. A numerical scoring system also provides all subjects with the same weightings for the two competing objectives of landing away from hazards yet close to the POI. To encourage expected-value decision making, subjects were told they should “maximize your total landing score summed over all your trials,” so competitive subjects would not take exceptional risks to achieve the highest single run score. Due to the difficulty of calculating the location of landings relative to the edge of hazardous region as seen on the display in the midst of an experiment, subjects were not told their landing scores.



The scoring system was as follows:

- Any part of vehicle intersecting hazardous region: -10 points
- Vehicle completely in nonhazardous region:  $5/x$  points, in which  $x$  = distance from center of POI to center of vehicle, in widths of the vehicle (vehicle width = 30 ft.). The maximum possible score is 10 points for a landing in which an edge of the vehicle is as close to the center of the POI as possible while barely touching the red region (score =  $5 / 0.5$  vehicle widths = 10).

During LPD in each trial, subjects are informed either that they will have to do the flying (100% manual) or that there is a 75% chance that flying will be automated (25% manual). Subjects were also told that the automated system always lands the vehicle no more than 10 feet off the center of a selected LAP. In manual flights, subjects must stay within safety limits on vehicle velocity, attitude, and fuel at touchdown (Bilimoria, 2008); otherwise the landing will be scored at -10 points. Subjects were shown a fuel contour map over the landing area which indicates that selecting LAPs towards the south and center of the landing area map would result in less fuel usage for the vehicle, allowing more hover time to reduce landing errors. Instructions to the subjects on the scoring system, touchdown safety limits, and the fuel contour map are shown in the subject training slides in Appendix 3.

The initial conditions of the vehicle at the start of each flight are listed below. Contact occurred when the vehicle's center of mass is 12 ft. above the ground.

- Horizontal range: 1,046 ft. south of the center of the landing area map. This range was chosen to be large enough so that no matter where a LAP was located on the landing area map, the vehicle's calculated trajectory would not have to overshoot and curve back southward towards the LAP.
- Altitude: 500 ft. altitude
- Attitude: pitched 19 degrees upward
- Horizontal velocity: 45 ft. per second (fps) northward
- Descent rate: -16 fps
- Attitude rate: none

In between parts of the experiment, the subjects were verbally asked what they perceived their manual landing dispersions to be, to describe their strategies in selecting LAPs, and to rate their personal risk-taking behavior on a scale from 1 to 5 (Appendix 6). Written notes on their responses were taken by the experimenter.

#### **4.3.4 Equipment and Displays**

Experiments were conducted in Draper Laboratory's lunar landing vehicle cockpit fixed-base simulator. Subjects were seated in a chair in front of three LCD screens arranged in a row at eye height, as shown in Figure 25, and interacted with the displays and simulation using a computer mouse and commercial gaming joysticks.



Figure 25: Experimenter seated in lunar landing simulator.

#### 4.3.4.1 Landing Aimpoint Selection and Manual Flying

For the training session and Part 1 of the experiment, the LCD screens displayed – from left to right – a flight display, the landing area map, and a landing rating screen, as shown in Figure 26.

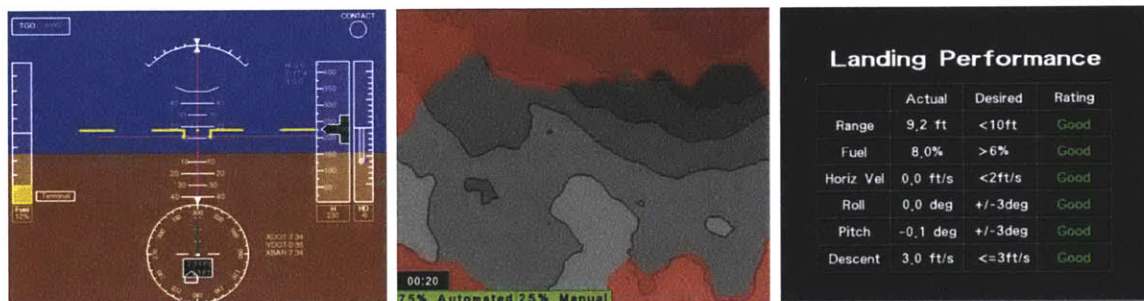


Figure 26: Displays for training and Part 1 of experiment. From left to right: flight display, landing area map, and landing rating screen.

As detailed in Appendix 3 instructions to subjects, the flight display, used by Stimpson in human subject cockpit experiments (Stimpson, 2011), provided real-time information about the vehicle to aid in manual flight. Such information included vehicle altitude (ft.), attitude (deg.), descent rate and horizontal velocities (fps), percentage of fuel remaining, and guidance cues for attitude and descent rate. In addition, the simulator provided recorded voice call-outs when 60 and 30 seconds of hover time remained before reaching the fuel safety limit.

The landing area map displayed a timer for the 20 seconds in which subjects must make a LAP selection before each flight. During that time, the flying task condition (100% or 25% chance of manual flight) was also displayed. After 20 s have expired, if the flight turns out to be manual, the vehicle began a flight trajectory from its initial dynamic conditions. A right-hand computer mouse was used to click on the map to select a LAP.

At the end of each manual flight, the rating screen showed the landing error and whether any safety limits were violated at vehicle contact.



Subjects controlled vehicle attitude with a right-hand joystick (Saitek Cyborg Evo). The mode of flight control was Rate Control – Attitude Hold, used in the Apollo lunar module (Hackler et al., 1968). A forward deflection of the joystick would result in the simulated vehicle pitching forward at a constant rate, which increases the rate of acceleration in the forward direction. (Note that the control system was modeled with a 2 ms time delay. This, combined with a large vehicle moment of inertia and thruster control power limits, resulted in approximately fourth order manual control – control of attitude acceleration rather than rate – when large attitude rates were commanded (Hainley, 2011)). If the subject lets go of the joystick, the vehicle’s attitude is held. Subjects were only required to control vehicle attitude in the pitch and roll axes; yaw was not exercised. Vehicle descent rate was controlled using a left-hand throttle button (Saitek X52), which could be clicked to increase or decrease the descent rate in increments of 1 fps.

For more information on the displays and use of the joystick and throttle button, see subject training slides in Appendix 3.

Vehicle dynamics were provided by a system model developed by Hainley for human subject cockpit experiments (Hainley, 2011), based on an early lunar vehicle design for the Constellation program (Duda et al., 2009). The model takes as input the starting location of the vehicle and commanded descent, roll, pitch, and yaw rates from the pilot. The descent engine was modeled to be fixed-gimbal. When the vehicle was commanded at a non-vertical attitude, the thrust was proportionally increased to maintain the commanded descent rate. Therefore, larger attitudes resulted in larger horizontal accelerations. Fuel usage was calculated based on the descent engine’s thrust output and specific impulse. The model’s outputs – vehicle fuel consumption, location, velocity, and attitude over time – were fed to the flight display.

Flying guidance cues for descent rate and attitude were generated by a slightly modified version of the automated guidance system in Hainley’s model, which calculated a reference trajectory based on guidance equations developed by Bilimoria (Bilimoria, 2008). The reference trajectory was calculated so that the horizontal velocity was proportional to the range to the selected landing point and the descent rate decreased linearly until the vehicle was at a near hover within a 150-ft. altitude and 22-ft. range of the LAP.

#### ***4.3.4.2 Scatter Plots of Synthesized Landing Errors***

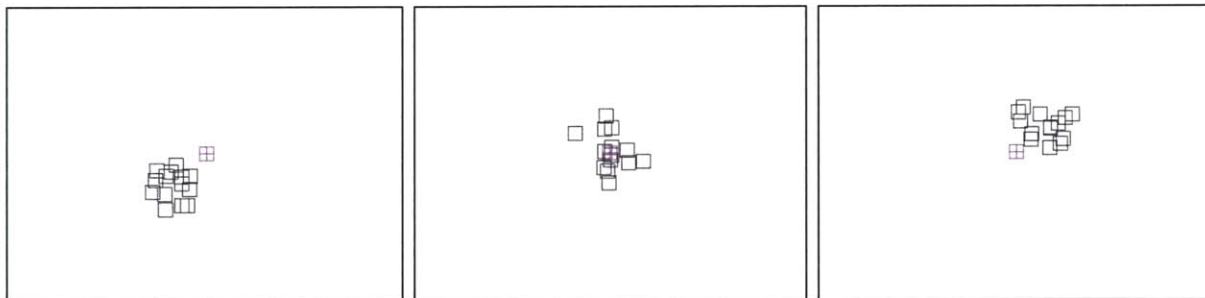
In Part 2 of the experiment, the left screen displayed a scatter plot of synthesized landing errors and the center screen displayed a landing area map as shown in Figure 27. The right screen was unused.

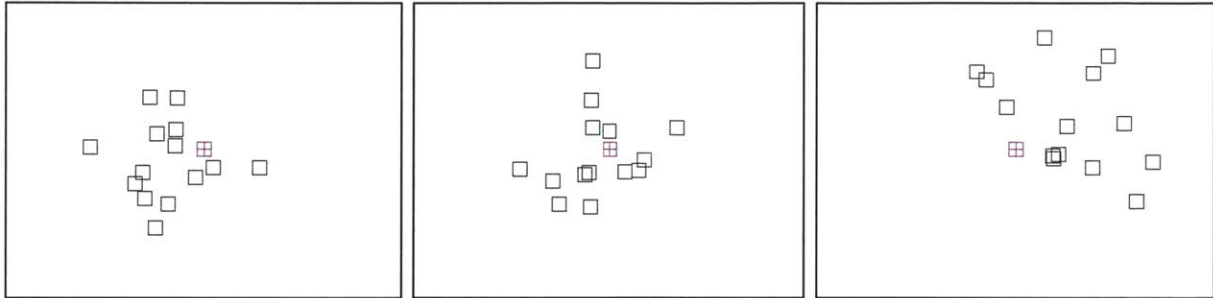


**Figure 27: Displays for Part 2 of experiment. From left to right: scatter plot of synthesized landing errors and landing area map.**

In each trial, the subject had 20 s to select a LAP that will maximize total future landing scores for a pilot whose landing performance is characterized by the scatter plot, for a given map and flying task allocation. The subject was not required to perform the task of flying. Verbal instructions read to the subject for Part 2 can be found in Appendix 5.

Six scatter plots were used in the experiment. Each contained 14 synthetic landing locations, relative to a LAP (corresponding to the 14 manual landings the subject performed in Part 1 of the experiment). The error of each landing from the LAP in the north-south (NS) (i.e., long-short) and east-west (EW) directions were derived from a normal distribution with a mean of -2, 0, or 2 vehicle widths (mean radial landing error =  $2\sqrt{2}$  to the southwest, 0, or  $2\sqrt{2}$  to the northeast) and a standard deviation of 1 or 3 vehicle widths. The six different scatter plots are shown in Figure 28 below. The SDs of 1 and 3 were chosen to create two visually distinct patterns of landing errors; scatter plots with SD=1 contained landings that were noticeably more consistent and clustered than those with SD=3. To help put these synthetically generated scatter plots into context before running Part 2 of the experiment, the subject was shown the scatter plot generated from his landing errors in Part 1.





