Human-Automated Path Planning Optimization and Decision Support

M. L. Cummings[1*], J. J. Marquez[*], and N. Roy[*]

[*]*Dept. of Aeronautics & Astronautics, Massachusetts Institute of Technology, Cambridge, MA.*
Emails: missyc@mit.edu, jessica.j.marquez@nasa.gov, nickroy@mit.edu

Path planning is a problem encountered in multiple domains, including unmanned vehicle control, air traffic control, and future exploration missions to the Moon and Mars. Due to the voluminous and complex nature of the data, path planning in such demanding environments requires the use of automated planners. In order to better understand how to support human operators in the task of path planning with computer aids, an experiment was conducted with a prototype path planner under various conditions to assess the effect on operator performance. Participants were asked to create and optimize paths based on increasingly complex path cost functions, using different map visualizations including a novel visualization based on a numerical potential field algorithm. They also planned paths under degraded automation conditions. Participants exhibited two types of analysis strategies, which were global path regeneration and local sensitivity analysis. No main effect due to visualization was detected, but results indicated that the type of optimizing cost function affected performance, as measured by metabolic costs, sun position, path distance and task time. Unexpectedly, participants were able to better optimize more complex cost functions as compared to a simple time-based cost function.

Keywords: human-automation interaction; path planning; decision support systems.

1. Introduction

Path planning and re-planning are essential tasks in planning flight trajectories for pilots and air traffic controllers, robotic agents exploring Mars, and unmanned air vehicle (UAV) operators re-routing aircraft. Path planning problems typically involve large data sets including multiple and interrelated variables, conflicting trade spaces, and often significant uncertainty for problems in dynamic environments. Human cognitive limitations such as limited speed in

---

[1] Corresponding author: M.L. Cummings, Ph.D.; missyc@mit.edu
Present Address: 77 Massachusetts Ave., 33-311
Cambridge, MA 02139
Phone: 617-252-1512
Fax: 617-253-4196

computation (Fitts, 1951) and the tendency towards biased decision making (Tversky & Kahneman, 1974) suggest that some form of automation-based path planning and decision support should be used to solve such complex problems, particularly in time-pressured environments.

However, automated planners are inherently brittle in that underlying models cannot account for all potential external conditions or relevant variables which could result in erroneous or misleading decision support (Smith, McCoy, & Layton, 1997). In uncertain path planning environments, which are typically caused by incomplete maps, inaccurate models (e.g., oxygen use rate) and shifting objective functions, it is critical that some balance of human-automated planner collaboration be achieved in order to leverage the computational power of automation, as well as the experience and judgment of human knowledge-based reasoning (Klein, 1989; Rasmussen, 1983; Sheridan, 2002). In such cases, an operator must recognize discrepancies between the expected state of the world and reality (i.e., the internal map representation is not entirely accurate), and determine whether to use the decision support aid under these degraded automation conditions.

Given the need to support operators in these complex, multivariate path planning tasks with some form of automation, however recognizing that automated decision support tools also have potential limitations, the primary research goal was to better understand how humans conduct complex optimizations during a geo-spatial problem solving task of path planning under time pressure when aided with a path planning algorithm and an associated visualization.

2. Background

There have been few studies examining how, from a cognitive perspective, humans generate and optimize paths. Several studies have focused on how people solve a traveling salesperson problem, TSP, where a decision maker must find the shortest path given a certain number of cities (or nodes) with the constraint of visiting each city only once (Graham, Joshi, & Pizlo, 2000; MacGregor, Ormerod, & Chronicle, 2000; Pizlo & Li, 2003).

In general, these studies confirmed a human's intrinsic ability to perceive relative size in determining shortest distance (Gibson, 1979). However, when the given path planning task involved an objective other than shortest distance, humans could not intuitively solve these problems. These studies did not investigate how automation could either aid or possibly hinder human performance during path planning.

2.1. A Cognitive Framework for Path Planning

To better understand how and why automation could influence the path planning task, it was first important to understand the basic cognitive steps. A notional framework for the cognitive process of path planning is summarized in Figure 1. This framework was developed through interviews with military personnel and orienteering hobbyists, both with extensive path planning experience, as well as observations of expert path planners in lunar outpost simulation exercises (Marquez, 2007).

First, a person must acquire external information about the world (e.g., flat vs. steep), and as well as those variables and constraints that must be accounted for in the path planning process (e.g., time, terrain grade, maximum distance limitations due to fuel, etc.) Then a person must consider all this information in light of a set of goals (e.g., I need to reach my destination in the shortest time without exceeding steep terrain limitations). A person creates an initial plan based on his or her integration of all these variables, as well as understanding the costs, i.e., the most direct path requires a steep ascent, which requires more fuel and may take more time than a slightly longer path.

Once this initial path is created, a person can create several alternative paths or adjust sections of already-generated paths that also potentially satisfy the goals. The actions constitute human-generated sensitivity analysis on the original solution. This is a complex iterative cognitive process since the human must mentally calculate trip time in terms of distance and path grade, which effectively represents the human generation of an internal cost function.

A final path is generated once all these processes are complete, and the operator is satisfied that the solution meets some minimal (as opposed to optimal) level of acceptable criteria, which is a well established phenomenon in the decision making under certainty literature known as satisificing (Simon, et al., 1986). The next section describes how automation can be leveraged to assist in this process.

[INSERT FIGURE 1 HERE]

2.2. AUTOMATED DECISION SUPPORT IN PATH PLANNING

Automation can be introduced into any or all of these three cognitive processes in Figure 1 in order to assist the human operator. Providing automation to support the information integration process in path planning is particularly useful when there are multiple variables and complex cost models. An automated path planner can quickly calculate total path costs and generate optimized paths for the operator, which generally is very time and resource consuming for users, especially given a large and complex problem space. For example, a recent study showed that even expert maritime human path planners take minutes to solve a simple path planning problem, whereas a heuristic algorithm only takes seconds (Cummings, et al., 2010). In time-pressured settings, this can mean the difference between mission success and failure. Furthermore, such calculations save time in the sensitivity analysis cognitive phase of path planning, allowing users to modify automated solutions if they do not reflect the objectives of the users. Such interactions may potentially afford users the ability to better understand the automation and underlying complexity of variables and models.

However, automation is not the complete solution. Previous research investigating human understanding of automated path planners in the aviation domain generally indicates that removal of a human operator from the path creation process leads to complacency, over-reliance on the automation, and automation bias (Chen & Pritchett, 2001; Johnson, Ren, Kuchar, & Oman, 2002; Layton, Smith, & McCoy, 1994). In a previous space exploration-based study, when automated path planning was used, human performance improved with respect to errors and task duration, but participants also experienced decreased situation awareness (Marquez & Cummings, 2008). Over-reliance, loss of situation awareness, and automation bias are known potential issues with the introduction of automated decision support tools (Parasuraman, Sheridan, & Wickens, 2000), which are particularly relevant when considering that automation can fail and operators may not react appropriately to these circumstances (for review, see Cummings, 2004; Parasuraman & Riley, 1997). Dependence on an automated system may result in the inability to recognize when automation has degraded, and inappropriate reactions to unexpected events.

2.3. The Role of Sensitivity Analysis in Path Planning

One design intervention that can be used to prevent inappropriate automation dependence is the inclusion of sensitivity analysis tools. Little attention has been given to developing such tools. Essentially, the focus of sensitivity analysis is to assess the effect of input variables on a solution. During path planning, changes in variables can vary path costs. Consequently, human operators perform sensitivity analysis by modifying a path, evaluating the change relative to the end goal, and iterating until a satisfactory solution is achieved, as depicted in Figure 1. Sensitivity analysis includes testing the robustness of an optimal solution, investigating sub-optimal solutions, and identifying sensitive model variables (Pannell, 1997). As a result, when a human conducts sensitivity analysis, he or she is learning how to optimize a path and this, in turn, may lead to a better understanding of how a solution is affected by changes in input variables (Saltelli, Chan, & Scott, 2000).