**Figure 28: Scatter plots of synthesized landing errors. Black squares represent where landings are located relative to a target LAP (magenta square with crosshairs). From left to right: mean landing deviation = -2, 0, and 2 vehicle widths in both north-south and east-west directions. Top row: SD of landing deviations = 1 vehicle width. Bottom row: SD of landing deviations = 3 vehicle widths.**

#### 4.3.5 Subjects

Eleven subjects were recruited for this experiment (2 female, 9 male), ranging in age from 19 to 34 (mean age = 23). Five subjects had some prior experience with computer-based flight simulators, and 2 had experience with lunar landing simulators. No subjects have had previous flight experience. All subjects gave informed consent in accordance with the MIT Committee on the Use of Humans as Experimental Subjects (see Appendix 7).

#### 4.3.6 Measurement Collection and Data Calculation

The following measurements were collected at 10Hz over the course of the experiment:

- Locations of LAPs selected by subject
- Vehicle parameters during flight
  - Altitude and horizontal position (ft)
  - Descent rate and horizontal velocity (fps)
  - Pitch and roll (deg)
  - Fuel percentage remaining
- Flying guidance cues shown to subject on flight display
  - Pitch and roll (deg)
  - Descent rate (fps)

The following parameters were then calculated:

- Locations of each selected LAP relative to POI on each map
- Vehicle's final landing error relative to selected LAP
- Estimated "risk level" associated with the LAP selection, or "risk t," as explained below
- Subjects' mean square error in descent rate, pitch, and roll relative to guidance cues during each manual flight

Figure 29 shows how LAP locations and the vehicle's landing error are referenced.



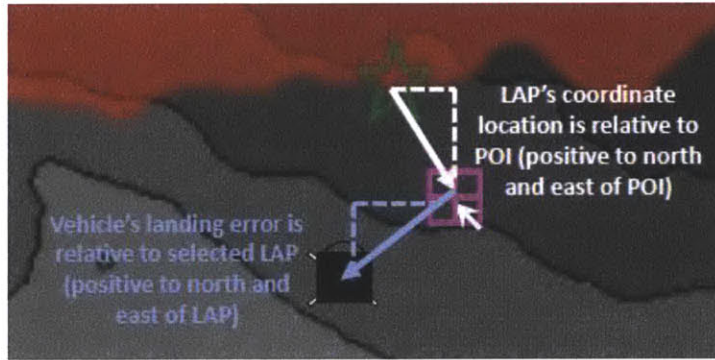


Figure 29: LAP location (crossed magenta square) and landing error (blue vector)

In this experiment, the POI was always located at the boundary of a hazard region. Each subject is assumed to have a landing error directional bias with some normally distributed dispersion around it. Hence, the risk of landing in the hazard region increases if subject places the LAP close to the POI, even if the LAP is in the non-hazard region. Quantitatively, the risk depends on the subject's directional bias and touchdown dispersion (measured in Part 1 or assumed based on the synthetic scatter plot data in Part 2). A measure of the overlap with the hazard zone was obtained, as illustrated in Figure 30, by overlaying the landing error distribution on the subject's LAP, and then computing the distance (in the north-south, along-approach-track direction) between the mean of the distribution to the edge of the hazard. This distance (labeled "x" in Figure 30) was computed in vehicle widths. To estimate the associated risk of hazardous landing, this distance was divided by the standard deviation (denoted "s" in Figure 30) of the landing dispersion. This non-dimensional ratio is referred to as "risk  $t$ " since it should be distributed as (Student's)  $t$  with  $n-1$  degrees of freedom for a landing error data sample containing  $n$  points. The more positive risk  $t$  is, the safer the landing. Conversely, if risk  $t$  is nearly zero or is negative, a hazardous area landing is more probable. Figure 30 illustrates how risk  $t$  values are calculated for two example LAP selections.



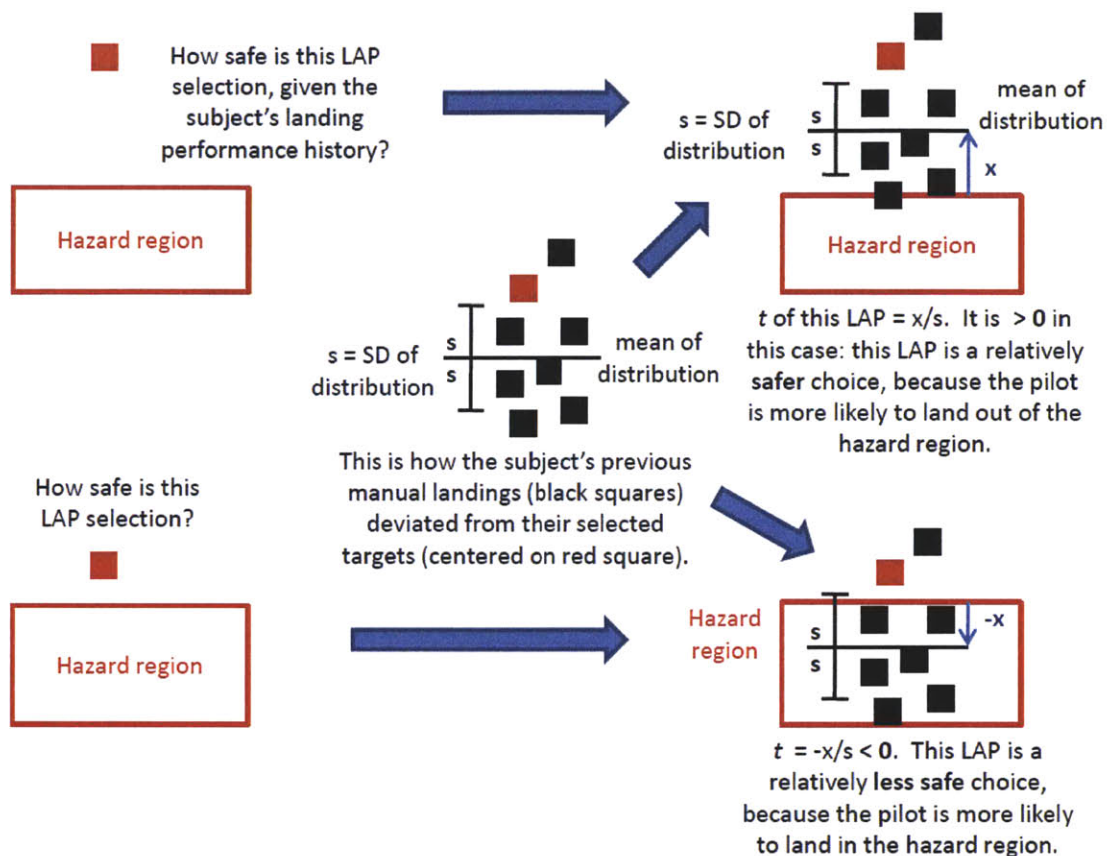


Figure 30: Risk  $t$  derived from the mean of a set of landings on a single map

For Part 1, risk  $t$ 's were calculated using the mean and standard deviation of the subject's own landing errors over all manual flights in this part of the experiment. The calculations were done separately for the two landing site locations. For Part 2, risk  $t$ 's were calculated using the known mean and standard deviation of the synthesized landing error scatter plots.

A corresponding " $t$ " parameter was also calculated for the across-track EW direction. Note that, because hazard edges are approximately horizontal, the along-track NS risk  $t$  can be interpreted as an indicator of the riskiness associated with a LAP. A positive  $t$  in the EW direction means only that the LAP was placed too far to the east given the distribution of the subject's EW landing errors; a negative  $t$  means the LAP was placed too far to the west.

Mean square error in descent rate, pitch, and roll relative to guidance cues for each manual flight was calculated by squaring the difference between actual and guidance values at all time points (measured at 10Hz) from the start of manual flight control to contact, then taking the average.

Vehicle width, 30 ft., was used as the unit of measurement for all data. All data analysis was performed using MATLAB (The MathWorks, Inc.) and Systat 13 (Systat Software, Inc.) software.

## 4.4 Results

### 4.4.1 Overview of Flying and Landing Performance in Part 1

#### 4.4.1.1 Mean Square Error, Landing Scores, and Safety Limit Violations

Table 3 below shows, among all subjects, the number of landings in Part 1 that exceeded safety limits upon contact and mean square errors in descent rate and attitude with respect to guidance cues (tallies of such landings by subject can be found in Appendix 8). Mean square errors for each manual flight were calculated from the start of the trajectory to contact.

**Table 3: Flying and landing performance over all subjects for Part 1**

Safety limits	Number of flights that exceeded safety limits at contact (11 subjects x 14 manual flights per subject= 154)	Mean square error from guidance cues = (root mean square error) <sup>2</sup>
Descent rate $\geq 3$ fps	9 / 154 (4 subjects)	(2.2) <sup>2</sup>
Pitch $\leq \pm 6$ deg.	12 / 154 (5 subjects)	(11.8) <sup>2</sup>
Roll $\leq \pm 6$ deg.	5 / 154 (2 subjects)	(2.4) <sup>2</sup>
Horizontal velocity $< 4$ fps	37 / 154 (7 subjects)	N/A
Fuel $> 5\%$ remaining	15 / 154 (5 subjects)	N/A
# of landings in hazard area	30 / 154 (10 subjects)	N/A

Landings that exceeded safety limits were included in the analysis, because subjects did take these landings into account when building an understanding of their landing dispersions in Part 1. Such landings give insight into which aspects of manual control subjects found the most difficult. The task of nulling the vehicle's horizontal velocity at contact to within safety limits was apparently the most difficult aspect of the flying task. Of the other three quantities – roll, pitch, and descent rate – on which subjects controlled during the flying task, they exhibited the greatest mean square error from guidance cues on pitch control. This is to be expected since the task primarily involved pitch maneuvers to slow the vehicle along the direction of flight and null horizontal velocities prior to touchdown.

Subjects achieved an average score of -2.2 per landing (SD of average scores among all subjects = 5.0). Since a violation of a safety limit earns -10 points, and it is extremely difficult to reach a score of 10 for a successful landing (highest individual landing score = 6.9), it is not surprising that the average landing score was negative. The lowest scoring subject (average score = -9.2) had 10 out of 14 manual flights that violated at least one safety limit; the highest scorer (average score = 5.4) had no flights that violated safety limits. A mixed hierarchical regression found a significant effect of the landing area map condition on score ( $p = 0.026$ ); scores that equaled -10 were excluded; residuals of fit were normally distributed and had stable variances). On average, scores were 1.2 points lower for the long map. One reason could be that, in the long map, the POI is located farther south than the POI in the short map (see Figure 23). Since the vehicle's starting location is south of the map (Section 4.3.3), LAPs selected for the long map would result in shorter flight trajectories, allowing less time for the pilot to close guidance errors or steer clear of the hazard region before contact. A significant effect of map is also reflected on risk  $t$ 's (LAP  $t$ 's in the NS direction), as shown in Section 4.4.2. There was no significant effect of flying task condition, or a cross-effect of map and flying task condition, on score.

#### 4.4.1.2 Landing Errors

There was no significant effect of the landing area map condition on landing error. The effect of map was examined for 22 combinations of subject and EW/NS directions (11 subjects x 2 directions = 22) by a two-sample t-test for each. Those differences were significant in 4 of the 22 cases (listed in Appendix 9), which is approximately what would be expected for a test applied at  $p=0.05$ . That number of significant results is also not significant according to the Sign test, which is consistent with the conclusion that there is no effect and that any apparent effect is a false positive. (Typically, one would apply a hierarchical regression with several independent variables to the analysis of data like this. That hierarchical fit was performed, but the results did not meet the requirements of normal distribution and stable variance among the residuals. As an alternative, less general methods were applied.)

Subjects' average landing error relative to the selected LAP for each flight in Part 1 was  $0.46 \pm 0.87$  (mean  $\pm$  SD) vehicle widths to the west and  $0.37 \pm 1.2$  vehicle widths to the south. Most subjects tended to land south-west of targeted LAPs, as seen in Figure 31, except for 2 subjects whose landing errors were substantially larger and had a different directional bias than the others.

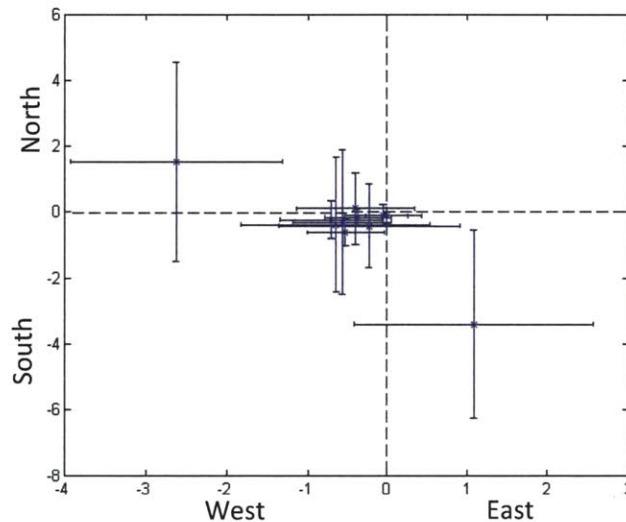


Figure 31: Mean landing errors for each subject. Error bars indicate SD in EW and NS directions. Units = vehicle lengths. Each data point contains landings for both landing area maps, as no significant effect of map was found on landing error.

Finally, landing accuracy remained consistent over Part 1 of the experiment; no significant effect of trial number on EW/NS landing error was found using linear least squares regression. Therefore, subjects were sufficiently trained so that their landing accuracy did not significantly increase during Part 1.

#### 4.4.2 Landing Aim Point $t$ 's

There was no effect of trial number on LAP  $t$ 's in the EW or NS direction in Part 1 of the experiment (for Part 1, there was a possible learning effect since subjects acquired knowledge on their manual landing dispersions over the course of trials, but no significant effect was found).

A linear least squares regression was performed on risk  $t$ 's, calculated in both the NS, and LAP  $t$ 's, calculated in the EW direction, versus trial number. This represents 88 combinations of subject, map,



and flying task condition (11 subjects x 2 maps x 2 flying task conditions x 2 directions (NS and EW) = 88 cases). Of these,  $88 \times .05 = 4.4$  false positives are to be expected at  $p = 0.05$ . Seven false positives were found (listed in Appendix 9), which is consistent with the expected number of false positives. The test did find significant effects of trial number on LAP or risk  $t$ . Therefore, the risk levels ( $t$ ) in Part 1 of the experiment are consistent with the null hypothesis that the risk taken is constant over time.

A mixed hierarchical regression test yielded a significant effect ( $p < 0.0005$ ) of map on risk  $t$ 's ( $t$ 's in the NS direction) in Part 1 of the experiment. On average, risk  $t$ 's for the long map were 0.45 less than those for the short map, which implies that subjects' LAP selections on the long map was slightly more risky than for the short map given their landing dispersions. This is consistent with the result found in Section 4.4.1.1 that subjects scored lower on the long map. Therefore, analysis was performed separately for the two landing area maps.

Hypothesis 1: The greater the subject's landing error, the greater will be the effect of flying task condition on the risk level of selected LAPs.

It is hypothesized that there is an effect of the size of the subject's landing error on the observed effect of flying task condition. The difference between the average  $t$ 's for the two flying task conditions – i.e., the effect of flying condition for each subject – was calculated for each subject separately for  $t$ 's in the NS and EW directions. A linear least-squares regression was then performed on the effects of flying condition against landing error. In Part 1, that landing error is the subject's average landing error. In Part 2, it is the mean of the landing errors in the scatter plot shown to the subject during LAP selection.

Results for Part 1 of the experiment are shown in Figure 32 below. A significant slope (in units of 1/ft.) was found for both the short ( $p=0.015$ ) and long ( $p=0.005$ ) maps in the NS direction (the deviations from the best-fit line were normally distributed and satisfied Levene's test). The large errors in the calculated slopes imply that they are not significantly different between maps. However, a mixed hierarchical regression yielded a significant effect of map on the difference in  $t$ 's due to flying condition in the NS ( $p=0.038$ ) and EW ( $p=0.05$ ) directions (residuals were normally distributed, variances of residuals for each map were stable over the range of predicted values). No significant slope was found in the EW direction. This suggests that subjects are more attentive to the NS rather than to EW positioning of their LAPs in Part 1. Since the edge of the hazardous region runs horizontally east to west on both maps, it is not surprising that LAP placement and landing error in the NS direction, which is relevant to hazard avoidance, may be more salient to subjects than in the EW direction, which is irrelevant to hazard avoidance.

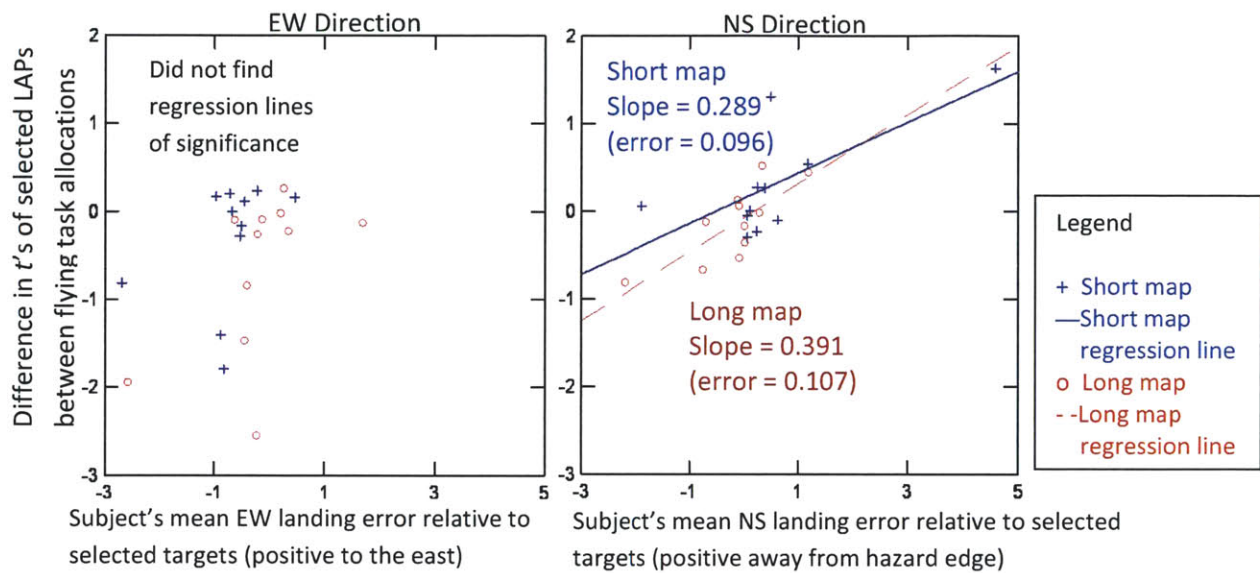
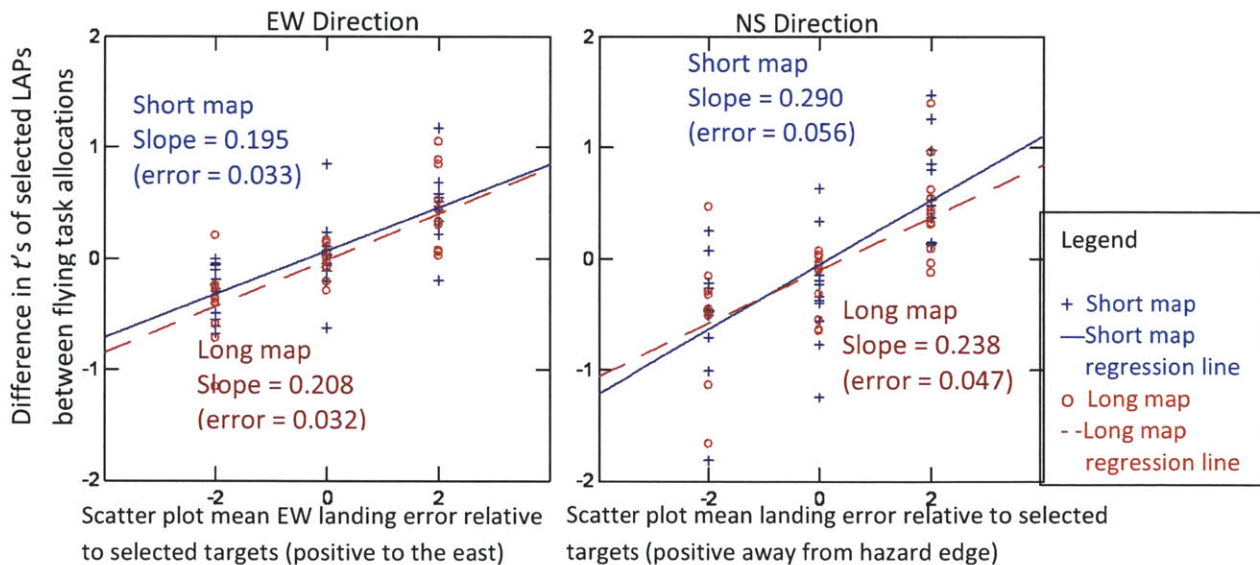
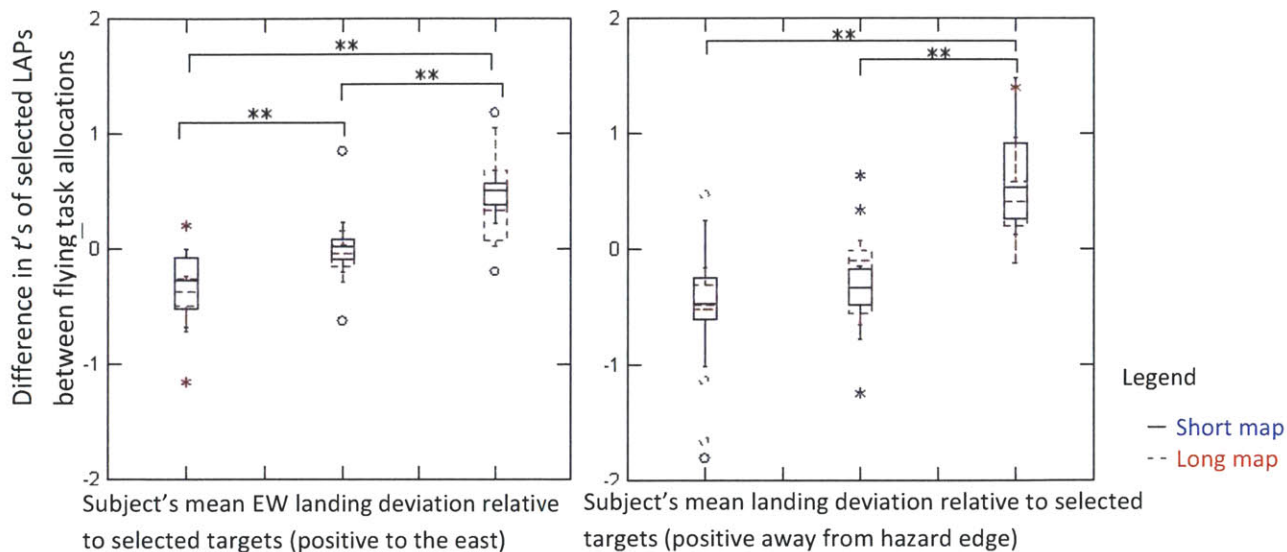