Most of the literature related to sensitivity analysis centers on its importance for model development and validation (e.g., Lu & Mohanty, 2001; McCarthy, Burgman, & Ferson, 1995), and the mathematical methodology of conducting sensitivity analysis (e.g., Frey & Patil, 2002; Saltelli & Bolado, 1998). A few articles have advocated for sensitivity analysis tools for multi-attribute decision making aids (e.g., Jimenez, Rios-Insua, & Mateos, 2003; Triantaphyllou & Sanchez, 1997). However, how these sensitivity analysis tools support cognitive processes or enable different kinds of problem solving strategies has generally not been addressed in the literature, especially in regard to path planning problems.

One series of studies has investigated interactive optimization, a form of sensitivity analysis, where the human and automation were both involved in an optimization task (Anderson, et al., 2000; Klau, Lesh, Marks, Mitzenmacher, & Schafer, 2002; Scott, Lesh, & Klau, 2002). These studies showed that allowing human operators the ability to change and modify automation-generated path solutions for vehicle routing with time windows afforded better operator understanding. However, with respect to implemented strategies by participants, the authors reported as many strategies as number of subjects, and thus trends were not assessed (Scott, et al., 2002). Aside from this, there is no other prior published research on human-automation sensitivity analysis and hence, there is still much to be learned about how best to support human sensitivity analysis of automated planning solutions.

In order to better understand how humans conduct complex optimizations during the geospatial problem-solving task of path planning under time pressure, an experiment was conducted using a prototype path planner. Human-automated path planning performance was evaluated across different visualizations, increasingly complex path cost functions, and under conditions of degraded automation performance. In addition, path planning performance metrics were defined in order to investigate the role of sensitivity analysis.

3. Methods

3.1. Path Planner Interface

A prototype path planning interface, Planetary Aid for Traversing Humans (PATH), was developed to examine human-path planning performance (Marquez, 2007) in the representative domain of Lunar exploration (Figure 2). Such a setting represents an extreme path planning example, because given limited resources (e.g., oxygen and time) and poor mapping (i.e., Lunar mapping resolution is only $5m^2$ at best), it is critical that the most optimal paths be found across a number of different variables. For future Mars and Lunar exploration, astronaut path planning during extravehicular activities (EVAs), both on foot and in rovers is expected, much like that of previous Apollo missions.

Planning an EVA is a task with a large problem space that includes planetary surface models, physiological models, life support constraints, and other operational constraints. In the Apollo era, astronauts had only paper maps to plan paths, much like present-day hiking maps. They relied on crude estimates of oxygen use, as well as inaccurate paper maps, which led to Lunar spacewalks preemptively shortened because the astronauts could not successfully navigate. Since the majority of future missions will require that astronauts be out of range of mission control, it is critical that automated path planning tools be leveraged since this large problem space is too complex to be solved by a human being.

The PATH interface (Figure 2) permits users to make, modify and submit perceived least-costly paths for geospatial path planning on the Moon. Participants are given a lunar terrain map with start and goal locations (beginning and end of a path) and a critical way-area. Analogous to a lunar exploration site, the critical way-area represents an area of interest for exploration. Once a complete path is constructed, the total path cost (based on a pre-determined cost function) is calculated and displayed in both a tabular and graphical manner on the lower part of Figure 2. In addition, a path cost profile (i.e., cost along the path) and a path elevation profile (i.e., elevations traversed along the path) are shown. Path elevation is a critical consideration since astronauts want to minimize any ascents while walking. Participants modify path solutions by moving, adding, and/or deleting waypoints on the path. The PATH program tracks participant interaction by storing every user mouse-click with corresponding time stamps and path costs. Additional details about the interface design can be found in Marquez (2007).

[INSERT FIGURE 2 HERE]

One common denominator in all path planning domains is the use of a map for generating paths. With advances in graphical display technology, possible map visualizations and associated variables extend far beyond standard topographical representations. Novel visualizations with various forms of fused data could dramatically influence performance, since they could visually represent the constraints depicted in Figure 1, as well as the environment. To this end, PATH supports three different types of visualizations, discussed below, and these were subsequently tested in the experiment.

For all PATH maps, obstacles (i.e., impassable terrain for astronauts in large, bulky suits and large rocks, etc.) are depicted in black. However, terrain that could be negotiated was represented in one of three ways. For the first visualization, typical of current topographical maps currently in use by astronauts-in-training in places like Haughton Crater, contour lines corresponded to changing terrain elevations (Figure 3a). For the second visualization, a color gradient map depicts levels of equal cost (LOEC) (Figure 3b), which will be discussed in more detail below. The third visualization of the PATH map was a combination of the topographical contour lines and the LOEC color gradient map (Figure 3c). Since such a display is envisioned to accompany astronauts on future EVAs, either in laptop form or an arm-mounted display, minimizing the size was a concern.

[INSERT FIGURE 3 HERE]

The LOEC visualization displays integrated cost information about areas on the map with relatively equal cost. The LOEC visualization is based on the calculations generated from the numerical potential field method (NPFM), which is an algorithm that calculates least cost paths (Barraquand, Langlois, & Latombe, 1992; Rimon & Koditschek, 1992). Each location on the map has a corresponding minimum total cost that is relative to the goal. Together, these costs can be considered a minimum total cost field (similar to a potential field), where the goal location is the lowest point (with a zero cost) and obstacles are peaks of high cost. The minimum total cost field is rendered in color, and locations depicted in the same color have the same minimum total cost to the goal. The color gradient ranges from yellow to purple, which represent low to high costs to the goal respectively and are static values.

The NPFM algorithm fundamentally relies on the idea of a “force field metaphor” (Frixione, Vercelli, & Zaccaria, 2001), where path solutions are “attracted” by the goal location and “repelled” by areas of high cost. This underlying metaphor provides human decision makers an intuitive, perceptually-based solution for a large problem space. Consequently, it was hypothesized that PATH users could solve the path optimization problem for complex cost functions more quickly and accurately by using this force field visualization since less mental integration is required as the cost field is directly shown to participants. In the context of the cognitive framework in Figure 1, such a visualization should reduce both the effort associated with mental integration of the costs and goals, as well as lessen the number of iterations in the sensitivity analysis phase.

3.2. Task

An experiment was conducted to investigate how humans conduct complex optimizations in the PATH environment. For training, all participants completed a self-paced PowerPoint® tutorial that explained the PATH tool and all possible factor level combinations. Participants then completed a total of 6 practice and 6 test trials, with each practice occurring before the test session. All sessions were recorded with Camtasia® and post-experiment debriefs were conducted. The experiment took approximately 90 min to complete.

In the training and test sessions, participants were asked to find an obstacle-free, least-costly path with at least one waypoint within the designated way-area, as fast as possible. Participants were free to make as many waypoints as they wished. They were given start and goal locations, a designated critical way-area, the sun’s position for that trial, and a lunar terrain map. Participants were told that, to the greatest degree possible, the sun should be high above the astronaut and at a cross-angle from the direction of travel in order to provide maximum contrast. Sun angle was a significant problem for Apollo astronauts (Carr, Newman, & Hodges, 2003), in that not enough contrast makes it difficult to navigate on the Moon due to the monochromatic setting.

Thirty-four participants volunteered for this experiment, and were compensated for their participation. Their average age was 25.0 ± 3.5 years. Participants were primarily graduate students, with 22 men and 12 women, equally distributed between visualization groups. There were no significant differences between average video-game usage and self-ratings on map use and hiking experience. In addition, all participants were given a map planning test that measures "speed in visually exploring a wide or complicated spatial field" (Ekstrom, French, & Harman, 1979), but there were no statistical differences across all independent variables for this test.

3.3. INDEPENDENT VARIABLES

Three independent variables were varied: visualization type (the three types discussed previously), cost function (four functions, discussed below), and type of scenario (nominal and off-nominal, i.e., degraded, automation). Four increasingly complex cost functions were used in this experiment since the number of variables and the complexity of the cost functions could influence how well a human understands both the problem and the proposed automated solution.

The cost functions used in this experiment (from simplest to most complex) were: 1) Distance: the shortest collision-free distance between the start and goal, 2) Time: the path that took the least time to complete, which depends on both distance and slope, as an astronaut's velocity depends on slope which typically ranged from 0.4 to 1.6 m/s in actual Apollo missions (Waligora & Horrigan, 1975) (Eqn. 1), 3) Metabolic: a combination of distance, time, and terrain slope which determines astronaut energy consumption (Eqn. 2) (Santee, Allison, Blanchard, & Small, 2001), and 4) Exploration (Marquez, 2007): this cost function combines metabolic costs and the Sun Score (SS) (Carr, et al., 2003) as an overall cost function (Eqn. 3). Sun Score is a quantifiable measure of favorable lighting conditions, and relates the sun's position (azimuth and elevation angles) with the observer's direction of travel.

$$T = d/v \tag{1}$$

$$\text{Metabolic Cost} = \text{Energy Rate (ER)}*\text{Time}, \tag{2}$$

$$\text{where } ER = (3.28m+71.1)(.661v(\cos\alpha)+.115)+(w_{p/n}\bullet m\bullet g\bullet v\bullet \sin\alpha)$$

$$\text{Exploration Cost} = \text{Metabolic Cost}*\text{Sun Score (SS)}, \tag{3}$$

$$\text{where } SS = (1+.5((\cos(2\theta)+2)(\cos(2\varphi)+2)))$$

T =time, d = distance, v = astronaut velocity, α = slope, m = mass, g = gravity of moon, w = weight for positive or negative slope, θ = azimuth angle, and φ = elevation angle.

The complexity of the cost functions was included as an independent variable since automated decision support for simple relationships may not be as effective as for more complex ones. The cost functions were increasingly complex in terms of number of variables and mathematical relationships, as each of the equations depends on the proceeding variables. Subject matter experts verified that they perceived path planning problems to be more difficult with increasing number of variables to consider. In addition, reviews of cognitive complexity literature indicate that a primary driver of perceived complexity is the number of entities that must be considered by a human (Edmonds, 1995; Li & Wieringa, 2000; Xing & Manning, 2005). The cost functions were given to the participants along with explanations of the variables in the PowerPoint® tutorial and included detailed equation definitions and illustrations.

The last independent variable, off-nominal vs. nominal automation, was included to examine the impact of off-nominal (i.e., degraded) automation, since the quality of automated path planner solutions depends on the accuracy of the state space representation (i.e., the map) and underlying energy models. In reality, Lunar maps and the energy and environmental models carry significant error likelihood, so flawed environment and cost representations in the cognitive framework in Figure 1 could lead to erroneous operator plans.

Participants were randomly distributed across the three possible visualization groups: elevation contours (EC) visualization, levels of equal cost (LOEC) visualization, or a combined EC/LOEC visualization. The visualization factor was a between-participants factor, but each participant experienced all four cost functions in increasing complexity order. Participants then experienced two off-nominal test sessions for the Time and Exploration cost functions, for a total of six experimental test sessions. All cost functions were not repeated in the off-nominal session in order to minimize testing time. All participants experienced both nominal and off-nominal scenarios, so the complexity and scenario type factors were within participants. Moreover, the two cost functions of Time and Exploration were selected for the off-nominal subset as they represented a relatively simple and complex function.

The cost functions were presented in the order of complexity since each equation relied on the previous one, which was part of the training. In addition, examining the effect of increasing complexity was a specific research question. Moreover, the cost functions were not randomized among participants because a previous pilot study showed a learning effect across functions, such that those participants who experienced the more difficulty functions first performed substantially worse than those who saw the function in increasing complexity order.

In the four nominal test session conditions, participants were told they could rely on the PATH interface to provide them with accurate path costs. In the off-nominal (degraded automation) condition, participants were informed that PATH's cost function models were inaccurate but that the interface still provided useful information. Before the degraded automation condition trials, participants were provided with examples that illustrated the differences between the expected costs (under the nominal scenario) and the observed costs (under the off-nominal scenario), in order to gain an understanding of the approximate distance between the expected and observed costs. In order to maintain scenario consistency between nominal and off-nominal trials, participants repeated the same start and goal locations for the Time and Exploration trials. In order to control for scenario complexity, the maps were rotated between scenarios to mask the fact that the same trial was presented.

3.4. DEPENDENT VARIABLES

Path planning performance was measured by *path cost errors* (path cost relative to least-costly path), as well as *total time to complete a trial path* (both time spent constructing the initial path and subsequently modifying it). This total time was weighted because participants were instructed that this was a time-pressured task, and they were penalized for taking longer than 4 minutes, calibrated in a pilot study. In the pilot, 18 participants were tested with an equivalent task but were instructed to take as much time as they felt necessary. Not under time pressure, 75% of participants finished in 4 minutes or less.

Because we hypothesized that some degree of sensitivity analysis would promote better performance, *time spent modifying a path* was also measured. Lastly, as a measure of participant confidence and sensitivity analysis behavior, *differential cost* was also measured, which was the path cost difference between the final submitted path and the first path constructed. The results are detailed in the next section.

4. Results

Repeated measures analyses of variance were used to analyze the dependent variables (if normality and homogeneity of error variance assumptions were met). Otherwise, appropriate non-parametric tests were used. An alpha of 0.05 for all statistical tests was applied, and the number of participants was selected to achieve an *a priori* power of 0.8. No covariates were used in the subsequent analysis as map planning scores and map and hiking experience were not significantly correlated with any dependent variables. Data greater than 3.29 times the standard deviation were considered outliers (Tabachnick & Fidell, 2001), and consequently removed from the analysis.

4.1. Path Cost Errors

Path cost errors were calculated by normalizing a participant's path cost to the minimum path cost. The range of path cost errors for the six scenarios is shown in Figure 4. There was no statistical difference between visualization groups, but the omnibus test was significant for cost function ($F(5,145) = 216.9$, $p < 0.001$). The nominal Distance function had the smallest path costs errors ($M = 1.03$, $SD = 0.006$). The Time cost function had the largest average path cost errors ($M = 1.26$, $SD = 0.14$), followed by the off-nominal Time function ($M = 1.22$, $SD = 0.20$). The Time functions were statistically no different given a Tukey pairwise comparison ($p = .927$), but these two were significantly higher than all other functions ($p < 0.001$). The off-nominal Exploration path error cost ($M = 1.105$, $SD = 0.006$) was greater than the nominal path Exploration costs ($M = 1.065$, $SD = 0.003$, $p = 0.006$).