Figure 32: Effect of flying task condition on  $t$  plotted against subjects' mean landing errors in Part 1. A positive effect (difference in  $t$ ) means that LAPs selected for the 25% manual condition were safer than those selected for the 100% manual condition. Each data point contains 10 LAP selections (5 for each allocation) by one subject for one map type.

For Part 2, however, there was a significant slope in both EW and NS directions ( $p < 0.0005$ ), as shown in Figure 33. Again here, as inferred from the error in the slopes relative to their difference, there is no significant effect of map. Mixed hierarchical regressions did not yield a significant effect of map either.





**Figure 33: The effect on the t-risk parameter of flying task condition and mean synthesized landing error in Part 2. A positive effect (difference in  $t$  values) means that LAPs selected for the 25% manual flying task condition were safer than those selected for the 100% manual condition. In the two upper charts, each data point contains 8 LAP selections (4 for each flying task condition) by one subject for one map type. The lower two charts contain box plots of the same data. \*\*=The effect of flying task condition (the difference in  $t$  values) is significant ( $p < 0.05$ ) for both maps.**

The positive slopes of these effects against subjects' landing deviations can be interpreted differently in the EW and NS directions.

In the EW direction, if subjects tended to land east of the targeted LAP, they placed LAPs farther to the west under the 100% manual flying condition. However, LAPs selected under the 25% manual condition were not moved westward. The result is a positive difference in  $t$ 's when subjects had eastward landing errors, and vice versa. This occurred only when they were viewing synthesized scatter plots of landing dispersions, and not in flight, in Part 1 of the experiment.

In the NS direction, if subjects had tended to land on the side away from hazards, they placed LAPs closer to the hazard edge under the 100% manual condition. LAPs selected under the 25% manual condition did not include this compensation. The result is a positive difference in  $t$ 's when subjects' landing errors were in the direction away from the hazard edge, and vice versa. This occurred no matter how subjects were presented with landing error information.

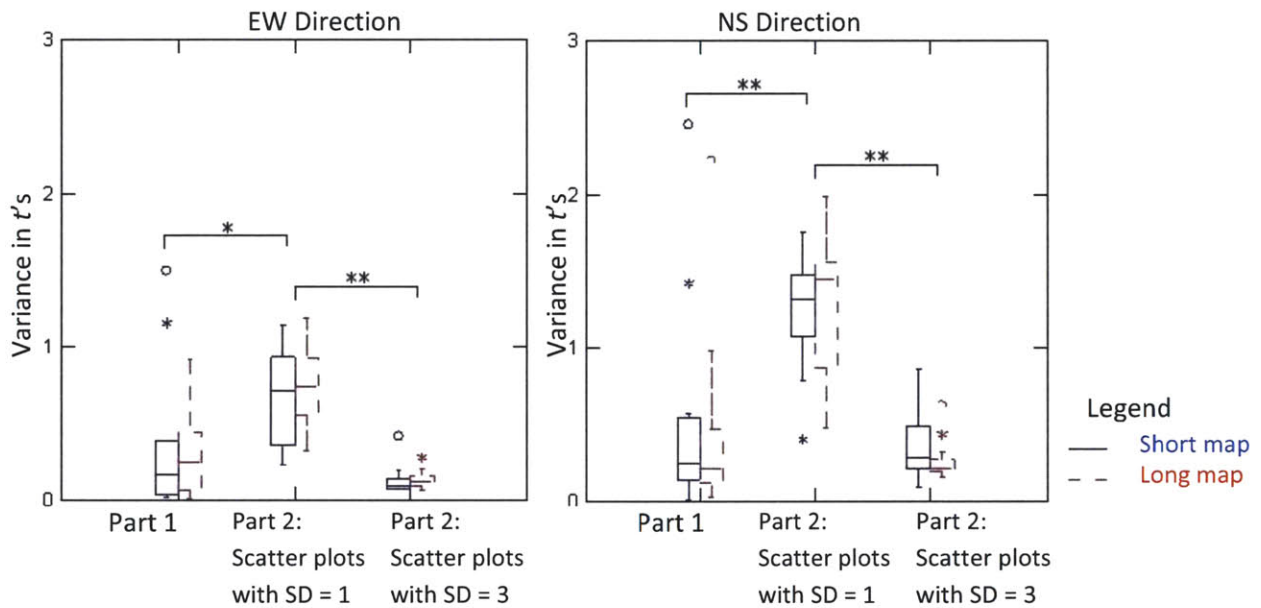
Therefore, subjects selected LAPs that compensated for their mean landing errors. The compensation was less when they knew there was only a 25% chance that the flight would be manually flown. The only case that showed no evidence of landing error compensation was in the EW direction in Part 1 of the experiment.

**Hypothesis 2:** The presentation of landing error information (successively observed landings vs. scatter plot of previous landings) affects the risk level of selected LAPs.



A paired t-test was performed between  $t$ 's of LAP selections made in Part 1 and in Part 2 of the experiment separately for the four combinations of map and direction (NS and EW). A significant difference (0.482,  $p=0.021$ ) was found between parts of the experiment only in the EW direction for the long map. Although the t-tests indicated a significant difference between Parts 1 and 2 for only one case (long map, EW direction), the results shown earlier in Figure 32 and Figure 33 suggest that scatter plots caused subjects to compensate more systematically for their landing errors in the EW direction. Therefore, there is no significant effect of the landing error information on LAP risk levels. Moreover, since Part 1 always took place before Part 2 (for all subjects) any effect observed might also be attributed to that effect of order.

The way landing error information was presented to the subject appeared to affect the consistency of  $t$ 's. The box plots in Figure 34 below show the range of subjects' variance in  $t$ 's for Part 1 and Part 2 of the experiment. In addition, for Part 2, separate variances in  $t$  are shown for the two different standard deviations (SD=1,3) used in separate sets of trials as the basis of the synthesized scatter plots.



**Figure 34: Variance in  $t$ 's of target selections vs. landing error information. Horizontal bars show significant difference by paired t-test,  $p < 0.05$ . Each box plot contains 11 data points, one for each subject (20 data points per subject for Part 1, 24 data points per subject for each Part 2 box plot). Note that the EW plot excludes an outlier subject for the short map. \*=Variances are significantly different ( $p < 0.05$ ) only for the long map. \*\*=Variances are significantly different ( $p < 0.05$ ) for both maps.**

Landing error scatter plots with the smaller SD of 1 vehicle width were apparently associated with larger variance in  $t$ 's. In fact, the variances of  $t$ 's in Part 1 are not significantly different from the variances of  $t$ 's selected using scatter plots of SD = 3 vehicle widths. One of the subjects had a very large outlier variance for the short map in the EW direction. That outlier subject, however, was not the source of the significant differences between variances (the significant results persisted when that subject was omitted). One possible explanation for this is that subjects may have been most willing to try different risk levels to maximize score when presented with highly consistent (small SD) landing dispersions, and

less so when presented with their own or inconsistent (large SD) landing dispersions. This is supported by the observation that 73% of subjects reported that they were more willing to put the LAP closer to the POI when viewing scatter plots with the smaller SD of 1 vehicle length. Understandably, these subjects felt that such distributions indicated more consistent and predictable landing performance; one subject, “assumed that they [pilots represented by landing error distributions of SD=1] were better pilots.”

#### 4.4.3 Subjective Responses

##### 4.4.3.1 *Self-Reported Landing Dispersions and Landing Point Designation (LPD) Strategy*

When asked about their landing errors, 45% of subjects were able to provide descriptions that encompassed landing errors for 100% of their manual flights in Part 1; subjects usually described their landings errors as being “no more than” and/or “no less than” some distance from selected LAPs. The other 55% provided descriptions that appeared to exclude a portion of their flights; 27% excluded the worst 16% (SD = 11%) of their flights, 18% performed better than stated, and 9% performed much worse than stated (i.e., all the actual flights were excluded by their description).

All subjects stated that they used different LAP selection strategies for the two flying task conditions in at least one part of the experiment (55% of subjects for Part 1, 82% of subjects for Part 2). Almost always, the difference in strategy put the LAP closer to the POI, which lies on the edge of a hazard region, under the 25% manual condition. 18% of subjects used the opposite strategy: in the 100% manual condition, they put the LAP closer to, or even into, the hazard zone to compensate for expected landing errors. Then, for the 25% manual condition, they made sure to keep the LAP clear of the hazard region in case of a precision automated landing.

##### 4.4.3.2 *Subsidiary Results*

A mixed hierarchical regression test was performed to find effects of the subjects’ self risk ratings on both  $t$  and the difference in  $t$  due to flying task condition (the map condition was included as an independent variable to account for effects not attributed to the risk rating). A significant effect of risk rating was found only on  $t$ 's in the EW direction for Part 1 ( $p = 0.005$ , residuals of fit were normally distributed and had stable variances, no significant effect of map was found). The higher the subject’s risk rating, the further to the east they placed their LAPs. This result is counter-intuitive, as most subjects tended to have westward landing errors (Figure 31); if anything, subjects who judged themselves to be more conservative (lower risk rating) should have placed their LAPs further to the east in compensation, rather than subjects who had higher risk ratings. Ratings are listed in 0.