[INSERT FIGURE 4 HERE]

Given the degraded performance in both the nominal and-off-nominal Time cost functions, a simple effects analysis revealed that the visualization type for these two cases was a significant factor (Figure 5). When the visualizations are compared between the nominal and off-nominal Time functions, participants' performance significantly improved with the LOEC visualization ($p = 0.005$) and with the mixed LOEC/EC displays ($p = 0.02$), but did not improve with just the EC display ($p =0.18$). This provides some evidence that in the off-nominal case for the Time functions, the LOEC display did help participants. This effect was not seen across any of the other conditions.

[INSERT FIGURE 5 HERE]

4.2. Planning Time Metrics

Two planning time metrics were measured, total planning time and the percent of time used to modify to each plan. Total planning time to complete each task was transformed to weighted time ($e^{t/\tau}$, where *t* is total time and τ equals 4 minutes (the maximum time to plan)). Weighted time was used, as opposed to raw total time, because participants were instructed that there was a penalty for spending too much time on the task, i.e., they were to optimize the path as fast as possible.

The repeated measures ANOVA for total time revealed a significant difference between cost functions ($F(5, 145) = 5.61$, $p < 0.001$), but there were no significant differences across visualizations nor interaction between treatments. Tukey pairwise comparisons were conducted, with results similar to the path cost error results. The Time cost function (M = 2.25, SD = 0.95) took the longest to optimize, which was statistically no different than the off-nominal Time function (M = 2.21, SD = 0.82). In descending order, the Exploration costs, both nominal (M = 2.25, SD = 0.95) and off-nominal (M = 2.25, SD = 0.95), were not statistically different from the Metabolic function (M=1.96, SD = 0.52), and the function with the least time to completion was the distance function (M=1.832, SD = 0.39). Unlike the path cost error metric, there was no significance for any of the displays across the nominal and off-nominal functions so no visualization was superior to another under any condition.

The percent time spent modifying paths was calculated as the ratio between time spent modifying a path to the total time from beginning a trial to path completion. This is an indicator of the amount of sensitivity analysis conducted during path optimization since it reflects that period of time that participants adjusted a path as compared to the creation of the path. A repeated measures ANOVA showed no overall difference across functions or visualization group, and no significant interaction between treatments. Thus, overall participants spent the same amount of time conducting sensitivity analyses.

However, a Helmert contrast test across the Time cost functions (nominal and off-nominal) determined that there was a significant difference between those participants with the LOEC visualization (either alone or combined) and those with the elevation contours ($p = 0.05$). While not part of our original hypothesis, given the surprising results regarding the difficulties with the Time cost functions, this test was warranted. These results indicate that participants with the LOEC visualization, either alone or combined, spent more of their time modifying the Time cost functions than the elevation contour participants. Furthermore, this effect was not seen with the other cost functions (i.e., Distance, Metabolic, and Exploration), and is likely the cause for the increased overall planning times seen in the previous total time analysis.

4.3. DIFFERENTIAL COST

The last performance metric analyzed was differential cost, which is the path cost error difference (in percent) between the first path cost participants generated for a particular trial and the final submitted path cost. This metric assesses how much path costs changed during the sensitivity analysis process. A small positive differential cost indicates that participants' final path attempts were close to their initial paths, which were close to the least-costly path. A large positive differential cost indicates that participants' significantly improved their initial solutions. A negative differential cost indicates that a participant submitted a path that was worse than the first one made. Non-parametric tests were used for analysis because of violations of normality and homogeneity ANOVA assumptions.

As with the other dependent variables, visualization had no significance across the functions, but the functions were significantly different given a Friedman test ($\chi^2(3, N = 34) = 41.37$, $p < 0.001$, Figure 6). All pair-wise comparisons (using Wilcoxon Sign tests) were statistically significant (all p values < 0.05), except between nominal and off-nominal Exploration functions ($p = 0.239$).

In terms of positive differentials, where larger values represent changes that significantly improved initial solutions (Figure 6), participants with the Time (nominal) function made the largest improvements (27.0%), followed by the off-nominal Time function (12.0%). The cost functions with the smallest average differential increases were the Exploration functions (3.0% and 1.5% for nominal and off-nominal respectively).

[INSERT FIGURE 6 HERE]

For the nominal functions, only 2% of all submitted solutions had negative differential costs. However, the Time off-nominal function produced the largest negative differential costs, indicating that a submitted path cost error was higher than the first path cost error. Six participants (18%) generated negative cost paths for the off-nominal Time function and 12 (35%) participants did so for the Exploration off-nominal function. The magnitude of these errors resulted in solutions that were, on average, 17% and 3% worse respectively. Thus, sensitivity analysis helped many participants improve initial paths in off-nominal conditions; however, it also led to significantly worse solutions, particularly problematic for participants in the off-nominal Time condition.

4.4. STRATEGY ANALYSIS

Path planning strategies were analyzed via recorded sessions and post-experiment debriefs. No strategies were briefed in advance, and two general strategies were identified *post hoc*, which were global path regeneration and local sensitivity analysis. Participants who chose to create entirely new paths to compare to the original path were classified as implementing global path regeneration sensitivity analysis, which were approximately one-quarter of the population. Participants who conducted local sensitivity analysis (modified the original path by adding, deleting, and/or moving individual waypoints) constituted three-fourths of the total number of participants.

These strategies can be interpreted in light of Figure 1, the proposed cognitive framework for path planning. Global path regeneration effectively required participants to define an entirely set of new constraints (as seen in the Information box), which meant defining entirely new requirements for the automation. In contrast, local sensitivity analysis only required participants to make slight path modifications to a path they generally felt was correct. These local modifications allowed them to stay focused on the path generation block. Indeed, one result of this study is that the cognitive framework in Figure 1 just represents the local sensitivity analysis strategy. To more accurately reflect the global sensitivity analysis strategy, an arrow would need to go from the “Goals Met?” decision diamond to the “Constraints” block in the Information box.

Even though participants preferred the local strategy as seen in Figure 1 by a large margin, choice of sensitivity analysis strategy was not visualization dependant, nor did it significantly affect performance (with respect to path cost errors, using Mann-Whitney tests). Thus, both strategies were equally effective, even under the automation off-nominal conditions.

5. Discussion

This experiment addressed several questions regarding human interaction with path planners, including the role of a novel visualization, the complexity of the costs functions, and the differences between solutions for nominal versus off-nominal (degraded automation) cases.

The implications of the previously presented results are discussed below, and Table 1 summarizes the major results for the dependent and independent variables. It was hypothesized that path planning performance would decrease with an increase of cost function variables, which represent increasing planning complexity. However, this decline was expected to be lessened by the LOEC visualization, since it provided users with an integrated elevation map and a cost function map that reduced the need for mental calculations.

5.1. PERFORMANCE WITH PERFECT AUTOMATION

For the conditions where the automation performed perfectly due to a correct underlying state space representation and accurate physiologic models, overall the visualization type did not influence path planning performance. In the debrief sessions, participants indicated that they generally applied a shortest distance heuristic regardless of their assigned visualization condition. Given that this heuristic generally followed the solutions of the NPFM, the visualization did not appear to add any additional *perceived* useful information. It remains to be seen if such the LOEC visualization could be useful when an automated solution is proposed that does not match the operator heuristic, e.g., a seemingly random solution from a probabilistic algorithm such as a rapidly expanding randomized tree (RRT). This is currently the subject of active research (Caves, 2010).

In terms of the complexities of the cost functions, Distance was projected to be the easiest to optimize, followed by the Time, Metabolic, and Exploration cost functions. In addition, performance degradation was expected to occur, as manifested by an increase of path costs errors and time spent on the task, as the cost functions grew more complex.

Whereas Distance was generally the easiest to optimize, unexpectedly, Time was the most difficult cost function, as supported consistently through every measure of performance. Participants, regardless of visualization group, had significantly higher path cost errors, weighted solution time (total time), and differential cost (a metric of solution improvement over time) for the Time cost function than the hypothesized more difficult cost functions (Metabolic and Exploration).