When shown the northern and southern POI maps (Figure 24) in the training session, subjects selected LAPs as close to the POI as possible most of the time. However, 36% (4 out of 11) of subjects chose LAPs on the other side of the hazard island 28% of the time for the northern POI map, and 18% (2 out of 11) of subjects did so 33% of the time for the southern POI training map.

Operationally, there are possible advantages to sacrificing proximity to a POI when presented with the terrain situations in these two maps. In the northern POI map, flying over the hazard island to land close to the POI would consume more fuel, and hence decrease available hover time for a precision landing.



In the southern POI map, one risks overshooting the POI and landing in the hazard region if the vehicle's northward velocity is not nulled in time (the vehicle always approaches the landing area from the south). Therefore, there may be an advantage in planning to overshoot the hazard island completely. However, if it is a scientific point of interest, there is the risk of contaminating the site with exhaust by flying over it. Also, the surface operations costs of landing across a hazard from the POI need to be considered, such as the amount of time and resources required to traverse through or around the hazard region to reach the POI.

#### 4.5 Conclusion and Discussion

The most significant result revealed by this experiment was that human operators systematically altered their LPD task performance according to system-level knowledge of the probabilities of manual versus automated flight. Subjects selected LAP locations that compensated for flying task condition, and the magnitude of that compensation was correlated with the size of subjects' own landing errors (Figure 32) or of synthesized pilot landing errors that were shown to them (Figure 33). (Note that when the landing errors are the subjects' own, the compensation was only along the NS direction – the direction of flight – as seen in Figure 32.)

There was no strong evidence of loss aversion or subjective probability biases described in Section 4.2.2. Had there been, it would have been expected that subjects would not have accounted for different flying task conditions in a way that is related to the sizes of human vs. manual landing errors, or they would have been insensitive to flying condition altogether. It is possible that in a real-life lunar landing scenario with higher stakes, astronauts may rely heavily on human risk and reward biases; rather than optimize expected reward, they might be more likely to plan for the worst-case scenario no matter what they know or estimate the probability of manual or automated flight to be. Here lies a limitation of this experiment: subjects were specifically instructed to maximize expected reward (proximity to POI over all landings) and were strongly cued to pay attention to the implications of different task allocations. This was done to elicit more variable decision making in the context of this experiment, but may not be representative of the “look and feel” of the stakes in a real lunar landing mission.

Memory heuristics, discussed in Section 4.2.1, may have had a small influence on subjects. There was no net significant effect of the way in which subjects acquired landing error information. However, there were other effects: subjects paid more attention to compensating in a less critical direction (EW) when making decisions based on scatter plots of landing errors (Figure 33), and the way in which landing error information is presented affected the consistency with which subjects adhered to a risk level in making LAP selections (Figure 34). This suggests that when relying on memory alone, subjects retain critical information such as landing error in the NS direction. When receiving information in a way that is not subject to memory heuristics, subjects appear to incorporate less salient information as well (landing error in EW direction, size of landing error distributions). The operational benefit of pilots incorporating as much information as possible at the moment of decision is clear; in this experiment, scatter plots allowed subjects to more fully compensate for different task allocations (EW as well as NS direction) and with varied sensitivity (SD of scatter plot affected consistency of  $t$ 's).



## 5 Conclusion

Simulation has the advantage of rapidly traversing an entire task allocation design space and providing high-level trends in system performance in response to different allocations. The simulation implemented in this project, with its current level of fidelity and for a given landing scenario, predicted that fuel and landing accuracy may be optimized if decision making tasks are allocated to the human operator and manual tasks such as control of vehicle dynamics are automated. The simulation also has the potential to predict which specific subtasks contribute to variability in fuel usage and landing accuracy.

However, ranges of system performance are lost when aspects of human performance simply cannot be modeled. Therefore, human subject experiments were conducted on a subset of the tasks modeled in simulation. It was observed that subjects are capable of incorporating system-level information – the allocation of the flying task – and making LPD compensations that were proportional to the difference in landing performance between human and automation. Also, subjects were sensitive to the way in which landing performance information was presented; they made LPD compensations only along the NS direction, the direction of flight, when estimating their own landing dispersions from manual flights, but made compensation in both NS and EW directions when presented with dispersions in plots. They were even found to be sensitive to the visual distribution of landing information in the plots; subjects exhibited the greatest variance in the risk level of their LAP selections when presenting with plots of tightly-clustered landing dispersions.

The work presented in this project is a functional demonstration of how computer simulation and human experiments can be used in tandem to analyze diverse aspects of task allocation design. Simulation provides broad sweeps of a full task allocation design space that would be extremely difficult to perform through live experiments. Moreover, experiments yield unpredicted or un-modeled human behaviors that need to be considered in making the task allocation decision.

## 6 Future Work

Individual future work for the simulation and experimental aspects of this project are discussed in Sections 6.1 and 6.2 below, respectively. The major next step for this project as a whole would be to incorporate human behaviors observed from the experiment into the simulation's human performance model of the LPD free selection task (subtask 3 in Figure 7). This closes the design loop of implementing a simulation model, running experiments to uncover human behaviors not accounted for in simulation, and modifying the simulation model to capture those behaviors.

Other future work items in tying simulation and experiments together are as follows:

- Implement the other subtasks modeled in simulation as human subject experiments to uncover un-modeled human behaviors in the performance of those subtasks. For example, how does visual perception of an OTW view affect the decision of whether or not to allow the automated system to select a LAP (subtask 1 in Figure 7)? How does actual human flying performance

adhere to the control loops used to model human performance of the same task (subtask 4 in Figure 7)?

- Model human and automated fault detection and error handling behavior. The human traits of flexible decision making and judgment may help keep system performance high when off-nominal situations are encountered; however, such situations are difficult to predict and model. One way to do this may be a backwards approach: implement off-nominal situations as experiments, and then develop human performance models that fit experimental data.

## 6.1 Future Work for Task Allocation Simulation

- Obtain valid ranges for parameters of human performance models, which allow for meaningful data analysis of system performance in response to human performance parameters. Validate structure of those models, using human subject experiments.
- Improve the fidelity of human performance models by implementing improved visual pattern search, and attention that switches among information signals according to task context. Models of different human decision making strategies can also be implemented, which include satisficing, minimax, and lexicographic ordering, among others (Lehto, 1997) .
- Preliminary modeling of cross-task aspects of human behavior such as memory and situational awareness, workload, and the performance of parallel tasks.
- Assemble a MATLAB/Simulink library of commonly used human cognitive behaviors that can be modified and used in a broad range of human performance modeling. This facilitates the formulation of the human as a component within a system model to allow for system-level analysis.
- Perform Monte Carlo simulation over human model parameter space and sensitivity analysis of multiple system-level metrics according to those parameters.

## 6.2 Future Work for Human Subject Experiment

One way to extend the human subject experiment is to explore subjects' responses to a continuum of task allocation probabilities, such as including 25% and 50% chances of automated flight. In this experiment, subjects produced a difference in  $t$  in response to a 75% chance of automated flight; would a proportion of that difference still appear with other probabilities of automated flight? Is there a probability of automation below which subjects would be insensitive to different task allocations?

An operational limitation of this experiment is the lack of a realistic OTW view. Free selection of a LAP is performed when the decision has been made to not use the automated system in LPD, and therefore takes advantage of the human pilot incorporating an OTW view in addition to the DEM (Figure 21). An experiment that provides subjects with this additional information source would uncover complexities in decision making related to visual perception of a terrain, responses to conflicting information between a DEM and OTW view, and the workload of perceiving two rich visual information sources.



## 7 References

- Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, 51(4), 355-365.
- Archer, R., Lebiere, C., Warwick, W., Schunk, D., & Biefeld, E. (2002). Integration of task network and cognitive models to evaluate system design. *1<sup>st</sup> Conference of the U.S. Army Research Laboratory Collaborative Technology Alliances (CTA's)*,
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 47) Academic Press.
- Bennett, F. V. (1972). *Apollo experience report - mission planning for lunar module descent and ascent* No. NASA TN D-6846). Houston, TX: National Aeronautics and Space Administration.
- Bilimoria, K. D. (2008). Effects of control power and guidance cues on lunar lander handling qualities. Paper presented at the *AIAA Space 2008 Conference*, San Diego, CA.
- Byrne, M. D., Kirlik, A., Allard, T., Foyle, D. C., Hooey, B. L., Gluck, K. A., et al. (2008). Issues and challenges in human performance modeling in aviation: Goals, advances, and gaps. *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*, , 52(13) 926-929.
- Chapanis, A., Frick, F. C., Garner, W. R., Gebhard, J. W., Grether, W. F., Henneman, R. H., et al. (1951). *Human engineering for an effective air-navigation and traffic-control system*. Washington, D. C.: National Research Council, Committee on Aviation Psychology.
- Chua, Z. K., & Major, L. M. (2009). Task modeling for lunar landing redesignation. *AIAA Infotech@Aerospace Conference*, Seattle, WA.
- Cohanim, B. E., Fill, T. J., Paschall, S. C., II, Major, L. M., & Brady, T. (2009). Approach phase V considerations for lunar landing. *2009 IEEE Aerospace Conference, March 7, 2009 - March 14*,
- Connelly, M. N., & Willis, J. E. (1969). *The development of a human effectiveness function allocation methodology (HEFAM)* No. SHM-70-11)
- Duda, K. R., Johnson, M. C., & Fill, T. J. (2009). Design and analysis of lunar lander manual control modes. *2009 IEEE Aerospace Conference*, Big Sky, MT.
- Epp, C. D., Robertson, E. A., & Brady, T. (2008). Autonomous landing and hazard avoidance technology (ALHAT). *2008 IEEE Aerospace Conference*, Big Sky, MT.
- Fitts, P. M. (1962). Functions of man in complex systems. *Aerospace Engineering*, 21(1), 34.