[INSERT TABLE 1 HERE]

The performance metrics clearly demonstrate participants generally had difficulty optimizing Time cost function paths, however, these metrics do not directly address *why* this occurred. One reason may be that humans more intuitively understand how to optimize Time, as it is a conceptually easier function to understand, as opposed to the Metabolic and Exploration cost functions, which were considerably more complex. This could lead to participants spending more effort and time optimizing the simpler function. However, if this theory were correct, then similar trends would have been seen for the Distance cost function, which were not.

Poor performance on the Time cost function scenarios is more likely a function of the larger sensitivity and increased saliency of the cost function to changes (i.e., slope), as compared to the other cost functions. This saliency appeared in the cost trend graph beneath the map, such that the costs between the old and new paths were quite different with every path created for the Time function, which was not the case for the other functions. When participants attempted to optimize a Time path, small path modifications resulted in large path cost differentials, which were much more visually salient than for the other functions.

This point is illustrated in Figure 7a, where a relatively small adjustment in a path led to a 26% cost difference between the old and new paths. In contrast, given the exact same original path and an even larger path adjustment, the change only resulted in a 2.6% cost difference, which is barely noticeable in the cost profile. As noted in Equations 1-3, time was calculated in exactly the same way across each. However, the added multiplicative Metabolic and Exploration cost variables made the time variable less of a contributor to the output.

Cost functions that visually represent large path cost changes due to small path modifications can be labeled "overly-salient" or "overly-sensitive" cost functions. Changes in the Time function paths were more prominent in the cost function display, which led to users making more path errors. This problem was also exacerbated by the LOEC visualization in that those participants with this tool spent more time modifying their paths than those with just the elevation contours. Especially in the LOEC conditions, participants could more easily assess the impact of a path modification on the overall path cost (i.e., changes were more visually salient). However, this saliency appears to have made reducing the path cost errors more challenging, despite the fact that participants spent more time in this condition. This implies that under time pressure, humans may have difficulty optimizing cost functions, such as the Time cost function, that are visually salient or sensitive to small changes in variables.

This result further highlights the unexpected influence of a secondary visualization. The cost trend graph was provided across all visualizations to provide participants with situation awareness in terms of how changes could affect possible future paths. While not a subject of investigation in this study, it appears that this visualization had more influence than the map visualization of cost, and how to better display this trend data for different cost functions to promote effective sensitivity analysis is an open research question.

[INSERT FIGURE 7A & B HERE]

5.2. PERFORMANCE WITH IMPERFECT AUTOMATION

In general, performance degraded with imperfect automation, and the results for the off-nominal Time costs function were significantly worse than for the off-nominal Exploration function. For the off-nominal Time function, more path costs errors were made, participants took longer, and also generated the paths with largest negative differential costs. For the off-nominal Time condition, participants generally did not know they had reached a minimum path cost, and spent additional valuable time decreasing their solution quality. In contrast (and unexpectedly), participants with the most complex cost function, the off-nominal Exploration case, performed markedly better than the less complex off-nominal Time function. However, it should be noted that those in the off-nominal Exploration condition made twice as many negative differential path cost errors that those in the off-nominal Time case. However, these errors were far smaller (3% vs. 17% for Exploration and Time respectively), providing more evidence that the Time sensitivity analysis was overly salient.

When compared to the nominal Time function, participants made fewer path cost errors with a LOEC visualization in the off-nominal Time conditions, but also took longer to generate a solution. However, this strategy adjustment was not seen across the nominal and off-nominal Exploration condition. It is likely that since changes were more salient with the Time function and operators knew there was a problem with the automation, they took more time and effort ensuring they had achieved a satisficing, i.e., good enough, solution. Since changes were not as salient with the Exploration condition, operators felt they achieved a good enough solution sooner.

In the post-experiment questionnaire comments, several of the LOEC participants mentioned leveraging the path elevation profiles (the bottom portion of Figure 2) during the off-nominal trials, more than they did during the nominal trials. It may be that the more detailed information of the path elevation profile was a more useful aid for determining changes in path slope, and that participants only integrated this part of the display into their decision-making process when they were wary of the automation. However, it should be noted that the same visual saliency problem in the map also existed for the cost profiles. So even if participants used the cost profiles more in the off-nominal conditions, the large visual changes led to worse performance for the Time cost function, as compared to the Exploration cost function.

6. Conclusion

An experiment was conducted to better understand how humans conduct path planning under time pressure, which investigated four cost functions and three visualizations in the domain of Lunar exploration. The cost functions tested were Distance, Time, Metabolic, and Exploration; each function depended on the previous function, growing in number of variables and complexity. The visualizations included elevation contours (EC), levels of equal cost (LOEC), and a combined visualization (both LOEC and elevation contour lines). Participants were tested in both nominal and degraded automation conditions.

Performance, as defined by errors and time spent planning, was driven primarily by the cost function, though not necessarily by the number of variables manipulated. Unexpectedly, the Time cost function was the most difficult cost function to optimize under both nominal and off-nominal automation conditions, even though this was hypothesized to be a simple function. The Time function depended on only two variables, distance and slope, but was more visually sensitive in that relatively small changes in a path resulted in perceived large path cost changes. These results suggest that if a cost function is visually sensitive to small changes in variables, users could have difficulty in optimizing paths due to an increase in sensitivity analysis, which requires more decision time without a guarantee of improved performance.

When reflecting on these results in the context of the cognitive framework in Figure 1, the inclusion of additional variables such as oxygen use rate and sun angle more accurately represented the environment, and also resulted in better performance despite the additional complexity of the cost function. This suggests that designers of visualizations and algorithms should not overlook details such as cost function variable selection as well as visual saliency of cost functions, as they can significantly impact operator performance in unintended ways. Moreover, these results demonstrate that a more complex cost function is not necessarily predictive of poor human performance.

Lastly, the visualization type did not show any main effects across a variety of dependent variables. The Levels of Equal Cost (LOEC) visualization added additional sensitivity analysis time, but appeared to provide some advantage under off-nominal conditions in terms of reducing path cost error. Furthermore, under degraded automation conditions, some participants adapted their path planning strategies by relying less on the aggregated multivariate visualization and more on a lower-level, more detailed cost profile. This result is important as it suggests that even if operators are given vague information about a decision support tool's ability to generate accurate solutions, they can and will adjust their strategies by seeking additional information. Thus, it is important that designers of such systems consider providing information that enables strategy adaptation, especially when computer-generated solution correctness cannot be guaranteed.

While these results are not reflective of the astronaut population performance (a population not easily accessed), they provide fundamental insights for any application where humans are expected to work with automation in a multivariate path planning scenario. Related applications include unmanned vehicle routing and air traffic control, which are inherently complex settings with potentially large numbers of variables, shifting internal and external constraints, expansive planning spaces, and dynamic goals. These results demonstrate that given appropriate visualizations, humans can reason efficiently and correctly, given both complex multivariate cost functions and imperfect automation. In fact, in some automated path planning cases, simpler problems representations do not guarantee better human performance.

Acknowledgments

We would like to acknowledge the NASA Harriett G. Jenkins Predoctoral Fellowship, the American Association of University Women (AAUW) Dissertation Fellowship, and the Office of Naval Research for sponsoring this research. We would also like to thank Justin Wong, Caroline Lowenthal, and the anonymous reviewers for their contributions.

References

Anderson, D., Anderson, E., Lesh, N., Marks, J., Mirtich, B., Ratajczak, D., et al. (2000). *Human-guided simple search.* Paper presented at AAAI 2000, Austin, TX.