- Forest, L. M., Cohanin, B. E., & Brady, T. (2008). Human interactive landing point redesignation for lunar landing. *2008 IEEE Aerospace Conference*,
- Forest, L. M., Kessler, L. J., & Homer, M. L. (2007). Design of a human-interactive autonomous flight manager (AFM) for crewed lunar landing. *2007 AIAA InfoTech at Aerospace Conference*, Rohnert Park, CA. , 1 64-71.
- Foyle, D. C., & Hooey, B. L. (Eds.). (2007). *Human performance modeling in aviation*. Boca Raton, FL: CRC Press/Taylor & Francis Group.
- Foyle, D. C., Hooey, B. L., Byrne, M. D., Corker, K. M., Deutsch, S., Lebiere, C., et al. (2005). Human performance models of pilot behavior. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, Santa Monica. 1109-1113.
- Hackler, C. T., Brickel, J. R., Smith, H. E., & Cheatham, D. C. (1968). *Lunar module pilot control considerations* No. NASA TN D-4131)NASA.
- Hainley, C. J., Jr. (2011). Lunar landing: Dynamic operator interaction with multi-modal automation systems. (M.S., Massachusetts Institute of Technology).
- Hollnagel, E., & Bye, A. (2000). Principles for modelling function allocation. *International Journal of Human-Computer Studies*, 52(2), 253-265.
- Jenkins, O. C., Mataric, M. J., & Weber, S. (2000). Primitive-based movement Classification for humanoid imitation. *Proceedings, First IEEE-RAS International Conference on Humanoid Robotics (Humanoids-2000)*,
- Jordan, N. (1963). Allocation of functions between man and machine in automated systems. *Journal of Applied Psychology*, 47(3), 161.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4), 341-350.
- Klumpp, A. R. (1974). Apollo lunar descent guidance. *Automatica*, 10, 133-146.
- Laughery, K. R., Jr., Lebiere, C., & Archer, S. (2006). Modeling human performance in complex systems. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (3rd ed., pp. 967). Hoboken, N.J.: John Wiley.
- Lehto, M. (1997). Classical decision theory. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd ed., pp. 1207). New York: John Wiley & Sons, Inc.

- Madni, A. M. (1988). HUMANE: A knowledge-based simulation environment for human-machine function allocation. *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference: NAECON*, Dayton, OH., 3 860-866.
- Marsden, P., & Kirby, M. (2005). Allocation of functions. In N. A. Stanton, A. Hedge, K. Brookhuis, E. Salas & H. Hendrick (Eds.), *Handbook of human factors and ergonomics methods* (). Boca Raton, FL: CRC Press LLC.
- McRuer, D. T., & Krendel, E. S. (1974). *Mathematical models of human pilot behavior* No. AGARD-AG-188)North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development.
- Mindell, D. A. (2008). *Digital apollo: Human and machine in spaceflight*. Cambridge, MA: MIT Press.
- NASA. (1971). *Apollo operations handbook: Lunar module (LM 11 and subsequent) vol. 2 operational procedures* No. LMA-790-3-LM-10-REV; LMA790-3-LM-11; NASA-CR-115269)NASA Johnson Space Center.
- Needham, J. M. (2008). *Human-automation interaction for lunar landing aimpoint redesignation*. Unpublished M.S., Massachusetts Institute of Technology, Cambridge, MA.
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 30(3), 286-297.
- Pisanich, G. M., & Corker, K. M. (1995). A predictive model of flight crew performance in automated air traffic control and flight management operations. *Proceedings of the Ohio State 8<sup>th</sup> International Symposium on Aviation Psychology*, Columbus, OH. 335-340.
- Sheridan, T. B. (2002). *Humans and automation: System design and research issues*. New York: John Wiley & Sons, Inc.
- Sheridan, T. B., & Parasuraman, R. (2000). Human versus automation in responding to failures: An expected-value analysis. *Human Factors*, 42(3), 403-407.
- Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. Cambridge, MA:
- Stanton, N. A. (2006). Hierarchical task analysis: Developments, applications and extensions. *Applied Ergonomics*, 37(1), 55-79.
- Stimpson, A. J. (2011). Design and evaluation of an achievability contour display for piloted lunar landing. (M.S., Massachusetts Institute of Technology).

- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.
- Tyler, S. W., Neukom, C., Logan, M., & Shively, J. (1998). The MIDAS human performance model. *Proceedings of the Human Factors and Ergonomics Society, 42<sup>nd</sup> Annual Meeting*, Chicago, IL., 1 320-324.
- Wen, H. Y., Duda, K. R., Slesnick, C. L., & Oman, C. M. (2011). Modeling human-automation task allocations in lunar landing. Big Sky, MT.
- Wickens, C. D., & Carswell, C. M. (2006). Information processing. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (3rd ed., pp. 111). Hoboken, NJ: John Wiley & Sons, Inc.
- Wickens, C. D., & Hollands, J. G. (2000). Decision making. *Engineering psychology and human performance* (3rd ed., ). Upper Saddle River, NJ: Prentice Hall.
- Wickens, C. D., McCarley, J., & Thomas, L. (2003). Attention-situation awareness (A-SA) model. *Proceedings of the 2003 Conference on Human Performance Modeling of Approach and Landing with Augmented Displays*, 189-225.



# Appendix 1 Preliminary Lunar Landing Task Analysis

H. Y. Wen, 10/21/09

**References:**

Numbers in brackets refer to the procedure number in the appropriate section of the "Apollo Lunar Operations Handbook: Lunar Module, LM 11 and Subsequent," Vol. II. Braking: Section 4.10.2.1. Approach: Section 4.10.2.2. Landing: Section 4.10.2.3  
 "AFM" = "Design of a Human-Interactive Autonomous Flight Manager (AFM) for Crewed Lunar Landing" by L. M. Forest, L. J. Kessler, and M. L. Homer  
 "DA" = "Digital Apollo" by David Mindell

KEY (Landing Phase)		
OVERARCHING TASKS AND DISPLAYS (Tasks and display info that apply to the entirety of the landing phase)		
CHRONOLOGICAL TASK SEQUENCE		
(Task 1)	(Task 2, also an AFM TARGET CONDITION which requires crew approval before starting)	(Task 3, also an AFM [Autonomous Flight Manager] FUNCTION)
(Subtask or display info related to Task 1)	(Subtask or display info related to Task 2)	(Subtask or display info related to Task 3)
<b>OVERARCHING TASKS FOR ENTIRE BRAKING-APPROACH-LANDING SEQUENCE</b>		
Can always enter manual throttle control	AFM FUNCTION: Monitor for program alarms [Braking 16, 22, 25; Approach 2; Landing 2] (AFM p. 7)	AFM FUNCTION: Monitor for abort situation [Landing 9], or such as if landing radar did not come in, or came in with too much of a difference from inertial navigation (DA p. 201), or because of program alarms (DA p. ~201) (AFM p. 7)
Display info: desired auto throttle % (do not exceed), H_dot, H		Decide to abort using descent or ascent state (DA p. 200)

BRAKING PHASE CHRONOLOGICAL TASK SEQUENCE					
Preparation	AFM TARGET CONDITION: Enter P63 [9,10]	Get into attitude for trimming engines [12]	Monitor attitude maneuver [13]. Visually cross-check attitude with out-of-window view [14].	Check landing radar position and settings. Ensure PGNCs is in automatic mode [14]	Monitor ullage maneuver [17]
Prepare subsystem settings [1,2]	Can change minimum altitude at which landing radar stops working [9]	Can also be done manually [12]			Display info: thrust %, engine switches in correct position [17]
Set up, verify abort system settings [3,8, 15]	Display info: time duration of burn, time from ignition, range (do visual cross-check) [10]				
Enable, verify manual hand controller systems [4] Prepare descent engine system for braking burn [6]					
BRAKING PHASE (cont.) CHRONOLOGICAL TASK SEQUENCE (cont.)					
	AFM TARGET CONDITION: Accept or reject engine ignition for braking burn [18]	Monitor burn [19]	Turn on descent engine command override [23]	If LM is not window-up, yaw around manually or automatically to allow landing radar lock [24]	Decide whether or not to accept landing radar data [25]
	Display info: inertial velocity, time from ignition, "Delta_Vm" (?) [18]	Monitor thrust %, error in altitude between landing radar and inertial calculations, descent rate, and altitude [22] If descent thrust fails, 1) approve automatic failure routine or 2) cancel failure routine and exit braking burn (then what?) [20] Can enter manual RCAH mode [22]			Display info: horizontal range to landing site, time-to-go to engine cutoff, inertial velocity [25] Display info: "altitude, descent rate and along-track velocity" not on DSKY (DA p. 201)
		Display info: inertial velocity, rate of descent, altitude [DA p. ~200]			

APPROACH PHASE			
OVERARCHING TASKS AND DISPLAYS			
Monitor fuel levels	Can transition to manual RCAH mode (P66) at any time (DA p. 206-208)	Display info: horizontal range to landing site, time-to-go to engine cutoff, inertial velocity	AFM FUNCTION: Make adjustments to descent/landing planning constraints (AFM p. 7)
CHRONOLOGICAL TASK SEQUENCE			
Observe that we are in Approach phase [1]	AFM TARGET CONDITION: Monitor pitchover [3]	AFM FUNCTION: Redesignate landing site if necessary [4] (AFM p. 7)	
	Display info: time remaining for redesignation, H rate, H [2]	Display info: See AFM papers for redesignation display design	
	Display info: designated landing point (DA p. 204-206)		

LANDING PHASE			
OVERARCHING TASKS AND DISPLAYS			
P66 can be entered at any time	If in RCAH mode, the desired automatic attitude to null horizontal velocities is displayed on FDAI.	Monitor fuel and oxidizer levels throughout [3]	Display info available: horizontal range to landing site, time-to-go to engine cutoff, inertial velocity
CHRONOLOGICAL TASK SEQUENCE			
AFM TARGET CONDITION: Observe that we are in Landing phase [1]	Monitor descent [3]	Monitor lunar contact light. Stop engine when light comes on [9]	Confirm touchdown (DA p. 206-208). Perform touchdown activities [10]
	Monitor display info: forward velocity, H rate, H [2]		
	Manually change rate of descent, attitude/horizontal velocity if desired [3]. Can also switch back to auto nulling of horizontal velocities [6]		
	Monitor thrust % to make sure it is not above limit [4]		
	Monitor to make sure horizontal velocities are not below limit -> lose Doppler tracking [8]		

## Appendix 2 Ranges for Human Model Parameters

Values for the following parameters were all drawn from estimated ranges. Perceptual gain and bias values were estimates, and will vary by pilot and according to the landing situation. Weights on decision factors, and limit values used in decision making (such as an acceptable “max terrain slope” or “max fraction of DEM that can be blacked out or hazardous”), were drawn from a range estimated to be wide enough to allow for varieties in pilot preference and landing situations for future automated systems. Gains and time delays for the “Touchdown” task were chosen so that a stable manual flight trajectory can be completed to LAPs selected almost anywhere within the landing area used for the simulation runs in this project.

“Decide to use automated system”

- Max terrain slope: [0.05 – 0.1]
- Max fraction of DEM that can be blacked out or hazardous: [0 – 0.5]
- Min fraction of perceived OTW hazards that must also be shown on DEM hazard map: [0.5 – 1]

“LPD using automated system”

- All gains: [0.75 – 1.25]
- All biases: [-10 – 10]
- Decision weights: ratios of 1, 2, or 3 for each decision factor. All weights must add up to 1.0

“LPD without automated system”

- All gains on proximity values: [1 – 1.5]
- Gain on OTW terrain elevations: [0.75 – 1.25]
- All biases: [-10 – 10]
- Decision weights: ratios of 1, 2, or 3 for each decision factor. All weights must add up to 1.0

“Touchdown”

- Time delays: [0.3 – 0.5 secs]
- Control of change in x, y accelerations
- Gain on location error: [0.04 – 0.06]
- Gains on velocity and attitude errors: [0.4 – 0.6]




# Appendix 3 Subject Training Slides

## INTRODUCTION

Your role in this experiment is to fly a lunar lander vehicle to the surface of the Moon. This requires you to perform 2 tasks:

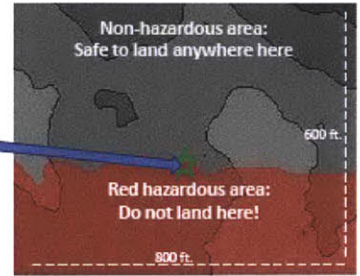
1. Select a landing target
2. Fly the vehicle down to the selected landing target



1

## INTRODUCTION (cont.)

This is the terrain area in which you will land:



Non-hazardous area:  
Safe to land anywhere here

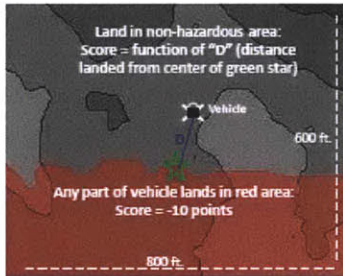
Point of scientific interest (green star):  
Land as close as possible to here

Red hazardous area:  
Do not land here!

2

## INTRODUCTION (cont.)

You will be scored according to where you land...



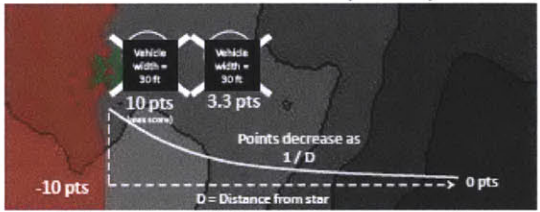
Land in non-hazardous area:  
Score = function of "D" (distance landed from center of green star)

Any part of vehicle lands in red area:  
Score = -10 points

... and according to whether you land within safety limits for fuel, vehicle attitude, and vehicle velocity (more on that later).

3

## INTRODUCTION (cont.)



Vehicle width = 30 ft

Vehicle width = 30 ft

10 pts (max score)

3.3 pts

Points decrease as  $1/D$

-10 pts

0 pts

D = Distance from star

Maximize your score by landing as close to green star as possible, without vehicle body intersecting the red region!

You will repeat this landing task multiple times in this experiment. Your goal is to maximize your total landing score, summed over all trials.

4

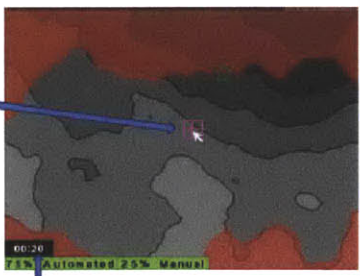
Now, you will learn how to perform the 2 tasks needed to land the vehicle:

1. Select a landing target
2. Fly the vehicle to the landing target

5

## 1. SELECTING A LANDING TARGET

- Use the mouse to click on the terrain map. A purple square will appear.
- When you fly, you will be provided with guidance cues to help you land your vehicle as accurately on the purple square as possible.
- You ARE allowed to place this target in the hazardous region. For example: if you notice that you tend to fly and land south of the target, you may place the target northward into the red zone here to compensate for your own landing bias.



66:20  
75% Automated, 25% Manual

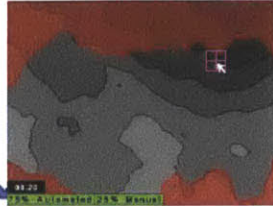
You have 20 secs. to select a landing target

6

## 1. SELECTING A LANDING TARGET (cont.)

In future lunar outpost re-supply and science missions, the landings are typically highly automated. However, manual control by the pilot may be required for the vehicle's final descent and touchdown.

If you see this, there is a 75% chance that the vehicle will be flown by an automated system, and a 25% chance that the vehicle will be flown manually. You will find out the flight mode (automated or manual) only after the 20 seconds is up.



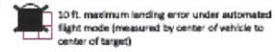
Or ...

This indicates that the vehicle will be flown to the target manually (by you, the pilot, with a joystick and throttle).



## 1. SELECTING A LANDING TARGET (cont.)

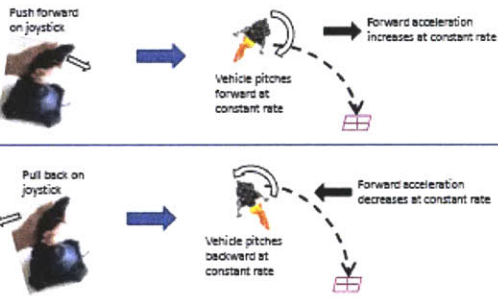
The automated system will always land on the purple square with an accuracy of 10 ft. or better.



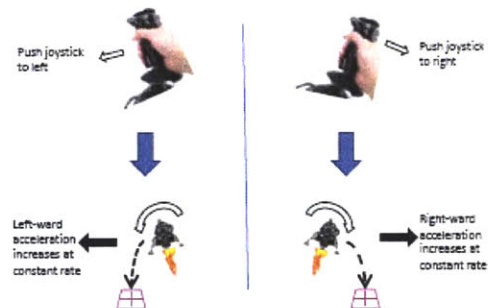
If the flight mode is manual, how accurately the vehicle lands on the purple square will depend on how accurately you can follow the guidance cues during flying.

You must maximize your total score summed over all your flights in this experiment. Since the automated system is always guaranteed to land no more than 10 ft. from your selected landing target, you may consider placing your target closer to the green star when there is a 75% chance that the flying will be automated.

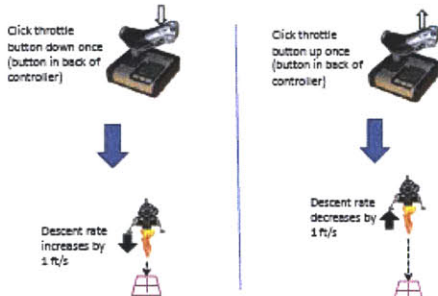
## 2. FLYING THE VEHICLE



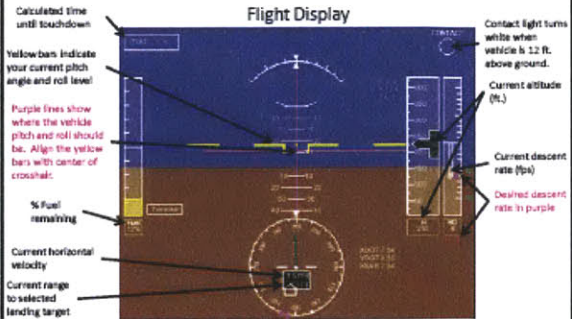
## 2. FLYING THE VEHICLE (cont.)



## 2. FLYING THE VEHICLE (cont.)



## 2. FLYING THE VEHICLE (cont.)



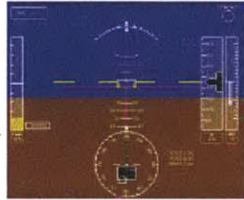
## 2. FLYING THE VEHICLE (cont.)

Do not deviate from the guidance cues provided to your selected landing target.

You will hear verbal warnings when 60 and 30 seconds of flight time remain before fuel decreases to 5% ("Bingo").

"Draper, 60 seconds till Bingo."

"Draper, 30 seconds till Bingo."



## 2. FLYING THE VEHICLE (cont.)

Rating Screen appears when vehicle has landed

### Landing Performance

	Actual	Desired	Rating
Range	9.2 ft	<10ft	Good
Fuel	8.0%	>5%	Good
Horiz. Vel	0.0 ft/s	<2ft/s	Good
Roll	0.0 deg	+/-30deg	Good
Pitch	-0.1 deg	+/-30deg	Good
Descent	3.0 ft/s	<=3ft/s	Good

At contact (12 ft. above ground), your conditions must meet the following:

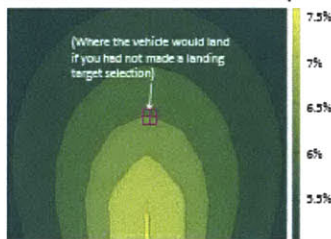
- % Fuel > 5%
- Descent rate <= 3 ft/s
- Horizontal velocity < 4 ft/s
- Roll and pitch < ±6 degrees

Otherwise, you will be deducted -10 points. Following the purple guidance cues on the Flight Display and listening to "Bingo" fuel callouts will help you meet these touchdown conditions.

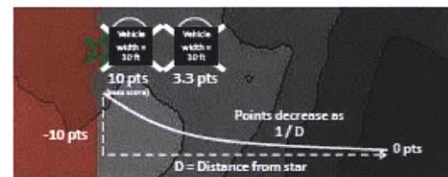
## 2. FLYING THE VEHICLE (cont.)

Remember, you will be penalized -10 points for landing with <5% fuel remaining. Below is a map of % fuel remaining when the automated system flies to different locations on the map:

Selecting a landing target closer to the bottom center of the map will provide you with more fuel – and more time – to stay in flight and improve the accuracy of your landing.



## REMEMBER



Maximize your score by landing as close to green star as possible, without vehicle body intersecting the red region!

You must maximize your total score summed over all your flights in this experiment. Since the automated system is always guaranteed to land no more than 10 ft. from your selected landing target, you may consider placing your target closer to the green star when there is a 75% chance that the flying will be automated.

10 ft. maximum landing error under automated flight mode (measured by center of vehicle to center of target)

16

## REMEMBER (cont.)

If you do not meet the following conditions at contact, you will be penalized -10 points:

- % Fuel > 5%
- Descent rate <= 3 ft/s
- Horizontal velocity < 4 ft/s
- Roll and pitch < ±6 degrees



17

## [End of training presentation]

Now, we will go into the vehicle cockpit simulator to practice selecting a landing target and – if the flight mode is manual – flying the vehicle.