Barraquand, J., Langlois, B., & Latombe, J. C. (1992). Numerical Potential Field Techniques for Robot Path Planning. *IEEE Transactions on Systems, Man, and Cybernetics, 22*(2), 224 - 241.

Carr, C. E., Newman, D. J., & Hodges, K. V. (2003). *Geologic Traverse Planning for Planetary EVA*. Paper presented at the 2003 International Conference on Environmental Systems, Vancouver, BC.

Caves, A. (2010), *Human-Automation Collaborative RRT for UAV Mission Path Planning*, Master of Engineering Thesis, MIT, Cambridge, MA.

Chen, T. L., & Pritchett, A. R. (2001). Development and evaluation of a cockpit decision-aid for emergency trajectory generation. *Journal of Aircraft, 38*(5), 935 - 943.

Cummings, M. L. (2004). *Automation Bias in Intelligent Time Critical Decision Support Systems.* Paper presented at the AIAA 3rd Intelligent Systems Conference, Chicago, IL.

Cummings, M.L., Buchin, M., Carrigan, G., & Donmez, B. (2010) Supporting Intelligent and Trustworthy Maritime Path Planning Decisions, *International Journal of Human Computer Studies*, *68*(10), 616-626.

Edmonds, B. (1995). What is complexity? - The philosophy of complexity per se with application to some examples in evolution. In F. Heylighen & D. Aerts (Eds.), *The Evolution of Complexity*, Kluwer, Dordrecht.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). Cognitive Factors: their identification and replication. In B. Fruchter (Ed.), *Multivariate Behavioral Research Monographs, No. 79-2*. Forth Worth, TX: Society of Multivariate Experimental Psychology.

Fitts, P. M. (1951). *Human engineering for an effective air navigation and traffic control system*. Washington, D.C.: National Research Council.

Frey, H. C., & Patil, S. R. (2002). Identification and review of sensitivity analysis methods. *Risk Analysis, 22*(3), 553 - 578.

Frixione, M., Vercelli, G., & Zaccaria, R. (2001). Diagrammatic reasoning for planning and intelligent control. *Control Systems Magazine, IEEE, 21*(2), 34 - 53.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Graham, S. M., Joshi, A., & Pizlo, Z. (2000). The traveling salesman problem: a hierarchical model. *Memory & Cognition, 28*, 1191 - 1204.

Jimenez, A., Rios-Insua, S., & Mateos, A. (2003). A decision support system for multiattribute utility evaluation based on imprecise assignments. *Decision Support Systems, 36*, 65 - 79.

Johnson, K., Ren, L., Kuchar, J. K., & Oman, C. M. (2002). *Interaction of Automation and Time Pressure in a Route Replanning Task.* Paper presented at the International Conference on Human-Computer Interaction in Aeronautics (HCI-Aero), Cambridge, MA.

Klau, G. W., Lesh, N., Marks, J., Mitzenmacher, M., & Schafer, G. T. (2002). *The HuGS platform: a toolkit for interactive optimization.* Paper presented at the Advanced Visual Interfaces conference, Trento, Italy.

Klein, G. (1989). Recognition-Primed Decisions. *Advances in Man-Machine Research, 5*, 47-92.

Layton, C., Smith, P. J., & McCoy, C. E. (1994). Design of a cooperative problem-solving system for en-route flight planning - an empirical evaluation. *Human Factors, 36*(1), 94 - 116.

Li, K. & Wieringa, P. A. (2000). Understanding Perceived Complexity in Human Supervisory Control. *Cognition, Technology & Work*, *2*, 75-88.

Lu, Y. C., & Mohanty, S. (2001). Sensitivity Analysis of a complex, proposed geological waste disposal system using the Fourier Amplitude Sensitivity Test Method. *Reliability Engineering and System Safety, 72*(3), 275 - 291.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2000). A model of human performance on the traveling salesperson problem. *Memory & Cognition, 28*(7), 1183 - 1190.

Marquez, J. J. (2007). *Human-Automation Collaboration: Decision Support for Lunar and Planetary Exploration.* Unpublished Ph.D., MIT, Cambridge, MA.

Marquez, J. J., & Cummings, M. L. (2008). Design and Evaluation of Path Planning Decision Support for Planetary Surface Exploration. *Journal of Aerospace Computing, Information, and Communication, 5*(3), p. 57-71.

McCarthy, M. A., Burgman, M. A., & Ferson, S. (1995). Sensitivity analysis for models of population viability. *Biological Conservation, 73*(1), 93 - 100.

Pannell, D. J. (1997). Sensitivity analysis of normative economic models: theoretical framework and practical strategies. *Agricultural Economics, 16*, 139 - 152.

Parasuraman, R., & Riley, V. A. (1997). Humans and automation: Use, misuse, disuse, and abuse. *Human Factors, 39*, 230 - 253.

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*, 286 - 297.

Pizlo, Z., & Li, Z. (2003). *Pyramid algorithms as models of human cognition.* Proceedings of SPIE-IS&T Electronic Imaging, Santa Clara, CA.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, 13*(3), 257-266.

Rimon, E., & Koditschek, D. (1992). Exact Robot Navigation Using Artificial Potential Functions. *IEEE Transactions on Robotics and Automation, 8*(5), 501 - 518.

Saltelli, A., & Bolado, R. (1998). An alternative way to compute Fourier Amplitude Sensitivity Test (FAST). *Computational Statistics and Data Analysis, 26*(4), 445 - 460.

Saltelli, A., Chan, K., & Scott, M. (Eds.). (2000). *Sensitivity Analysis*. Chichester, England: John Wiley & Sons.

Santee, W. R., Allison, W. F., Blanchard, L. A., & Small, M. G. (2001). A proposed model for load carriage on sloped terrain. *Aviation, Space, and Environmental Medicine, 72*(6).

Scott, S. D., Lesh, N., & Klau, G. W. (2002). *Investigating human-computer optimization.* Paper presented at CHI, Minneapolis, MN.

Sheridan, T. B. (2002). *Humans and Automation: System Design and Research Issues*. Santa Monica, CA: John Wiley & Sons, Inc.

Simon, H. A., Hogarth, R., Piott, C. R., Raiffa, H., Schelling, K. A., Thaier, R., et al. (1986). *Decision Making and Problem Solving*. Washington D.C.: National Academy Press,.

Smith, P., McCoy, E., & Layton, C. (1997). Brittleness in the design of cooperative problem-solving systems: The effects on user performance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 27*, 360-371.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics*. New York: Harper Collins.

Triantaphyllou, E., & Sanchez, A. (1997). A sensitivity analysis approach for some deterministic multi-criteria decision-making methods. *Decision Sciences, 28*(1), 151 - 194.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

Waligora, J. M., & Horrigan, D. J. (1975). Metabolism and Heat Dissipation during Apollo EVA Periods. In R. S. Johnston, L. F. Dietlein & C. A. Berry (Eds.), *Biomedical Results of Apollo (NASA SP-368),* Washington, DC: NASA.

Xing, J., & Manning, C. A. (2005). *Complexity and automation displays of air traffic control: Literature review and analysis.* Technical Report DOT/FAA/AM-05/4, FAA Office of Aerospace Medicine, Washington, DC.

Figure Captions

Figure 1: A cognitive model for path planning

Figure 2: PATH interface

Figure 3: Experimental visualizations (left to right): elevation contours (EC), levels of equal cost (LOEC), and combined visualization (both elevation contours and LOEC)

Figure 4: Path cost errors across cost functions

Figure 5: Path cost errors for nominal and off-nominal Time cost functions

Figure 6: Differential path cost errors across cost functions

Figure 7: Cost differences for path modifications within cost function, Time (a) and Exploration (b)

Table 1: Summary of Results

		Independent Variable		
Dependent Variable		Visualization (EC, LOEC, Combined)	Cost Function	Comments
Path cost errors	Lowest	No difference	Distance	For off-nominal Time, LOEC & Combined < EC
	Highest		Both Time functions	
Weighted time	Lowest	No difference	Distance	None
	Highest		Both Time functions	
Percent time modifying	Lowest	No difference	No difference	For both Time functions, LOEC & Combined > EC
	Highest			
Differential cost	Lowest	No difference	Both Exploration functions	Off-nominal functions produced high negative costs
	Highest		Both Time functions	

Figures

Figure 1

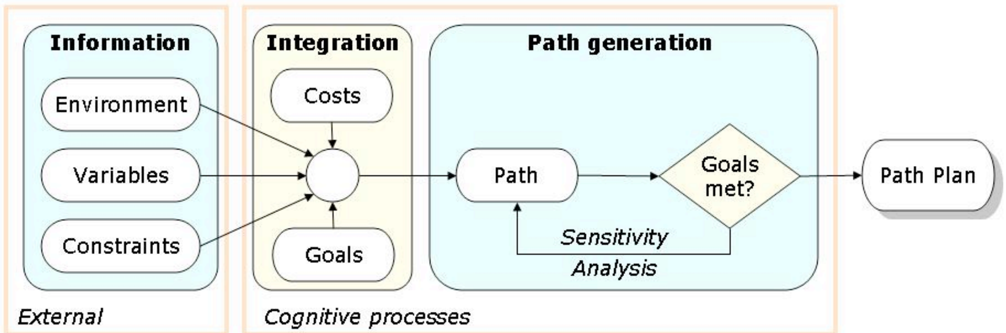


Figure 2 (Color required)

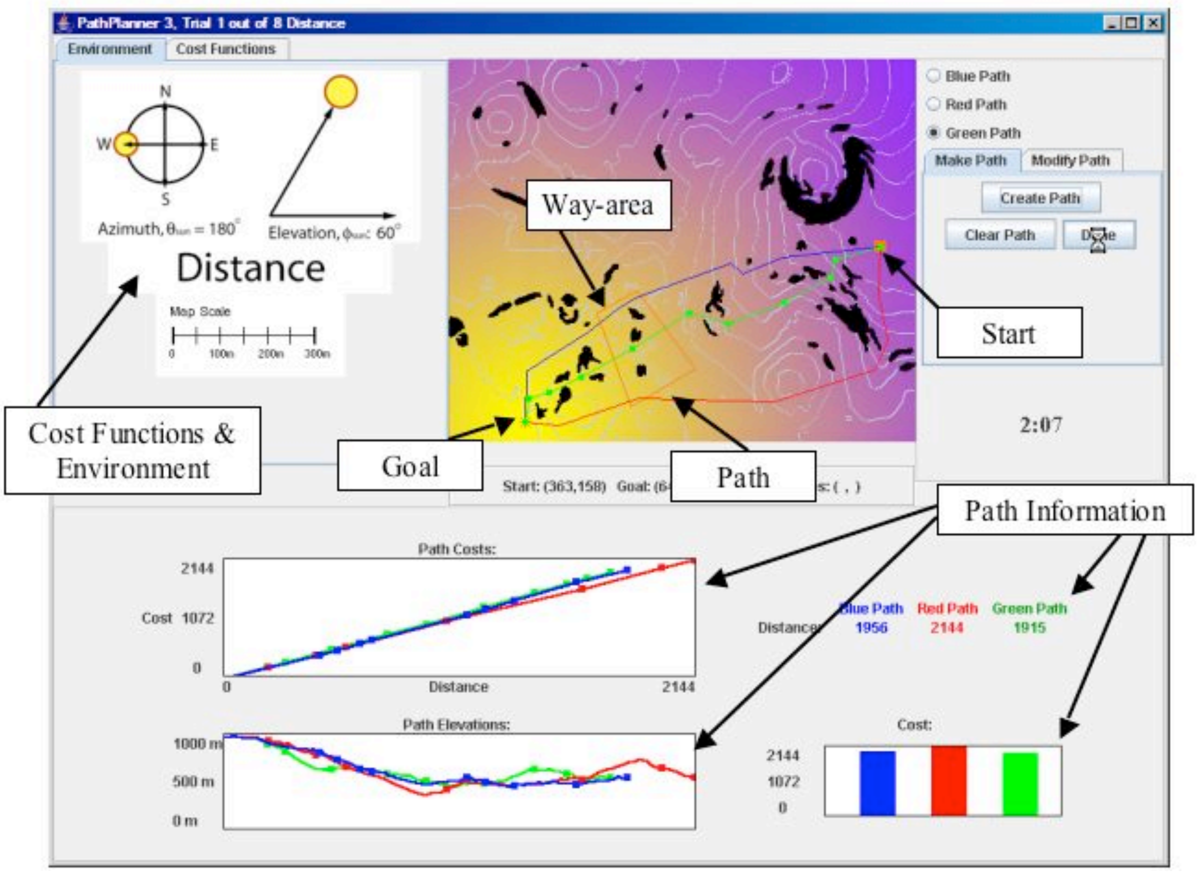


Figure 3 (Color required)