18



## Appendix 4      Schedule of Trials

### Training

Trial	Map	Flying task condition shown during LAP selection (% chance of manual flight)	Flying task actually turns out to be automated (A) or manual (M)
1	short	25%	A
2	long	100%	M
3	short	100%	M
4	short	100%	M
5	long	25%	A
6	long	100%	M
7	northern POI	100%	M
8	southern POI	25%	M
9	southern POI	100%	M
10	northern POI	100%	M
11	northern POI	25%	A
12	southern POI	100%	M

### Part 1

Trial	Map	Flying task condition shown during LAP selection (% chance of manual flight)	Flying task actually turns out to be automated (A) or manual (M)
1	short	100%	M
2	long	25%	A
3	long	100%	M
4	short	25%	M
5	long	100%	M
6	short	25%	A
7	short	100%	M
8	long	25%	M
9	long	100%	M
10	short	25%	A
11	short	100%	M
12	long	25%	A
13	long	25%	M
14	short	100%	M
15	short	25%	A
16	long	100%	M
17	short	25%	M
18	long	100%	M
19	long	25%	A
20	short	100%	M

Part 2

Trial	Map	Flying task condition shown during LAP selection (% chance of manual flight)	Scatter plot shown (<mean>, <SD> in vehicle widths)	Trial	Map	Flying task condition shown during LAP selection (% chance of manual flight)	Scatter plot shown (<mean>, <SD> in vehicle widths)
1	short	100%	(-2, 1)	25	short	25%	(-2, 1)
2	long	100%	(0, 1)	26	long	100%	(0, 3)
3	short	25%	(-2, 3)	27	long	100%	(2, 3)
4	short	25%	(0, 3)	28	long	25%	(2, 1)
5	short	100%	(2, 1)	29	long	100%	(-2, 3)
6	long	25%	(2, 3)	30	short	25%	(0, 1)
7	short	100%	(-2, 3)	31	long	100%	(-2, 1)
8	long	25%	(-2, 1)	32	short	25%	(2, 3)
9	short	100%	(0, 3)	33	long	100%	(2, 1)
10	long	25%	(0, 1)	34	long	25%	(0, 3)
11	short	25%	(2, 1)	35	long	25%	(-2, 3)
12	short	100%	(0, 1)	36	short	100%	(2, 3)
13	short	100%	(2, 3)	37	short	100%	(0, 1)
14	long	25%	(-2, 3)	38	short	25%	(2, 1)
15	long	25%	(0, 3)	39	long	25%	(0, 1)
16	long	100%	(2, 1)	40	short	100%	(0, 3)
17	short	25%	(2, 3)	41	long	25%	(-2, 1)
18	long	100%	(-2, 1)	42	short	100%	(-2, 3)
19	short	25%	(0, 1)	43	long	25%	(2, 3)
20	long	100%	(-2, 3)	44	short	100%	(2, 1)
21	long	25%	(2, 1)	45	short	25%	(0, 3)
22	short	25%	(-2, 1)	46	short	25%	(-2, 3)
23	long	100%	(0, 3)	47	long	100%	(0, 1)
24	long	100%	(2, 3)	48	short	100%	(-2, 1)

## Appendix 5      Verbal instructions for scatter plot portion of experiment

Now, we will do the last part of this experiment. On this scatter plot, each of the 14 black squares represents where another pilot (let's call him or her P) landed relative to a landing target on his or her manual runs on normal terrain, with no hazards nearby, and there was no effort to land close to any point of interest (green star).

To put this into context, let me show you your own scatter plot, the one you generated in the 14 flights you just flew in the first part of this experiment: *<switch to subject's scatter plot>*

*<switch back to training scatter plot>* Now, you will select a landing target for P's next flight. (P—not you—will fly the trajectory.) Once again, you have 20 seconds and you will be shown one of the same two flight conditions: (1) With a 25% chance that P will be landing manually and a 75% chance that the flight will be automated or (2) With a 100% chance that P pilot will be landing manually. In each trial, you will be shown a scatter plot for P's prior landings and a map of the terrain P will have to land in. Pilot P always landed with greater than 5% fuel, less than or equal to 3 ft/s descent rate, had a horizontal drift less than 4 ft/s, was less than 6 degrees in pitch and roll at touchdown – they were only offset from the landing site, no other touchdown parameters were violated.

Use P's history of landing accuracy to choose a landing target that will maximize his or her total score *over all future landings in the same terrain and flight conditions*. Under manual flight, pilot P may fly slightly better than what is represented in the scatter plot since now there are hazards nearby. It is your job to estimate the likely decreased landing variance of pilot P, if any, and factor that in when selecting the landing target for them. In addition, you might choose to place the landing target closer to the green star when there is a 75% chance that the landing will be automated – therefore, giving allowing P an overall better chance of scoring higher over all trials. The landing would be more accurate under automation, if that turns out to be the landing condition, and would potentially give a higher score—with less risk of landing in the hazard zone. Remember that the automated system is guaranteed to land no more than 10 ft from your selected target. A landing in the hazard zone results in a score of -10 points. The best possible landing - as close to the star as possible without touching the hazard zone - earns a maximum reward of 10 points. Then, the points decrease by 1 / the distance from the star. Also, selecting targets closer to the center of the map allows more time for P to hover and improve the accuracy of the landing before hitting the 5% fuel limit.

Let's do this once for practice.

*Do training trial.* Do you have any questions, or would you like another practice run?



## **Appendix 6      Verbal questions between parts of experiment**

After training:

“From your flights in this training session, about how accurately do you think you’re able to land the vehicle (in feet or vehicle widths relative to the target you choose, and in what direction\*)?”

\* Whether the landing error lay to the north, south, east, or west of the selected LAP.

After Part 1:

1. “Again, about how accurately do you think you’re able to land the vehicle (in feet or vehicle widths relative to the target you choose, and in what direction)?”
2. “Can you speak a little bit about how you selected targets? Any mental rules you might have used for yourself?”

After Part 2:

1. “Can you speak a little bit about how you selected targets? Any mental rules you might have used for yourself?”
2. “How much of a risk-taker would you consider yourself, on a scale of 1-5, with 1 meaning that you avoid risks whenever possible, and 5 meaning that you greatly enjoy taking risks? When I say risk-taking, I mean your behavior during driving, whether you enjoy doing extreme sports, playing games of chance for large sums of money, or making financially risky investments.”

## Appendix 7 COUHES Forms

3/8/11

### CONSENT TO PARTICIPATE IN NON-BIOMEDICAL RESEARCH

*Lunar Landing Decision Making and Manual Flight Control  
under Different Human-Automation Task Allocations*



You are asked to participate in a research study conducted by Charles M. Oman, Ph.D. from the Department of Aeronautics and Astronautics Man-Vehicle Laboratory at the Massachusetts Institute of Technology (M.I.T.), Kevin R. Duda, Ph.D. from The Charles Stark Draper Laboratory, Inc., and Hui Ying Wen from the Man-Vehicle Laboratory at M.I.T. You were selected as a possible participant in this study because NASA and the National Space Biomedical Research Institute are interested in understanding how to best design the human-machine interface used to control the lunar lander for future lunar missions. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

#### • PARTICIPATION AND WITHDRAWAL

Your participation in this study is completely voluntary and you are free to choose whether to be in it or not. If you choose to be in this study, you may subsequently withdraw from it at any time without penalty or consequences of any kind. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

#### • PURPOSE OF THE STUDY

The goal of this experiment is to study aspects of pilot decision making and manual flying in a lunar landing mission when different types of aid from an automated system are provided. The proposed experiments are designed to understand how, during a landing, pilots 1) select a landing site with and without the computer aiding in the decision, 2) how the expectation of automation flying affects that decision, and 3) make joystick and throttle control inputs to fly a simulated vehicle in a fixed-base simulator to touchdown. The results will help in understanding the effects of different human-automation task allocations in lunar landing.

#### • PROCEDURES

If you volunteer to participate in this study, we would ask you to do the following things:

You will be seated in a cockpit simulator. You will select a landing site on a digital terrain map displayed on an LCD monitor. After you have done so, you may be asked to “fly” the simulated vehicle to the landing site you have selected. In performing these tasks, you will use a joystick, throttle, mouse pointer, and provided LCD displays.

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law.

No personal information will be collected in this experiment. All simulated flight performance data collected in this experiment will be coded to prevent the identification of the data with a specific person. All data reported in journal or conference papers will be group data or de-identified. The data will be archived when the project is completed and papers published (about 2013). No identifying information will be kept with the data

- **IDENTIFICATION OF INVESTIGATORS**

If you have any questions or concerns about the research, please feel free to contact:

Principal Investigator: Charles M. Oman, Ph.D., (617) 253-7508, [coman@mit.edu](mailto:coman@mit.edu)

Co-Investigator: Kevin R. Duda, Ph.D., (617) 258-4385, [kduda@draper.com](mailto:kduda@draper.com)

Research Assistant: Hui Ying Wen, (617) 258-2216, [hwen@draper.com](mailto:hwen@draper.com)

- **EMERGENCY CARE AND COMPENSATION FOR INJURY**

If you feel you have suffered an injury, which may include emotional trauma, as a result of participating in this study, please contact the person in charge of the study as soon as possible.

In the event you suffer such an injury, M.I.T. may provide itself, or arrange for the provision of, emergency transport or medical treatment, including emergency treatment and follow-up care, as needed, or reimbursement for such medical services. M.I.T. does not provide any other form of compensation for injury. In any case, neither the offer to provide medical assistance, nor the actual provision of medical services shall be considered an admission of fault or acceptance of liability. Questions regarding this policy may be directed to MIT's Insurance Office, (617) 253-2823. Your insurance carrier may be billed for the cost of emergency transport or medical treatment, if such services are determined not to be directly related to your participation in this study.

- **RIGHTS OF RESEARCH SUBJECTS**

You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you feel you have been treated unfairly, or you have questions regarding your rights as a research subject, you may contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T., Room E25-143B, 77 Massachusetts Ave, Cambridge, MA 02139, phone 1-617-253 6787.

## Appendix 8 Landings in Experiment that Violated Safety Limits at Contact

	Number of landings that exceeded safety limit (out of 14 manual flights in Part 1 of experiment)										
	Subj. 1	Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	Subj. 8	Subj. 9	Subj. 10	Subj. 11
Descent rate (fps)	0	0	1	0	0	4	0	0	0	1	3
Pitch (deg)	3	1	2	2	0	4	0	0	0	0	0
Roll (deg)	1	0	0	4	0	0	0	0	0	0	0
Horizontal velocity (fps)	8	1	2	5	4	7	10	0	0	0	0
Fuel (% remaining)	5	4	0	4	1	0	0	1	0	0	0
Landings in hazard zone	10	1	2	3	2	2	6	2	0	1	1

## Appendix 9 False Positives in Human Experiment Results

Cases in which landing errors were different by map (two-sample t-test, $p = 0.05$ )	
Subject	Direction
5	NS
8	EW
10	EW
11	EW

Cases in which significant correlation was found between $t$ and trial number in Part 1 (linear least squares regression, $p = 0.05$ )		
Subject	Map	Direction
4	Short	EW
4	Long	EW
11	Short	EW
5	Short	NS
6	Short	NS
7	Long	NS
8	Short	NS



## Appendix 10 Subjects' Self Risk Ratings

	Subj. 1	Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	Subj. 8	Subj. 9	Subj. 10	Subj. 11
Rating	2	1	2	3.5	3	3	4	2	1, 2*	2	3

\* Subject voluntarily gave an additional rating of his/her risk behavior within the context of this experiment.