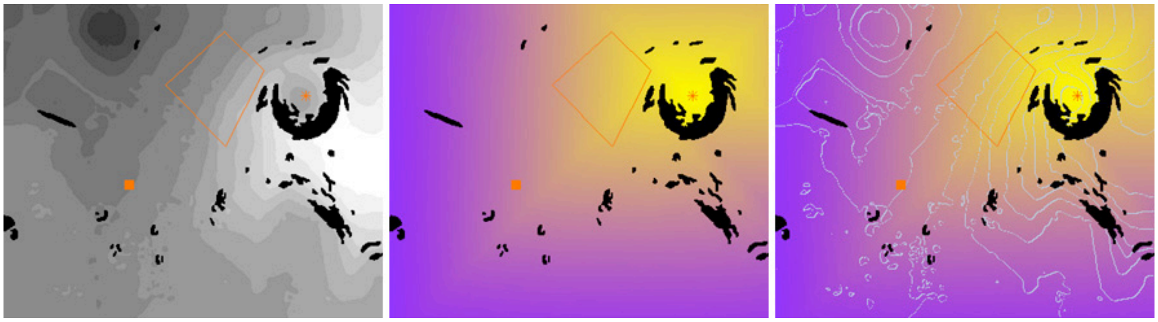


Figure 4

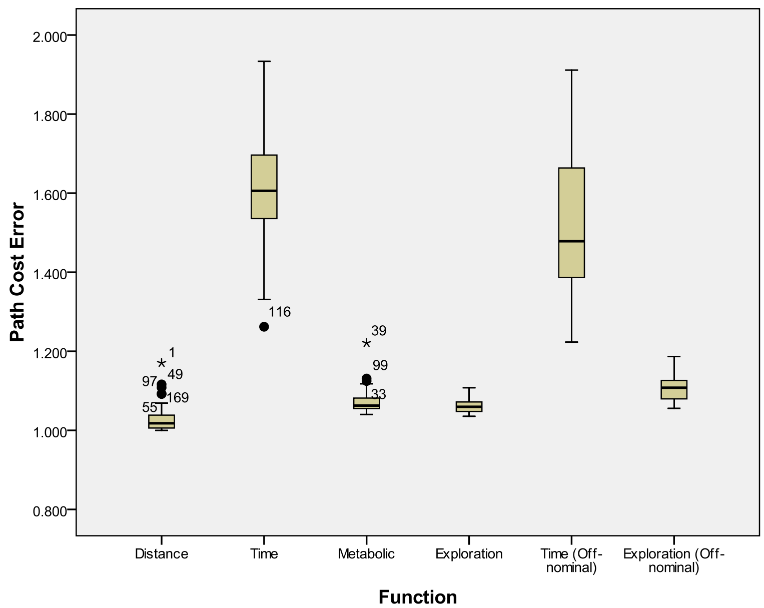


Figure 5

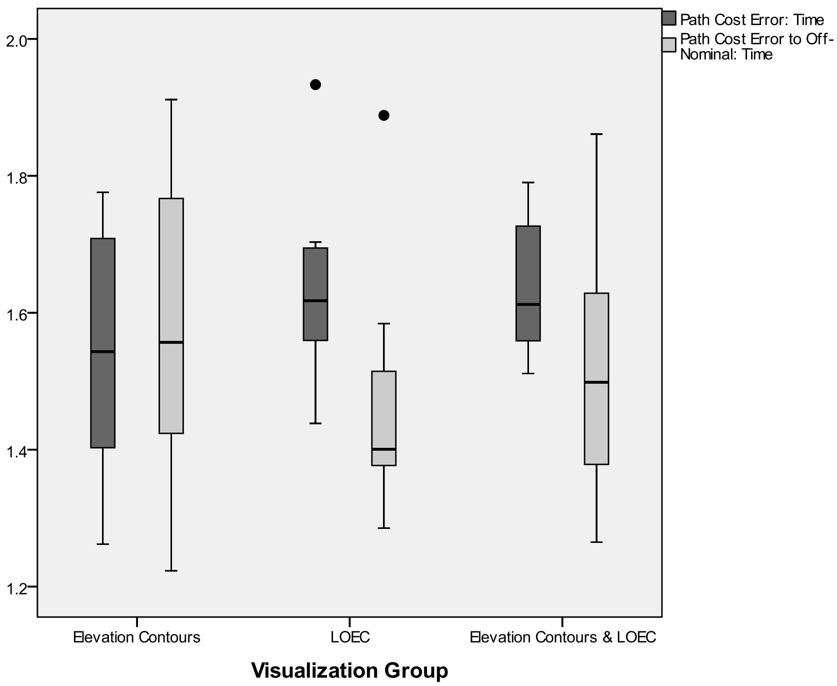


Figure 6

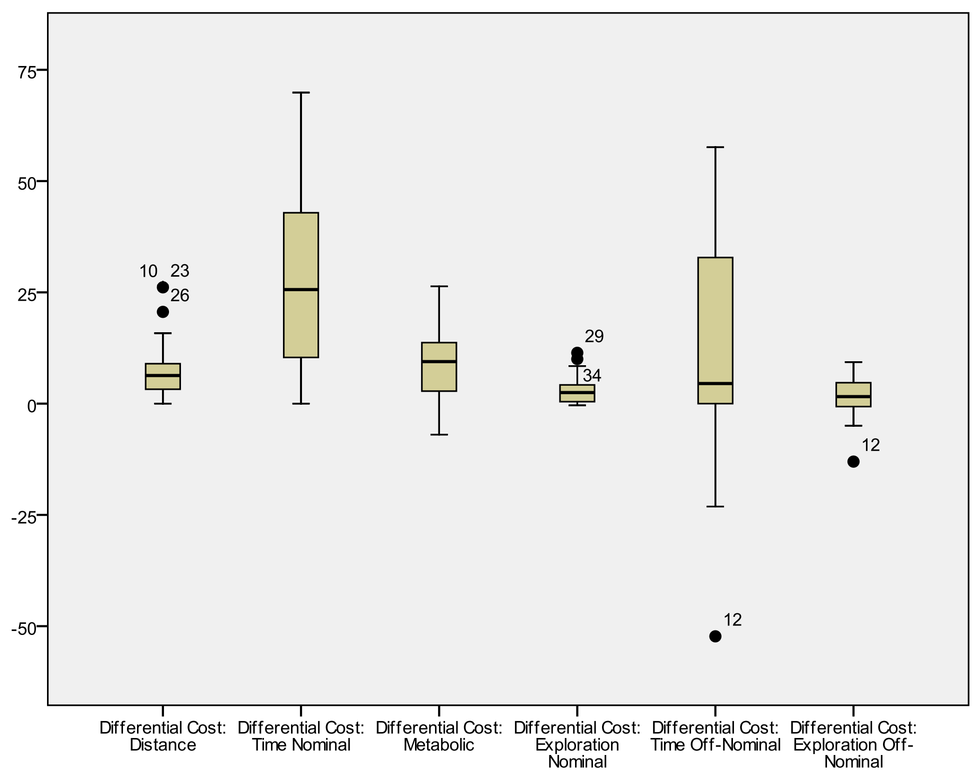
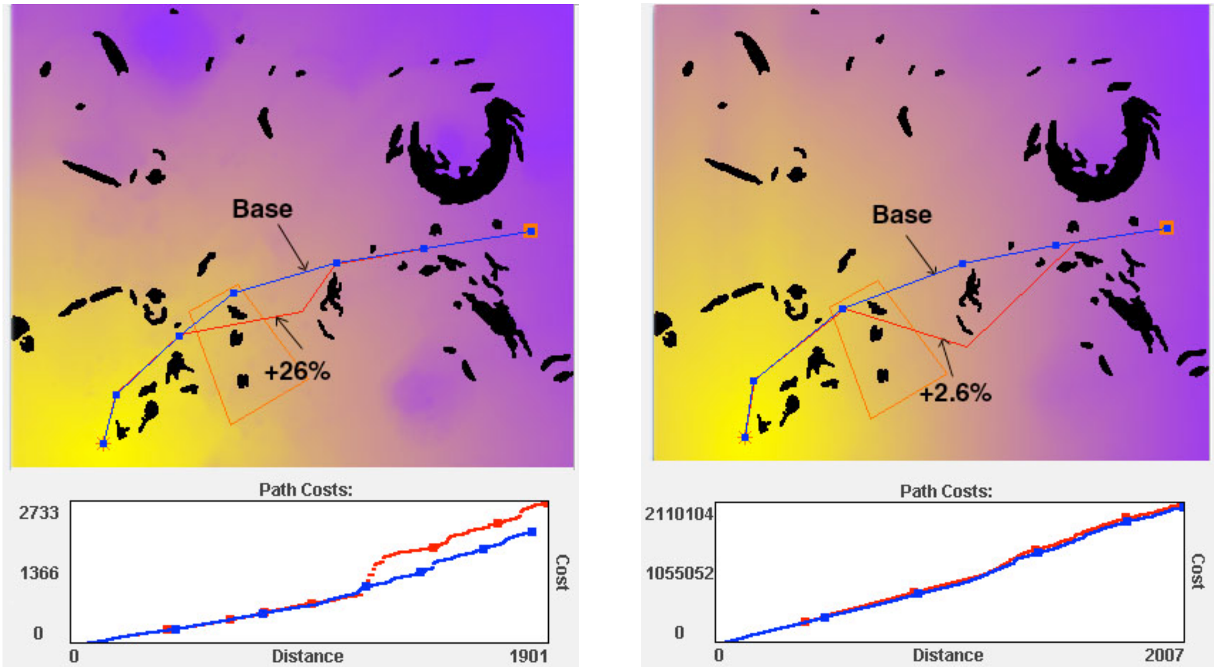


Figure 7a & 7b (Color required)



(a) (b)